

**UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA**



**TRABAJO DE GRADUACION PARA OPTAR AL GRADO DE:
MAESTRO(A) EN ESTADÍSTICA**

TITULO:

APLICACIÓN DE LA ESTADÍSTICA AL ANÁLISIS QUÍMICO

ASESORA: Doctora Carmen Capilla Romá

PRESENTADO POR: Tania Torres Rivera y Mirna Lorena Sorto Alvarez

CIUDAD UNIVERSITARIA, NOVIEMBRE DEL 2003.

UNIVERSIDAD DE EL SALVADOR

AUTORIDADES:

Rectora: Dra. María Isabel Rodríguez

Vicerrector Académico: Ing. Joaquín Orlando Machuca Gómez

Vicerrectora Administrativa: Dra. Carmen Rodríguez de Rivas

Secretaria General: Lic. Margarita Muñoz

Ciudad Universitaria, noviembre del 2003.

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

AUTORIDADES:

Decano en funciones: Msc. Francisco Antonio Chicas Batres

Vicedecano: Msc. Francisco Antonio Chicas Batres

ESCUELA DE MATEMÁTICA

Director en funciones: Lic. Ramón Arístides Paz Sánchez

Coordinador de la Maestría en Estadística: Dr. José Nerys Funes Torres

Ciudad Universitaria, noviembre del 2003.

**UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA**

**TRABAJO DE GRADUACION PARA OPTAR AL GRADO DE:
MAESTRO(A) EN ESTADÍSTICA**

TITULO:

APLICACIÓN DE LA ESTADÍSTICA AL ANÁLISIS QUÍMICO

ASESORA: Doctora Carmen Capilla Romá, Profesora Titular de la Universidad Politécnica de Valencia España, Departamento de Estadística e I. Operativa

PRESENTADO POR: Tania Torres Rivera y Mirna Lorena Sorto Álvarez

DEDICATORIAS

*A la Dra. Carmen Capilla Romá por su invaluable
colaboración y generosidad*

INTRODUCCION

Actualmente, las tendencias tecnológicas, el desarrollo sostenible y el curso de globalización que se están siguiendo en El Salvador, exigen que la perspectiva del uso de la Estadística por parte de los profesionales de las áreas técnicas y de ciencias cambie.

En El Salvador la aplicación de la Estadística, se encuentra todavía en pañales. Si bien es cierto existen muchos buenos profesionales en las áreas técnicas y de ciencias que no ignoran por completo el uso de los métodos estadísticos, estos no tienen la disciplina de su aplicación, porque muchas veces desconocen como o bien a donde se emplea la estadística como herramienta, y por tanto también no perciben el apoyo que el uso correcto de la estadística, sus técnicas y sus innovadores avances pueden brindarle en sus investigaciones y análisis.

Gran parte de la responsabilidad de este hecho radica, en que escasamente se enseña a la estadística desde la perspectiva de su aplicación en las diferentes áreas de desarrollo científico-técnico en el país.

Los profesionales de las ciencias y técnicas químicas no son la excepción en el contexto anterior, y son estos los que practican una de las ciencias más cuantitativas de todas las existentes.

Ante la problemática expuesta, se puede plantear que es imprescindible que a parte de los cursos básicos de probabilidad y estadística que se imparten a nivel de pre-grado, es necesario que se enseñe a los alumnos avanzados así como a profesionales, sobre el uso de la estadística en su campo de trabajo.

Con este trabajo se persigue presentar un material que beneficie al profesional en las áreas químicas, al proporcionarle un condenso de herramientas estadísticas que le permitan diseñar y realizar sus experimentos y mediciones de forma fiable y objetiva. Además se introducen conceptos y procedimientos que son necesarios dentro del contexto de la normalización y certificación de los laboratorios de ensayo y análisis químico siempre desde el punto de vista estadístico.

Entre los aspectos que destacan de este documento se encuentran:

1. Una rica investigación bibliográfica en literatura propia de análisis químico y de Estadística presentada de una forma sencilla. En algunos casos la información se ha

estructurado en forma de cuadros resumen.

2. La resolución y análisis detallado de ejercicios de aplicación práctica para el químico analista sobre los diferentes temas presentados en el documento. Los ejercicios han sido extraídos de la literatura, de documentos técnicos y otros se han creado con datos generados en algunos laboratorios del país.
3. La aplicación de hojas de cálculo como Excel y de programas de calculadoras científicas en la resolución de algunos ejemplos. La resolución en Excel de algunos ejercicios se presenta como archivo anexo, a la cual puede accederse mediante el uso del CD incluido.

ORGANIZACION

El documento se compone de ocho capítulos y cada capítulo está compuesto de tres partes que son: *Primera parte:* Presentación de la teoría base sobre el tema; *Segunda parte:* Presentación de ejercicios resueltos, y *Tercera parte:* Referencias bibliográficas y anexos.

El capítulo 1 va enfocado a lograr que los profesionales en las áreas químicas comprendan la naturaleza aleatoria de sus datos de trabajo dentro de la función del laboratorio de análisis químico, a través de la identificación de los métodos de la estadística univariante y su aplicación a la presentación descriptiva de datos analíticos. En el capítulo 2 se destacan los conceptos de muestras aleatorias, distribuciones muestrales y pruebas de normalidad. En el capítulo 3, se continúa con la estadística inferencial en sus dos procesos: la estimación y el contraste de hipótesis. En este capítulo se hace énfasis en la importancia de la creación de intervalos de confianza para el reporte de resultados provenientes de análisis químico, en el proceso de toma de decisiones basado en datos muestrales para concluir respecto al comportamiento de un sistema de medición química y en la determinación y análisis de la presencia de datos anómalos o "outliers".

En los capítulos 4 y 5, se exponen los principios y aplicaciones al análisis químico de metodologías estadísticas que permiten la realización de experimentos válidos. Entendiendo en sí al proceso de medición química como un experimento se presentan

sobre todo, metodologías estadísticas que permiten estudiar relaciones de dependencia entre una serie de factores que intervienen en los procesos de medición y los resultados de dichos procesos como son la determinación de las concentraciones de analito o bien la cuantificación de alguna variable fisicoquímica, entre estas metodologías se encuentran el diseño de experimentos y el análisis de regresión. En el capítulo 4, se presenta el diseño estadístico de experimentos comenzando con los diseños unifactoriales completamente aleatorizados, y aquellos con restricciones en la aleatorización y finalizando con los diseños factoriales completos y el especial tipo de los diseños factoriales 2^k . El capítulo 5, se refiere al análisis de regresión en sus diferentes tipos lineal y no lineal, presentando aplicaciones específicas de importancia dentro de los procesos de medición química.

En los capítulos 6 y 7, se presenta una perspectiva estadística del aseguramiento de la calidad en los laboratorios de análisis químico, en sus dos componentes la valoración de la calidad y el control de calidad. En el capítulo 6, se presenta el importante tema del muestreo como la llave de la calidad y temas relativos a la valoración de la calidad como el desarrollo y análisis de pruebas de intercomparación de laboratorios. En el capítulo 7, se desarrolla el tema del control estadístico de calidad aplicado a procesos de medición en análisis químico.

Finalmente, en el capítulo 8, se presenta a la estadística no paramétrica como otra área de interés en el análisis químico. En este capítulo se explican y desarrollan métodos no paramétricos de carácter cuantitativo como una herramienta alterna para efectuar contrastes de hipótesis sobre todo cuando no se cumplen los requisitos básicos para llevar a cabo una prueba paramétrica. Este tema ha sido contemplado al final del documento por presentar métodos que no obstante su utilidad carecen de la rigurosidad estadística de los métodos paramétricos que siempre serán los más recomendados.

AGRADECIMIENTOS

Queremos agradecer sinceramente a las personas e instituciones que sin su valiosa colaboración no hubiera sido posible la culminación de este trabajo:

Dra. Carmen Capilla Romá

Lic. Alfredo Aguilar

Coordinación de la maestría de la Escuela de Matemática

Cooperación Universitaria Española (CUES)










Esperamos que con este documento podamos contribuir a cambiar la mentalidad del profesional en química sobre el uso de técnicas estadísticas y que se logre superar el desconocimiento de la importancia de su aplicación.

Tania Torres Rivera

Mirna Sorto Alvarez

APLICACION DE LA ESTADISTICA AL ANALISIS QUIMICO

INDICE

	Pág.
I. Introducción a la Estadística para análisis químico	1-1  capitulo 1.PDF
II. Muestras aleatorias y distribuciones muestrales	2-1  capitulo 2.PDF
III. Inferencia estadística aplicada al análisis químico	3-1  capitulo 3.PDF
IV. Diseño de experimentos aplicado al análisis químico	4-1  capitulo 4.PDF  capitulo 4 Anexos.PDF
V. Análisis de regresión aplicado al análisis químico	5-1  capitulo 5.PDF
VI. Aseguramiento de calidad en laboratorios de análisis químico, una perspectiva estadística	6-1  capitulo 6.PDF
VII. Control de Calidad en laboratorios de análisis químico	7-1  capitulo 7.PDF
VIII. Métodos no paramétricos para análisis químico	8-1  capitulo 8.PDF

CAPITULO I

INTRODUCCION A LA ESTADISTICA PARA ANALISIS QUIMICO

CONTENIDO

	Pág.
1.0 Función del laboratorio de análisis químico	1-2
1.1 Fiabilidad de los resultados en análisis químico	1-3
1.2 Análisis químico y Estadística. Quimiometría	1-5
1.3 Marco conceptual para el estudio de la Quimiometría	1-6
1.4 Variable aleatoria. Población. Clasificación de las poblaciones	1-14
1.5 Descripción de la variabilidad	1-15
1.6 Significado práctico de la desviación estándar	1-32
1.7 La distribución normal	1-33
1.8 Aproximaciones normales	1-37
REFERENCIAS BIBLIOGRÁFICAS	1-39
ANEXO A1. Tabla de la distribución normal estándar	1-40



1.0 FUNCION DEL LABORATORIO DE ANALISIS QUIMICO

La química analítica forma parte inseparable de la mejora de las condiciones de vida y de la preservación del medio ambiente, a causa de su enorme potencial de información sobre sistemas, procesos, productos, materiales y sustancias. Las demandas de análisis cada vez más complejos, exactos, rápidos y competitivos en costos constituyen una necesidad primordial en el mundo globalizado de la actualidad. Para ello, la Química analítica ha de intensificar su evolución. Por una parte, mejorando y aplicando los sistemas de calidad y por otra, incidiendo en los aspectos científicos y técnicos que condicionan la calidad intrínseca de las metodologías de medición, sin la que no puede tener éxito ningún manual de calidad por perfecto que sea.

Como puede intuirse, un factor fundamental para la aplicación efectiva del análisis químico es el conocimiento y capacidad científico-técnica de los gestores y técnicos del laboratorio de ensayo, puesto que ellos definen el proceso analítico a resolver, diseñan el proceso analítico y ajustan sus elementos principales, o los más característicos, de tal modo que los resultados producidos cumplan con los objetivos de la información requerida, con respecto a las leyes, normas o control de calidad de productos, procesos y sistemas.

En general, la función de un laboratorio de ensayo debe involucrar dos procesos secuenciales que son (ver figura 1.1):

- I. El proceso de investigación-medición científico-tecnológico: que puede considerarse como un proceso primario y que involucra aspectos como la generación de metodologías de medición, validación de métodos de análisis químico, cuantificación de la incertidumbre de los resultados y formar parte de equipos de investigación en áreas de ciencias químicas e ingeniería.

- II. El proceso de medición por servicio: que puede considerarse un proceso de carácter secundario y cuya calidad depende de la efectividad del primer proceso.

Estos dos procesos en la función del laboratorio, se manejan con frecuencia inconsistentemente, produciéndose, a veces, confusión en el trabajo y en la comunicación con los clientes o usuarios del servicio.

Las etapas generales involucradas en los procesos de la función del laboratorio de ensayo son las siguientes (GARCÍA, 1998, págs. 45-51):

1. Planteamiento del problema analítico.
2. Gestión de acceso y ejecución del muestreo.
3. Estabilidad, conservación y tratamiento previo de las muestras.
4. Materiales de referencia y preparación de patrones.
5. Calibración de instrumentos.
6. Selección y aplicación de las condiciones experimentales.
7. Adquisición y tratamiento de datos.

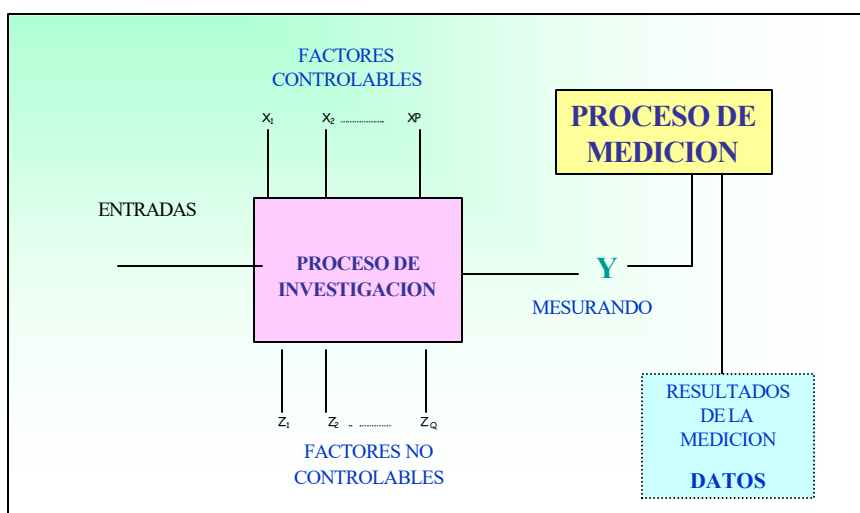


FIGURA 1.1 FUNCION DEL LABORATORIO DE ENSAYO

1.1 FIABILIDAD DE LOS RESULTADOS EN ANÁLISIS QUÍMICO

Muchas decisiones importantes están basadas sobre los resultados de análisis químicos cuantitativos. Los resultados se usan por ejemplo, para estimar rendimientos, comparar materiales contra especificaciones o límites establecidos, o estimar valores monetarios. Es importante, por tanto, tener indicadores de la calidad de los resultados, esto es, la extensión en la que ellos pueden ser confiables para lograr los propósitos buscados. Usualmente, los indicadores de calidad de los resultados, se refieren a ciertos requisitos

establecidos por el usuario y que se denominan *Parámetros de Calidad o Características significativas del método*. En general, estos pueden ser de tres tipos (RAMIS, 1997):

- a) **Estadísticos**: como la exactitud, representatividad, sensibilidad, selectividad, incertidumbre, etc.
- b) **Operativos**: como rapidez, requerimientos energéticos, capacidad para el trabajo de campo, complejidad en la preparación de la muestra analítica, capacidad para el microanálisis, etc.
- c) **Económicos**: como inversión inicial, costo de mantenimiento, preparación del personal, etc.

La calidad en los resultados, puede verse afectada, por múltiples factores externos e internos al laboratorio, y en especial de los siguientes (GARCIA, 1998):

- Adecuación de la metodología a la información que se desea obtener.
- Equipamiento del laboratorio y calidad de las calibraciones.
- Calidad de los patrones y materiales de referencia.
- Representatividad de la muestra.
- Experiencia profesional del personal.
- Tratamiento estadístico de resultados, aplicando los medios más convenientes.
- Organización del laboratorio (sistema de gestión de calidad real aplicado).

La fiabilidad o confiabilidad de las medidas, en análisis químico, debe ser un objetivo prioritario para el proceso de medición en el laboratorio de ensayo. Sin embargo, aunque esta confiabilidad se traduzca, en primera instancia como:

“calidad de las medidas”, no debe dejarse de lado el hecho de que ésta, depende fundamentalmente, de la calidad del proceso analítico, la cual depende a su vez de la efectividad del sistema de calidad implementado en el laboratorio de análisis químico. De acuerdo a lo anterior, es importante, que para el alcance de la “calidad total” se cuente con una clara comprensión de los parámetros de calidad así como de su relación con el sistema de calidad en el laboratorio de ensayo.

El sistema de calidad total, para el laboratorio de análisis químico, se puede definir como el conjunto interrelacionado de planes, políticas, procesos, procedimientos, personal y

tecnología que se necesitan para cumplir con los objetivos del laboratorio. Este sistema, comprende diferentes elementos que son: *el aseguramiento de la calidad, control de calidad y la evaluación de la calidad*. Los capítulos 6 y 7 del texto comprenden esta temática.

1.2 ANÁLISIS QUÍMICO Y ESTADÍSTICA. QUIMIOMETRÍA

“La aportación a largo plazo de la estadística depende no tanto de obtener gran cantidad de estadísticos con estudios profundos en la industria, como lo es crear una gran generación de físicos, químicos, ingenieros y otros que tengan conocimientos de la estadística y que en una u otra forma puedan ayudar al establecimiento y dirección de los procesos de producción del mañana”.

Shewhart y Deming

Los procesos de investigación y medición en análisis químico constituyen dos de los desarrollos más cuantitativos en el área de la química.

La estadística interviene en estos procesos a través del suministro de un conjunto de herramientas que al ser usadas adecuadamente, conducen a la obtención de resultados confiables. En análisis químico, la ciencia que aplica los conocimientos estadísticos y otros conocimientos de las matemáticas y de las ciencias de la computación al análisis de los sistemas de su interés se conoce como *quimiometría*. En gran medida, la *quimiometría* es estadística, aunque con unas características que la diferencian claramente de la estadística aplicada a otros campos del saber. El objetivo general de ésta, es extraer en forma óptima la máxima cantidad de información relevante acerca de los sistemas naturales, utilizando para ello datos químicos, en forma exclusiva, o en combinación con datos de cualquier otra índole que aporten información relacionada con el problema que se desea resolver (RAMIS, 1997).

Dado que la *quimiometría* permite potenciar y aumentar el rendimiento de un proceso de análisis químico, esta debe actuar sobre todas y cada una de las etapas de dicho proceso, es decir planeación, recolección, análisis y control. Así, por ejemplo, antes de obtener los datos, la aplicación de los métodos quimiométricos permite el diseño experimental, de modo que se obtenga la información requerida con el mínimo de recursos. Otras, técnicas

quimiométricas aumentan la eficacia del proceso de adquisición de datos. Finalmente, una vez obtenidos los datos se trata de extraer de los mismos la mayor cantidad de información posible, y establecer la confiabilidad de la información obtenida.

Entre las fuentes de dificultad más generales a las que se enfrenta el químico analista y que lo conducen de forma inevitable al uso de la estadística se encuentran:

1. La variabilidad de los resultados obtenidos de la aplicación de los procesos de medición y/o investigación.
2. La necesidad de determinar los efectos de los factores controlables y no controlables sobre los resultados obtenidos, y la complejidad de las relaciones entre los mismos y la respuesta.
3. La necesidad de establecer la confiabilidad de los resultados obtenidos del proceso de investigación y/o medición.

1.3 MARCO CONCEPTUAL PARA EL ESTUDIO DE LA QUIMIOMETRIA

La determinación del sistema analítico implica en general, describir las etapas o procesos involucrados, definir las características o propiedades de importancia para los mismos, y clasificar los factores que influyen sobre los procesos así como sus efectos. La estadística proporciona una serie de herramientas que auxilian cada uno de los aspectos anteriores, tal y como se presenta en la tabla 1.1. Sin embargo, antes de entrar en el estudio de dichas herramientas es necesario definir conceptos como los siguientes:

- a) Objetos o unidades experimentales.
- b) Variables y datos.
- c) Medición.
- d) Poblaciones y variabilidad.

En las secciones de la 1.3.1 a la 1.3.5 se presenta una descripción detallada de dichos conceptos.

1.3.1 OBJETOS O UNIDADES EXPERIMENTALES

En quimiometría los *objetos* son los distintos entes materiales sobre los cuales se quiere obtener información como espacios naturales en análisis ambiental, partidas de materia prima o de producción en análisis industrial, análisis de fluidos fisiológicos en análisis clínico y veterinario, etc. Los objetos vendrán representados por las muestras, utilizando la palabra muestra en el sentido de química analítica y no en el estadístico, esto es que la muestra es una parte representativa del objeto de análisis. En estadística los objetos también se llaman unidades experimentales.

TABLA 1.1 RELACION ENTRE LAS ETAPAS DEL PROCESO ANALITICO Y LAS HERRAMIENTAS ESTADÍSTICAS.

Etapas del proceso de análisis químico	Técnicas estadísticas aplicables
<i>Problema de análisis</i>	Selección de variables Estadística del muestreo
<i>Tratamiento de la muestra:</i> -Gestión de acceso y ejecución del muestreo -Estabilidad, conservación y tratamiento previo de las muestras	Selección de factores y variables Diseño experimental Análisis de regresión Optimización
<i>Medición:</i> -Materiales de referencia y preparación de patrones -Calibración de instrumentos -Selección y aplicación de las condiciones experimentales	Selección de factores y variables Diseño experimental Análisis de regresión Optimización
<i>Resultados:</i> Adquisición y tratamiento de datos	Análisis exploratorio de datos Procesado de datos uni y multivariable Estimación de intervalo Pruebas de significación Análisis de regresión Análisis multivariable

1.3.2 CARACTERÍSTICAS Y VARIABLES

Buscar resultados confiables como salidas de los procesos del sistema del laboratorio de análisis químico, significa, en primera instancia, la determinación y cuantificación de sus *características o propiedades descriptoras* y de sus relaciones con los factores que afectan dichos procesos.

Cuando estas características descriptoras se expresan numéricamente son llamadas *variables, datos o resultados*. En análisis químico las variables pueden ser de carácter químico (ej.: concentraciones), fisicoquímico y físico (ej.: ph, conductividad, etc).

Por otra parte, también se puede disponer de información adicional a la información propiamente de química analítica, como costos económicos por unidad de producción, estimación del grado de aceptación del producto por el consumidor, calidad organoléptica del agua potable, etc. Todas estas variables pueden contener información que se debe tratar de aprovechar conjuntamente con la información aportada por los datos químico-analíticos. Es importante hacer notar que los datos solo serán útiles si su determinación surge a la luz de un proceso sistemático de búsqueda de la información necesaria (proceso analítico).

1.3.3 MEDICIÓN Y ESCALAS DE MEDICIÓN

La *medición* puede definirse como la asignación de números a las características de los objetos o sucesos de acuerdo a ciertas reglas. En el proceso de medición se desarrolla una relación entre el sistema empírico (estudio de cosas físicas, sociedades y personas) y el sistema abstracto (sistemas numéricos), siendo el resultado del mismo conocido como *medida* (KINNEAR y TAYLOR, 1998, cap. 7). En la figura 1.2, se presenta el esquema del proceso de medición.

La definición de medición sugiere que las características de los objetos o eventos son las que se miden y no los objetos o eventos en sí. Por ejemplo, se puede medir la dureza, la densidad, el contenido de sulfatos, la temperatura, etc en una muestra del agua potable.

Es importante destacar que solo es posible una medición efectiva cuando las relaciones entre las propiedades de los objetos o sucesos en el sistema empírico se relacionan directamente con el sistema numérico. Si esta relación es tergiversada se presenta el error de medición (KINNEAR y TAYLOR, 1998, cap. 7).

En la definición de medición, el término "número" impone ciertas restricciones sobre el tipo de manipulaciones numéricas admisibles (KINNEAR y TAYLOR, 1998, cap. 7).

La naturaleza de las relaciones existentes en el sistema empírico determina el tipo de manipulaciones numéricas que son válidas en el sistema abstracto. Hay una gran tentación de emplear todas las características del sistema numérico en el análisis de datos y hacer caso omiso de las restricciones impuestas por los fenómenos empíricos en estudio (KINNEAR y TAYLOR, 1998, cap. 7).

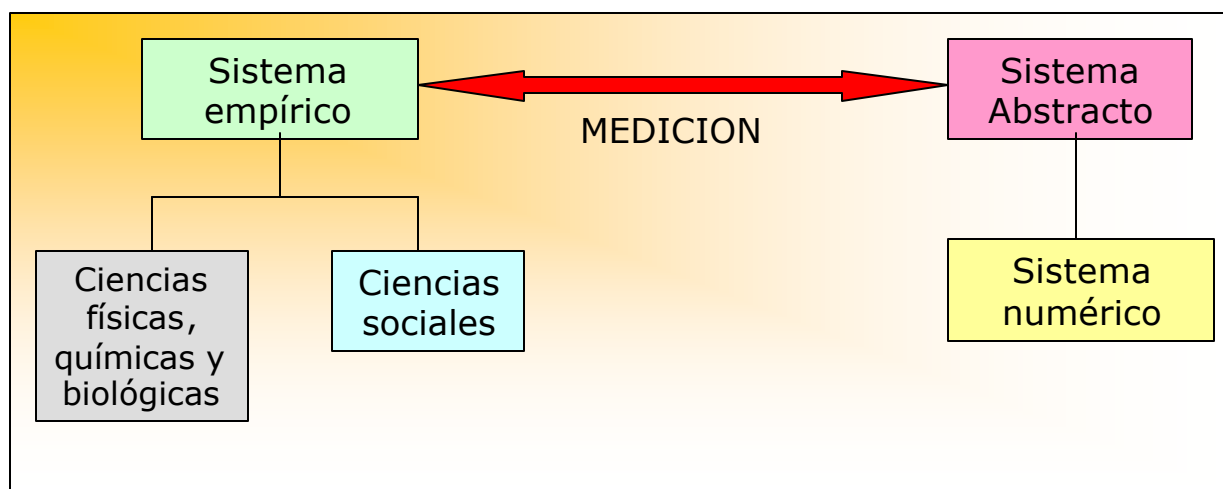


FIGURA 1.2 PROCESO DE MEDICION (KINNEAR y TAYLOR, 1998)

En general, se habla de cuatro características del sistema numérico (KINNEAR y TAYLOR, 1998, cap. 7):

1. Cada número de la serie (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) es único y hay diez.
2. La ordenación de los números es convencional.
3. Se pueden precisar sustracciones equivalentes.
4. Se pueden definir cocientes equivalentes.

Las escalas de medición se han clasificado en dos tipos generales, en términos de las cuatro características del sistema numérico, estas son (KINNEAR y TAYLOR, 1998, cap. 7):

- I. Escalas no métricas: que a su vez se clasifican en escalas *nominales* y *ordinales*.

Las medidas con una escala nominal, también conocida como escala de categoría, es una escala en la que el número cumple únicamente con la propiedad uno del sistema numérico y proporciona el número de ocurrencias en cada clase o categoría de la variable que se está estudiando. Por lo tanto los números asignados a los objetos no tienen más significado cuantitativo que indicar la ausencia o presencia de una característica o atributo bajo investigación. Un ejemplo de medición en esta escala es identificar cualitativamente la presencia o no de un componente químico en una muestra.

Las escalas ordinales representan un nivel superior de precisión en la medida, con el cumplimiento de las propiedades uno y dos. Una escala ordinal define la relación ordenada entre los objetos o sucesos. Cada subclase puede ser comparada con otra en términos de una relación "mayor que" o "menor que". Los números utilizados en escalas como las ordinales no son cuantitativos, dado que indican posiciones relativas en series ordenadas. Por ejemplo la cuantificación organoléptica de un agua purificada.

- II. Escalas métricas: Estas a su vez se clasifican en escalas *de intervalo* y *de razón*. Ambas proporcionan el nivel más alto de precisión en la medida, permitiendo realizar todas las operaciones matemáticas.

Una escala de intervalo comprende la utilización de números para clasificar las características de los objetos o sucesos, de manera que las distancias entre los numerales correspondan a las distancias entre los mismos con relación a lo que se está midiendo. Las escalas de intervalo poseen todas los requisitos de una escala ordinal, más la característica de "igualdad de diferencia" del sistema numérico. En una escala de intervalo existe una asignación arbitraria del punto "cero", y éste no indica la ausencia absoluta de la propiedad en la muestra. Ejemplos de mediciones en esta escala son las medidas de temperatura en grados Celsius o Fahrenheit.

Una escala de razón tiene todas las propiedades de una escala de intervalo más un punto cero absoluto. En la medición de razón sólo se puede asignar arbitrariamente un número, es decir, la unidad de medición o distancia. Una vez determinado éste número, las asignaciones numéricas restantes están completamente determinadas. El punto cero absoluto o natural se refiere a la asignación del número cero a la

ausencia de la característica que se está midiendo. Un ejemplo de medición en esta escala es la concentración de un analito en una muestra.

Como puede observarse la diferencia real entre las escalas de intervalo y de razón es que las de intervalo tienen un punto cero arbitrario, mientras que las escalas de razón tienen un punto de cero absoluto (KINNEAR y TAYLOR, 1998, cap. 7).

En el esquema de la figura 1.3 se presenta la relación entre el tipo de escala y las propiedades del sistema numérico para el manejo de datos.

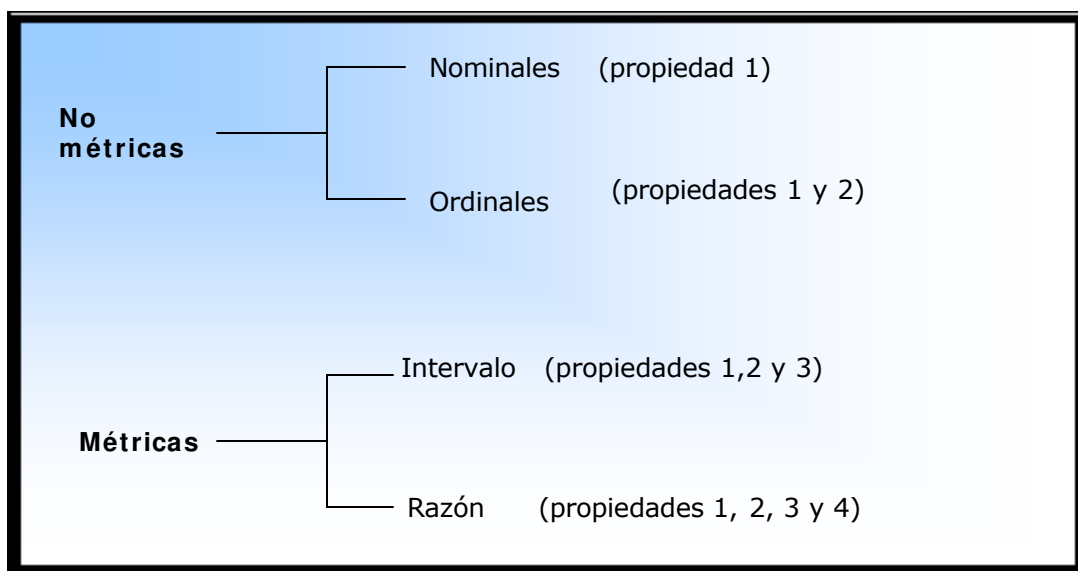


FIGURA 1.3 RELACION ENTRE ESCALAS Y PROPIEDADES DEL SISTEMA NUMERICO

1.3.4 VARIABILIDAD

Considere el proceso de preparación de un refresco de frutas por un restaurante (que es famoso por este refresco). En principio, cada vez que se prepara, se efectúa una repetición de ciertos pasos del proceso: la selección y corte de la fruta, el pesado de la fruta picada, la medición del volumen de agua, el pesado del azúcar, el licuado de la fruta y por último el mezclado de los materiales. La evidencia indica que a pesar de que se intente repetir todos los pasos en forma idéntica, el resultado no es el mismo todas las veces que se produzca el refresco. Esta variación en el "producto" (el refresco, en este caso) recibe el nombre de *variabilidad* y está presente en todo proceso real, de modo que

no se puede predecir con exactitud como “quedará el refresco” antes de que se fabrique. Ello no significa que la *variabilidad* no se pueda medir. En este ejemplo si puede ser posible “medir” la dulzura con que acostumbra quedar el refresco, o bien la apariencia que suele tener el mismo.

En general, los resultados de cualquier trabajo que se realice contienen variación y su distribución sigue un cierto patrón. La *variabilidad* es un fenómeno de la naturaleza y un hecho en el entorno de cualquier proceso de producción¹. Los datos (medidas) que se obtienen siempre están “dispersos” y casi nunca tienen un único valor constante. Esta “dispersión” en los datos, es causada por un número infinito de factores controlables o no controlables que afectan el proceso. Ver figura 1.1.

Si se ignora la existencia de la *variabilidad*, o si se racionaliza en forma falsa que es pequeña, se pueden llegar a tomar *decisiones incorrectas sobre problemas importantes* (PEREZ, 1999, p. 79).

La toma de decisiones en presencia de “ruido”, conduce al uso de *la estadística* como ciencia especializada, ya que esta ayuda a analizar los datos en forma adecuada y a obtener conclusiones tomando en cuenta “lo inevitable”, la existencia de la *variación o dispersión*.

Muchas veces se dice que “*La estadística es al químico analista, lo que un bisturí es a un médico cirujano*”.

1.4 VARIABLE ALEATORIA. POBLACION. CLASIFICACION DE LAS POBLACIONES

Los métodos estadísticos se aplican a fenómenos en cuya toma de decisiones prevalecen las condiciones de incertidumbre y hay *variabilidad* (tal como se estableció en la sección 1.3.4). Esta incertidumbre plantea al químico analista la necesidad de medirla, siendo la medida de la incertidumbre la *probabilidad*.

Asociado al concepto de *probabilidad* está el concepto de *variable aleatoria*. Una *variable aleatoria* representa el conjunto de valores que pueden observarse en un fenómeno aleatorio (tal como lo es el proceso de medición), valores que dependen del azar y sobre

¹ “proceso de producción”, cualquier proceso que genere un bien de manufactura, servicio o software. De acuerdo a esto el proceso de “medición”, es en sí, también un proceso de producción.

los cuales es posible establecer una medida de su probabilidad. Según la escala de valores que pueda tomar la *variable aleatoria* esta puede ser de dos tipos *discreta* y *continua*. Una *variable aleatoria discreta* sólo puede tener ciertos valores claramente separados, que resultan de contar algún elemento de interés, en este sentido son conteos de números pertenecientes a escalas nominales. Una *variable aleatoria continua* es aquella que puede adquirir cualquier número real, en este sentido pertenecen a resultados en escalas de tipo métricas. Al menos teóricamente (aunque no en la práctica), la variable continua puede expresarse con un número infinito de valores decimales.

En general, en *estadística* se entenderá como *población* al conjunto de todos los individuos o elementos que constituyen el objeto de un determinado estudio y sobre los que se desea obtener ciertas conclusiones. Para la *quimiometría* la *población* estará constituida por el conjunto de toda la materia sobre la que se aplica un método de análisis químico de una forma específica.

Dado que el proceso de medición es un fenómeno aleatorio puede decirse que en *toda población que provenga de este proceso existirá variabilidad y que sus resultados por tanto son variables aleatorias*.

Las poblaciones pueden clasificarse en función del número y tipo de elementos que la constituyen dividiéndose, en poblaciones *finitas e infinitas* en base al número de elementos, y en poblaciones *reales y conceptuales* en base al tipo de elementos. Las *poblaciones finitas* son aquellas que tienen un número contable de elementos mientras que las *poblaciones infinitas* son las que no tienen un número contable o "concebible" de elementos (ROMERO y ZUNICA, 1993, cap. 1).

Si los elementos que contienen las características en estudio tienen una existencia real, previa a la realización del estudio las poblaciones son de tipo *real*. Por otra parte, si los elementos que contienen a las características son generados por la realización de un determinado proceso las poblaciones son de tipo *conceptual*. A estos procesos, que en sucesivas realizaciones pueden ir generando los diferentes elementos de la población se les denomina *experimentos aleatorios* (ROMERO y ZUNICA, 1993, cap. 1). En este sentido el análisis químico es un experimento aleatorio.

Como ejemplos de poblaciones de carácter finito y real pueden mencionarse: la determinación de desnutrición de los niños de una comunidad, las características de los

lotes o partidas de producción que se estudian para la aceptación o rechazo de las mismas, etc.

Normalmente, en análisis químico, las poblaciones se consideran de carácter *infinito* y *conceptual*.

Cuando sobre cada resultado de la población se estudian k características diferentes (todas ellas expresadas numéricamente) se obtiene una variable aleatoria k -dimensional (ROMERO y ZUNICA, 1993, cap. 1). Por ejemplo, si en una porción de suelo se le lee su contenido de hierro, su pH , su densidad, se tiene una variable aleatoria de dimensión 3. Es importante, distinguir, entre una variable aleatoria k -dimensional, en que k variables se miden sobre los elementos de una población y un conjunto de k variables aleatorias unidimensionales, definidas sobre k poblaciones distintas (ROMERO y ZUNICA, 1993, cap. 1).

1.4.1 QUI M I O M E T R I A Y E S T A D I S T I C A U N I V A R I A B L E Y M U L T I V A R I A B L E

En realidad ningún problema en análisis químico es intrínsecamente univariado. Por el contrario, un problema siempre puede ser descrito por varias variables. En algunos casos será necesaria una sola variable (ejemplo: conductividad eléctrica para establecer la concentración salina de una muestra de agua). En este caso el problema deberá resolverse en base a la información univariante. Sin embargo, en otros casos no será conveniente ni posible resolver el problema sin recurrir a dos, tres o más variables: concentraciones de ión cloruro, de ión sodio, de ión sulfato, conductividad, densidad, etc. para establecer la calidad de un producto para una determinada aplicación industrial, o para cualquier otro problema que se quiera resolver (RAMIS, 1997).

Cuando se dispone de una sola variable los métodos estadísticos que habrá que aplicar serán más sencillos que en el caso en que se disponga de más variables. Se habla entonces de estadística *univariante*, y de *métodos univariantes*. Si se dispone de dos o más variables, se hablará de estadística multivariada y se aplicarán métodos multivariados. En este documento se hará énfasis sobre los métodos univariados de análisis de datos (RAMIS, 1997).

1.4.2 RECURSOS INFORMÁTICOS EN QUIMIOMETRÍA

Actualmente, para trabajar en quimiometría se utilizan recursos informáticos de diversa procedencia, como son:

- a) *Paquetes estadísticos multipropósito:* como el STATGRAPHICS, SPSS, MINITAB, SAS, etc.
- b) *Entornos asistidos para matemáticas y estadística:* como el MatLab, Visual Basic, etc.
- c) *Hojas de cálculo:* como Excel, Lotus, etc.
- d) *Paquetes diseñados especialmente para quimiometría:* como PARVUS, UNSCRAMBLER, etc.

El uso de cualquiera de estos recursos dependerá de la disponibilidad del mismo, así como de los requerimientos por parte del químico analista.

1.5 DESCRIPCIÓN DE LA VARIABILIDAD.

Una vez que se ha establecido que los datos relativos a los elementos de la población van a estar dispersos, surge el problema de detectar el patrón en la variabilidad de los datos. La pauta de variabilidad de una *variable aleatoria* puede describirse de tres formas que son:

- Forma gráfica.
- Mediante expresiones cuantitativas de la distribución.
- De forma sintética recurriendo a un tipo de modelo matemático.

1.5.1 DESCRIPCIÓN GRÁFICA DE LA VARIABILIDAD

A menudo se utilizan métodos gráficos simples para facilitar el análisis de los datos. Entre estos están (MONTGOMERY, 1991, cap.2):

- a) **Diagramas de puntos:** los diagramas de puntos son útiles para representar una serie pequeña de datos (menos de 15 datos). Tal y como se presenta en la figura 1.4, el diagrama de puntos permite ver rápidamente la localización general o tendencia central de las medidas y su dispersión. Para construir este tipo de gráficos se procede de la siguiente forma: a) se traza una línea horizontal en el que se marcan la escala de valores de la variable aleatoria y b) se colocan todos los datos en la gráfica, identificándolos con puntos, uno por cada dato. En el caso de que los valores de dos datos sean similares o iguales se colocan uno encima del otro, esto es verticalmente.

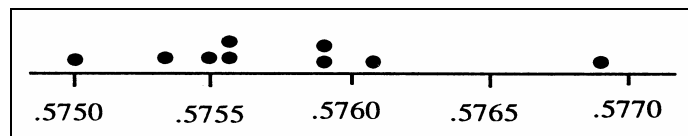


FIGURA 1.4 DIAGRAMA DE PUNTOS

- b) **Histogramas:** El histograma se construye dividiendo el eje horizontal en intervalos (por lo general de la misma magnitud) y tomando sobre el j -ésimo intervalo un rectángulo proporcional a n_j , el número de observación que caen en el intervalo. El histograma representa la tendencia central, la dispersión y la forma general de la distribución de los datos, tal como se presenta en la figura 1.5.

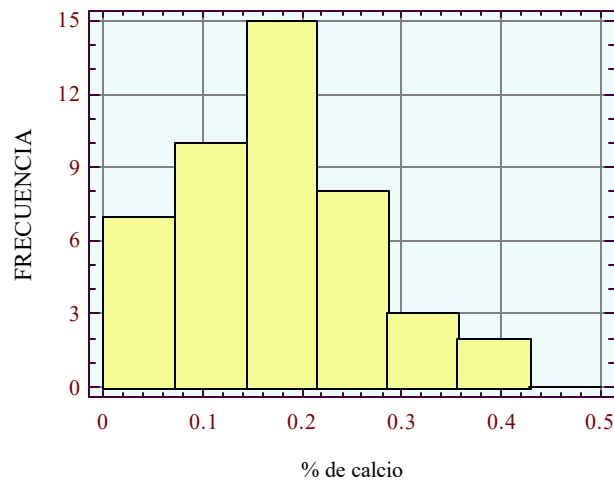


FIGURA 1.5 HISTOGRAMA

- c) **Diagrama de frecuencia acumulada:** Otro gráfico que resulta interesante, es el diagrama de frecuencias acumuladas. En este caso, las abscisas levantadas sobre el límite superior de cada intervalo corresponden a la frecuencia acumulada, es decir a la suma de las frecuencias constatadas en todos los intervalos anteriores al considerado (incluyendo las de éste). La gráfica, tal como se observa en la figura 1.6, tiene forma de línea quebrada creciente. En general, se opera con frecuencias relativas y la altura final es, por tanto, igual a 1. El diagrama de frecuencias acumuladas permite responder a preguntas como ¿qué porcentaje de elementos de entre los estudiados tienen valores inferiores a un determinado valor? o ¿qué valor es esperado por un porcentaje de la muestra?.

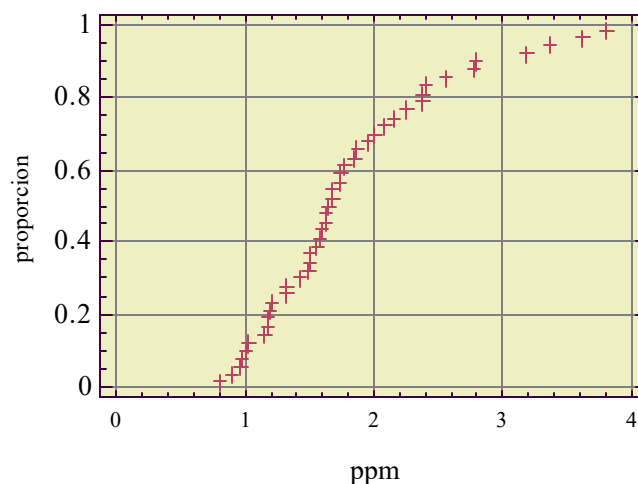


FIGURA 1.6 DIAGRAMA DE FRECUENCIAS ACUMULADAS

- d) **Diagramas de cajas o Box-Whisker:** Es un medio útil para representar datos. En dichos diagramas (ver figura 1.7), los valores máximo y mínimo, los cuartiles inferior y superior (percentil 25 y 75 respectivamente) y la mediana (percentil 50) se representan en una caja rectangular alineada ya sea horizontal o verticalmente. La caja se extiende del cuartil inferior al superior, y es atravesada de un lado a otro por la mediana. A partir de los extremos de la caja se extienden líneas ("bigotes") hasta los valores máximo y mínimo. Si un datos dista del cuartil más cercano en más de 1,5 veces el intervalo intercuartílico se grafica separado de los bigotes.

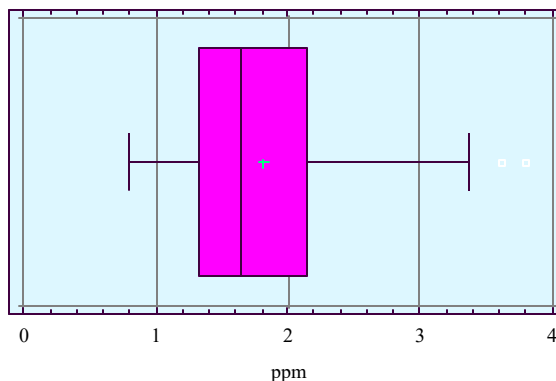


FIGURA 1.7 DIAGRAMA DE CAJA O BOX-WHISKER

1.5.2 EXPRESION CUANTITATIVA DE LAS DISTRIBUCIONES

Las expresiones cuantitativas de las distribuciones de variabilidad de los datos permiten definir tres propiedades importantes de dichos datos que son: forma de la distribución, tendencia central y dispersión. De acuerdo a lo anterior estas expresiones se clasifican en tres tipos generales que son (PEREZ, 1999, cap. 3):

- a) **Medidas de posición:** En vez de manejar todos los datos sobre la variable, tarea que puede ser pesada, se puede caracterizar su distribución de frecuencias mediante algunos valores numéricos, eligiendo como resumen de los datos un valor central alrededor del cual se encuentren distribuidos los valores de la variable. El valor de la variable para representar la distribución se llama medida de posición y es un valor representativo de todos los valores que toma la variable. Debe encontrarse entre el mayor y el menor valor de los datos. Pero estas medidas de posición de los datos deben cumplir determinadas condiciones para que sean verdaderamente representativas de la variable a la que resumen. Toda síntesis de una distribución se considerará como operativa si intervienen en su determinación todos y cada uno de los valores de la distribución y siendo siempre calculable y de fácil obtención. Entre las medidas de posición más comunes se encuentran: *la media aritmética, media aritmética ponderada, media armónica, la media cuadrática, la mediana, moda, cuantiles, momentos potenciales*. En la tabla 1.2 se presenta un resumen de las formulas de cálculo de las medidas de posición más usadas en análisis químico.

- b) **Medidas de dispersión:** Las medidas de dispersión permiten calcular la representatividad de una medida de posición, para lo cual será preciso cuantificar la distancia de los diferentes valores de la distribución respecto a dicha medida (que se pretende que sea su síntesis). A tal distancia es a lo que, en términos estadísticos se le denomina *variabilidad o dispersión* de la distribución. Entre menor sea la distancia, más representativa es la medida de posición, y entre mayor sea dicha distancia, será menos representativa. Resulta pues necesario para completar la información que pueda deducirse de una medida de posición o centralización, acompañarla de uno o varios coeficientes que nos midan el grado de dispersión de la distribución de la variable respecto de esa medida de centralización. Estos coeficientes son los que se conocen como *medidas de dispersión*. Inicialmente se distingue entre medidas de dispersión absolutas y relativas, entendiéndose por relativas las que no dependen de las unidades de medida. Posteriormente se clasifican las medidas absolutas y relativas según sean medidas referentes a promedios o no los sean. Entre las medidas de dispersión absolutas no referentes a promedios se tienen el *recorrido o rango* o diferencia entre el mayor y menor de los datos y el *recorrido intercuartílico* o diferencia existente entre el tercer cuartil y el primero. Entre las medidas de dispersión relativas no referentes a promedios se tienen *el coeficiente de apertura* o cociente entre el mayor valor y el menor valor de una distribución y *el recorrido relativo* o cociente entre el recorrido y la media, así como el *recorrido semintercuartílico* o cociente entre el recorrido intercuartílico y la suma del primer y tercer cuartil. Entre las medidas de dispersión absolutas referentes a promedios se tienen la *varianza* y *la desviación típica*. Entre las medidas de dispersión relativas utilizadas para comparar medidas de posición o promedios, se tienen el *índice de dispersión respecto a la mediana* y el *coeficiente de variación de Pearson*. En la tabla 1.3, se presenta el resumen de las formulas de cálculo de las medidas de dispersión más usadas en quimiometría.
- c) **Medidas de forma:** No se pueden basar las conclusiones únicamente en expresiones que vengan dadas en términos de medidas de posición y dispersión de una distribución, se necesita conocer más sobre el comportamiento de la misma. Si bien intentamos globalizar el comportamiento del colectivo que sea objeto de estudio, para lo cual las medidas de posición son el mejor instrumento, no se debe

proceder a una interpretación que implique un comportamiento de todos los elementos del colectivo uniformemente constante e igual a la medida de posición en cuestión con un error dado por la correspondiente medida de dispersión. Este error o disparidad se hace más ostensible al analizar la representación gráfica de la distribución. Pues bien las medidas de forma de una distribución se basan en su representación grafica, sin llegar a realizar la misma. Las medidas de forma se clasifican en *medidas de asimetría* y *medidas de curtosis o apuntamiento*. Las medidas de asimetría tienen como finalidad el elaborar un indicador que permita establecer el grado de simetría (o asimetría) que presenta una distribución. Estos indicadores pueden ser de dos tipos *coeficiente de asimetría* y *coeficiente de asimetría estandarizado*. Las medidas de curtosis estudian la distribución de frecuencias en la zona central de la misma. La mayor o menor concentración de frecuencias alrededor de la media y de la zona central de la distribución dará lugar a una distribución más o menos apuntada. Por esta razón a las medidas de curtosis se les llama también de apuntamiento o concentración central. Las medidas de curtosis se aplican a distribuciones campaniformes, es decir unimodales simétricas o con ligera asimetría. Los indicadores de curtosis son los *coeficientes de curtosis* y *el coeficiente de curtosis estandarizado*. En la tabla 1.3 se presentan las formulas para las medidas de forma de una distribución.

TABLA 1.2 RESUMEN DE FORMULAS DE MEDIDAS DE POSICIÓN

MEDIDAS DE TENDENCIA CENTRAL	FORMULA DE CALCULO
Media aritmética	$\bar{X} = \frac{\sum x_i}{n}$ \bar{X} = promedio x_i = observaciones n = número de observaciones
Media geométrica	$G = \sqrt{x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}}$ G = media geométrica
mediana	$\tilde{X} = x_{\lfloor \frac{n+1}{2} \rfloor}$ para n impar $\tilde{X} = \frac{x_{(n/2)} + x_{((n/2)+1)}}{2}$ para n par
Cuartiles	$Q_y = \frac{n}{4} y$ y = puede ser 1, 2 o 3

TABLA 1.3 FORMULAS PARA LAS MEDIDAS DE DISPERSION Y FORMA

MEDIDAS DE DISPERSION	ABSOLUTAS	RELATIVAS
Varianza	$s^2 = \frac{\sum_{i=1}^N (x_i - \bar{X})^2}{n-1}$	
Desviación estándar	$s = \sqrt{s^2}$	Coefficiente de variación $CV = \frac{s}{\bar{X}}(100)$
Recorrido o rango	$R = \max(x_i) - \min(x_i)$	Indice de dispersión respecto a la mediana $R_{rel} = \frac{R}{\tilde{X}}$
MEDIDAS DE FORMA	COEFICIENTES	COEFICIENTES ESTANDARIZADOS
Asimetría	$CA = \frac{1}{n-1} \frac{\sum_{i=1}^N (x_i - \bar{X})^3}{s^3}$	$CA_{est} = \frac{CA}{\sqrt{6/n}}$
Curtosis	$CC = \left[\frac{1}{n-1} \frac{\sum_{i=1}^N (x_i - \bar{X})^4}{s^4} \right] - 3$	$CC_{est} = \frac{CC}{\sqrt{6/n}}$

1.5.3 DISTRIBUCIONES DE PROBABILIDAD

La *variabilidad* de una variable aleatoria determinada puede expresarse de forma sintética recurriendo a un tipo de modelo matemático: **la distribución de probabilidad.**

La distribución de probabilidad, manifiesta la forma en que se distribuye la probabilidad de que la variable aleatoria tome un valor o intervalo definido.

Cuando una variable aleatoria es *discreta*, es decir cuando su conjunto de valores posibles es discreto, su distribución de probabilidad viene caracterizada por la **ley de probabilidad o función de cuantía P(x)**, que da la probabilidad de cada uno de los valores posibles de la variable. Cuando el conjunto de valores posibles de la variable es de tipo continuo la caracterización de dicha pauta de variabilidad se efectúa mediante **función de densidad f(x)**, que da la probabilidad de que la variable aleatoria este en un intervalo (MONTGOMERY, 1991, cap.2).

En la figura 1.8, se ilustran distribuciones de probabilidad discretas y continuas hipotéticas. Nótese que en la distribución de probabilidad discreta la probabilidad está representada por la altura de la función, mientras que en el caso continuo es el área bajo la curva $f(x)$, asociada a un intervalo dado, la que representa la probabilidad.

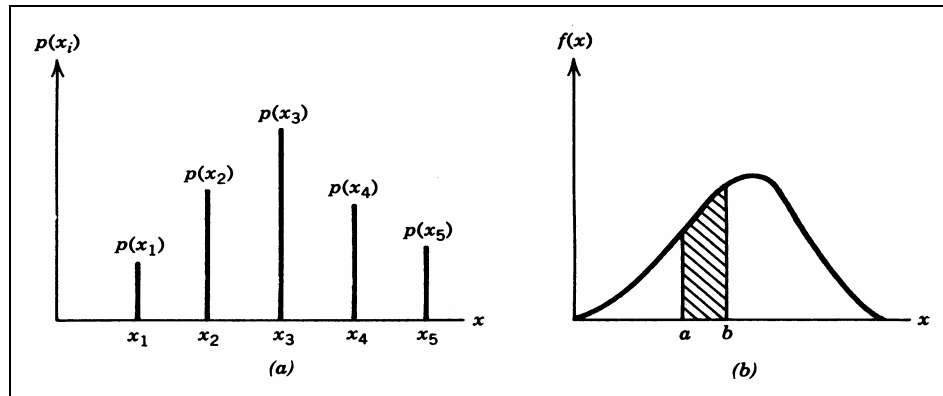


FIGURA 1.8 DISTRIBUCIONES DE PROBABILIDAD (a) caso discreto; (b) caso continuo

Las propiedades de las distribuciones de probabilidad pueden resumirse cuantitativamente como sigue (MONTGOMERY, 1991, cap.2):

- Si la variable aleatoria x es discreta:
 - $0 \leq p(x) \leq 1$ para todos los valores x_i
 - $P(X=x_i) = p(x_i)$ para todos los valores x_i
 - $F(x) = P(X \leq x) = \sum_{\forall x_i \leq x} p(X = x_i)$
(función de distribución)
 - $R(x) = P(X > x) = 1 - F(x)$
(función complementaria)
 - $\sum_{\forall x_i} p(x_i) = 1$
- Si la variable aleatoria x es continua:
 - $0 \leq f(x)$
 - $P(a \leq x \leq b) = \int_a^b f(x) dx$
 - $F(x) = P(X \leq x) = \int_{-\infty}^x f(x) dx$
(función de distribución)
 - $R(x) = P(X > x) = 1 - F(x)$

(función complementaria)

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

1.5.3.1 VALORES ESPERADOS, MEDIA VARIANZA Y DESVIACIÓN TÍPICA

La **media** (μ) para una distribución de probabilidades es una medida de su tendencia central, matemáticamente la media se define como (MONTGOMERY, 1991, cap.2):

$$\mu = \int_{-\infty}^{+\infty} xf(x)dx \quad \text{para } x \text{ continua}$$

$$\mu = \sum_{\forall xi} xp(x) \quad \text{para } x \text{ discreta}$$

También puede expresarse la media en términos del valor esperado o promedio a largo plazo de la variable aleatoria y como $E(x) = \mu$. Donde E es el operador valor esperado.

La extensión o dispersión de una distribución de probabilidad se mide por la **varianza o variancia**, que se define como (MONTGOMERY, 1991, cap.2):

$$\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx \quad \text{para } x \text{ continua}$$

$$\sigma^2 = \sum_{\forall xi} (x - \mu)^2 p(x) \quad \text{para } x \text{ discreta}$$

En términos del valor esperado la variancia puede expresarse como:

$$\sigma^2 = E[(x-\mu)^2]$$

El operador variancia se define como $V(x) \equiv \sigma^2$. La raíz cuadrada de la variancia se conoce como desviación típica. En la tabla 1.4 se presentan algunas propiedades para μ y σ .

Los conceptos expuestos para variables aleatorias unidimensionales pueden generalizarse para el caso k-dimensional. Surgen en este contexto nuevos conceptos, como distribuciones marginales y distribuciones condicionales. Estrechamente relacionada con el

concepto de independencia de sucesos, aparece también la idea de independencia de variables aleatorias.

TABLA 1.4 PROPIEDADES DE LA MEDIA Y DE LA VARIANZA

1. $E(c) = c$
2. $E(x) = \mu$
3. $E(cx) = cE(x) = c\mu$
4. $V(c) = 0$
5. $V(y) = \sigma^2$
6. $V(cx) = c^2 V(x) = c^2 \sigma^2$
Si se tienen dos variables aleatorias por ejemplo, x_1 con μ_1 y σ_1^2 y x_2 con μ_2 y σ_2^2 , entonces:
7. $E(x_1 + x_2) = \mu_1 + \mu_2$
Si x_1 y x_2 son independientes, entonces,
7. $V(x_2 \pm x_1) = \sigma_1^2 + \sigma_2^2$
8. $E(x_1 \cdot x_2) = \mu_1 \cdot \mu_2$

1.5.3.2 EL CASO BIVARIANTE

Considérese que a cada elemento o unidad experimental se le asignan un par de valores (x, y) , en el que x es por ejemplo, la concentración de ión cloro, e y la concentración de ión calcio para determinar la composición iónica en una muestra de agua. Matemáticamente hablando, esta asignación estaría representada por una función (X, Y) que hace corresponder a cada elemento un conjunto de valores (x, y) de R^2 . Dicha función recibe el nombre de variable aleatoria bidimensional o vector aleatorio de dimensión 2.

Los conceptos de variabilidad poblacional y su representación, discutidos en los apartados anteriores pueden generalizarse de manera inmediata al caso multivariante (espacio R^p) y en particular al caso bivalente, como se verá a continuación.

a) DENSIDADES DE PROBABILIDAD CONJUNTA (PRAT et al., 1999, cap. 3)

Cuando se considera la población conceptual, en un histograma generalizado, las variables aleatorias X e Y convergen en general hacia una superficie regular $f(x,y)$ denominada *densidad de probabilidad conjunta* (d.p.c), tal como puede observarse en la figura 1.9. De nuevo, no todas las funciones matemáticas $f(x,y)$ pueden ser una densidad de probabilidad conjunta. Para ello es necesario que:

- I. $f(x,y) \geq 0$ para todo $(x,y) \in \mathbb{R}^2$
- II. $\iint_{\mathbb{R}^2} f(x,y) dx dy = 1$

En el caso bivalente se tendrá que:

$$P(x_o \leq X \leq x_o + dx, y_o \leq Y \leq y_o + dy) = f(x_o, y_o) dx dy$$

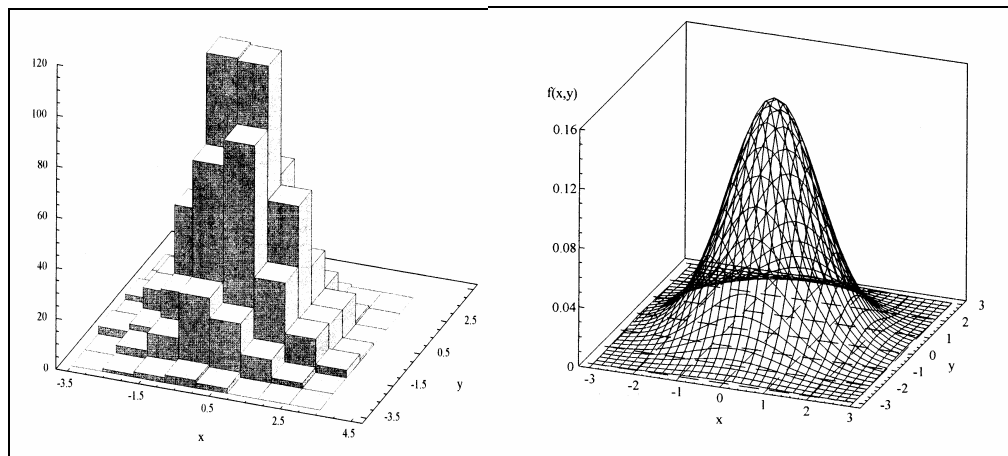


FIGURA 1.9 HISTOGRAMA GENERALIZADO Y CONVERGENCIA A UNA SUPERFICIE REGULAR

b) DENSIDADES MARGINALES (PRAT et al., 1999, cap. 3)

A partir del conocimiento de la d.p.c, $f(x,y)$ de un vector aleatorio (X,Y) , siempre es posible obtener la densidad de probabilidad univariante correspondiente a cada una de las dos variables aleatorias que definen el vector.

La distribución univariante de X se conoce como *densidad marginal de X* y está definida por:

$$f_x(x) = \int_R f(x, y) dy$$

La distribución univariante de Y se conoce como *densidad marginal de Y* y está definida por:

$$f_y(y) = \int_R f(x, y) dx$$

En general, no siempre es posible obtener una densidad conjunta a partir de las densidades marginales, aunque si los será en el caso en que las variables aleatorias X y Y sean estocásticamente independientes.

c) DENSIDADES CONDICIONALES E INDEPENDENCIA DE VARIABLES (PRAT et al., 1999, cap. 3)

Sea (X, Y) un vector aleatorio cuya d.p.c. es $f(x, y)$ y sea y_o un cierto valor de Y tal que $f_y(y_o) \neq 0$. Se define como densidad de X condicional al valor de y_o de Y a la densidad univariante (PRAT et al., 1999, cap. 3):

$$f(x / Y = y_o) = f(x / y_o) = \frac{f(x, y)}{f_y(y_o)}$$

Si x_o es un valor de X tal que $f_x(x_o) \neq 0$, se puede definir también la densidad de Y condicionada al valor de x_o de X como:

$$f(y / X = x_o) = f(y / x_o) = \frac{f(x, y)}{f_x(x_o)}$$

Así pues, en general se tiene que:

$$f(x, y) = f(y / x_o) f_x(x_o) = f(x / y_o) f_y(y_o)$$

El concepto de densidad condicional permite definir la noción de independencia entre variables aleatorias. Dado el vector aleatorio (X, Y) , se dirá que X, Y son *variables aleatorias independientes* si y sólo si $f(x / y_o) = f_x(x)$. Debe notarse que en este caso:

$$f(x, y) = f_x(x) * f_y(y)$$

y por tanto, si las variables X, Y son independientes, es posible obtener la densidad como producto de las marginales.

Los conceptos de densidad conjunta, marginal y condicional así como la independencia estocástica pueden extenderse fácilmente a los vectores de dimensión p . En particular, si (X_1, X_2, \dots, X_p) es un vector aleatorio, se dirá que las variables que lo componen X_1, X_2, \dots, X_p son variables aleatorias independientes si y sólo si la densidad de probabilidad conjunta es tal que:

$$f(x_1, x_2, \dots, x_p) = f_{X_1}(x_1) f_{X_2}(x_2) \dots f_{X_p}(x_p)$$

d) COVARIANZA (PRAT et al., 1999, cap. 3)

Sea (X, Y) un vector aleatorio. La variabilidad de dicho vector será representada por su función de densidad de probabilidad conjunta $f(x, y)$.

Un parámetro de interés para caracterizar la dependencia lineal entre las variables X e Y es la covarianza definida de la forma siguiente:

$$\mu_{11} = COV(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$$

Donde μ_x y μ_y son las medias de las densidades marginales.

1.5.3.3 MODELOS PROBABILÍSTICOS

En las tablas 1.5 y 1.6 se presentan las distribuciones de probabilidad más generalizadas para variables aleatorias discretas y continuas respectivamente.

En la tabla 1.7, se presenta un resumen de los modelos gráficos, matemáticos, así como también del contexto del uso para las distribuciones cuyo conocimiento es fundamental para las aplicaciones estadísticas en análisis químico, entre las que se encuentran: *la distribución binomial, hipergeométrica y la de Poisson* para las discretas y las distribuciones *exponencial, triangular, uniforme, log-normal y normal* (merecedora de comentarios aparte, sección 1.6) para las continuas.

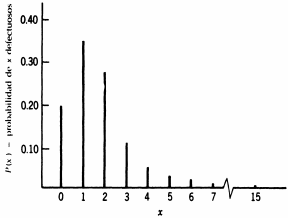
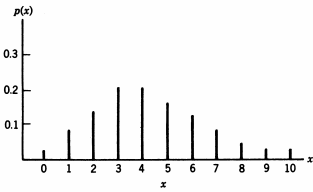
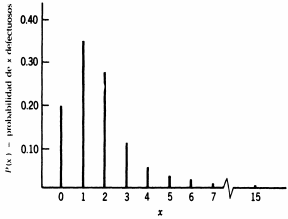
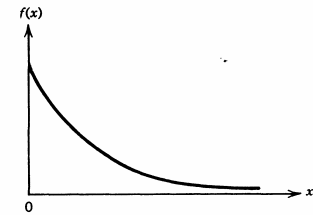
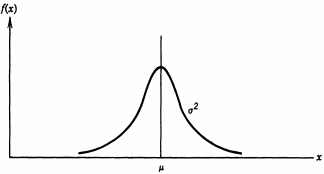
TABLA 1.5 DISTRIBUCIONES DISCRETAS MAS GENERALIZADAS

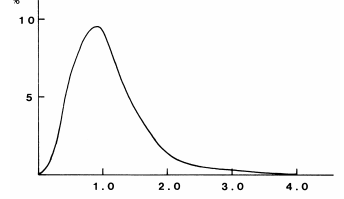
<i>Distribución</i>	<i>Ley de probabilidad</i>	<i>Parámetros</i>	<i>Media</i>	<i>Varianza</i>
<i>Bernoulli</i>	$p^x(1-p)^{1-x}$ $x=0,1$	p $0 < p < 1$	p	$p(1-p)$
<i>Binomial</i>	$\binom{n}{x} p^x(1-p)^{n-x}$ $x=0,1,\dots,n$	n, p $n=1,2,\dots$ $0 < p < 1$	np	$np(1-p)$
<i>Poisson</i>	$\frac{\lambda^x}{x!} e^{-\lambda}$	$\lambda > 0$	λ	λ
<i>Geométrica</i>	$p(1-p)^{x-1}$ $x=1,2,\dots$	$0 \leq p \leq 1$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
<i>Binomial negativa</i>	$\binom{n+x-1}{x} p^x(1-p)^n$ $x=0,1,\dots$	$n > 0$ $0 \leq p \leq 1$	$\frac{n(1-p)}{p}$	$\frac{n(1-p)}{p^2}$
<i>Hiper-geométrica</i>	$\frac{\binom{Np}{x} \binom{N(1-p)}{n-x}}{\binom{N}{n}}$ $x=0,1,\dots,n$	$N=1,2,\dots$ $n=1,2,\dots,N$ $p=0, \frac{1}{N}, \dots, 1$	np	$np(1-p)$ * $\frac{N-n}{N-1}$

TABLA 1.6 DISTRIBUCIONES CONTINUAS MAS GENERALIZADAS

Distribuc.	Función de densidad		Media	Varianza
Uniforme	$\frac{1}{b-a}$ $a < x < b$	$a \in R$ $b \in R$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	$\frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}$	$\mu \in R$ $\sigma > 0$	μ	σ^2
Exponen.	$\lambda e^{-\lambda x}$	$\lambda, x > 0$	$1/\lambda$	$1/\lambda^2$
Gamma	$\frac{x^{a-1} e^{-px}}{\Gamma(a)}$ $\Gamma(a) = \int_0^{\infty} e^{-x} x^{a-1} dx$	$a, p > 0$ $x > 0$	$\frac{a}{p}$	$\frac{a}{p^2}$
Beta	$\frac{x^{p-1} (1-x)^{q-1}}{B(p, q)}$ $B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$	$p, q > 0$ $0 < x < 1$	$\frac{p}{p+q}$	$\frac{pq}{(p+q)^2}$
Chi - cuadrado	$\frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2}$	$x \geq 0$ $n = 1, \dots$	n	$2n$
t de Student	$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$	$n = 1, \dots$	0	$\frac{\Gamma\left(\frac{n-2}{2}\right) \Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{n-1}{2}\right)}$
F de Fisher	$\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{\frac{m}{2}-1} \left(1 + \frac{mx}{n}\right)^{-\frac{m+n}{2}}$	$x > 0$ $m = 1, \dots$ $n = 1, \dots$	$\frac{n}{n-2}$	$\frac{n^2(2m+2n-4)}{m(n-2)^2(n-4)}$
Weibull	$\frac{\beta}{\eta} (x-\gamma)^{\beta-1} e^{-\left(\frac{x-\gamma}{\eta}\right)^\beta}$	$x \geq \gamma$ $\beta, \eta > 0$ $\gamma \in R$	$\gamma + \eta \Gamma\left(1 + \frac{1}{\beta}\right)$	$\frac{\beta}{\alpha} \left[2\Gamma\left(\frac{2}{\alpha}\right) - \frac{1}{\alpha}\right]$
Log - normal	$\frac{1}{x\sigma\sqrt{2\pi}} e^{-(\ln x - \mu)^2 / 2\sigma^2}$	$\mu \in R$ $x > 0$	$e^{\mu + \frac{\sigma^2}{2}}$	$e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$
Laplace	$\frac{\lambda}{2} e^{-\lambda x-\alpha }$	$\alpha \in R$ $\lambda > 0$	α	$\frac{2}{\lambda^2}$

TABLA 1.7 DISTRIBUCIONES MAS FRECUENTES EN ANALISIS QUIMICO

Distribución	Forma	Distribución	Uso
Binomial		$P(X=x) = \binom{n}{x} p^x q^{n-x}$ <p>La media de la distribución es:</p> $\bar{x} = np$ <p>y la varianza:</p> $\sigma^2 = np(1-p)$	Se presenta si en una población finita se realiza un muestreo sin reemplazo. Por ejemplo, sea una variable que consiste en la presencia o no de una característica. Si se toman n mediciones independientes (pruebas de Bernoulli), y la probabilidad de la presencia de la característica es p , y esta es constante, entonces el número de "éxitos" (presencia de la característica) x en n pruebas de Bernoulli tendrá una distribución binomial.
Poisson		$P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!}$ <p>En esta distribución:</p> $\bar{x} = \sigma^2 = \lambda = np$	Si n es grande ($n > 50$) y p pequeño ($p < 0.1$), las distribuciones binomiales suelen aproximarse a las de Poisson. Este es un modelo del número de veces que aparece una característica, en la unidad de medición, que puede ser el tiempo, área, volumen, etc., como ocurre en las técnicas de conteos intensidades radiactivas.
Hipergeométrica		$P(X=x) = \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}}$ <p>En esta distribución:</p> $\bar{x} = nD/N$ $\sigma^2 = \bar{x} \left(1 - \frac{D}{N} \right) \left(\frac{N-n}{N-1} \right)$	Aplicable para definir la probabilidad de x ocurrencias en n pruebas de un suceso cuando hay un total de D ocurrencias en una población de N elementos.
Exponencial		$f(x) = \lambda e^{-\lambda x}$ <p>para esta distribución:</p> $\bar{x} = \frac{1}{\lambda}$ $\sigma^2 = \frac{1}{\lambda^2}$	Aplicable cuando es probable que se produzcan más observaciones por debajo que por encima de la media. La distribución exponencial se utiliza ampliamente en el campo de la ingeniería de la confiabilidad como un modelo para el tiempo hasta el fallo de un componente o sistema. Esta cuenta con ser una información que debe proveerse a un analista, para el mantenimiento de sus equipos.
Normal		$P(x \leq a) = F(a) = \int_{-\infty}^a \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$	Aplicable cuando hay concentración de observaciones alrededor del promedio y es igualmente probable que las observaciones sucedan arriba y abajo del promedio. La variación en las observaciones, por lo general, es el resultado de muchas causas pequeñas. Es una de las distribuciones de mayor importancia en análisis químico.

<p>Log-Normal</p>		$\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$	<p>Se cumple que si X es $N(\mu, \sigma)$ entonces $\ln x$ es log-normal. Esta distribución es de amplio uso en la teoría de los errores. Y en la teoría de la supervivencia como modelo de ajuste de vida. Algunas concentraciones de analito tienen distribución log-normal</p>
-------------------	---	---	---

1.5.3.4 ALGUNAS APROXIMACIONES UTILES

En ciertos problemas es útil a veces aproximar una distribución de probabilidad con otra. Esto es en especial provechoso en situaciones en las que se hace difícil manipular analíticamente la distribución original, o ésta no está bien tabulada. Entre las aproximaciones más útiles se encuentran:(1) la aproximación binomial a la hipergeométrica; (2) la aproximación de Poisson a la binomial; (3) la aproximación normal a la binomial. En la figura 1.10, se presenta un resumen de las distintas aproximaciones antes mencionadas. En esta figura, H, B, P y N representan las distribuciones hipergeométrica, binomial, de Poisson y normal respectivamente. La disponibilidad de computadoras y calculadoras modernas ha hecho innecesarias dichas aproximaciones, pero existen casos esporádicos en que sí son útiles (MONTGOMERY, 1991, cap. 2).

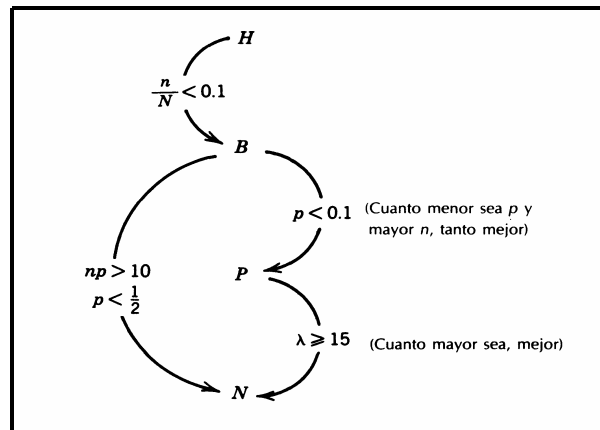


FIGURA 1.10 APROXIMACIONES A LAS DISTRIBUCIONES DE PROBABILIDAD (MONTGOMERY, 1991, cap. 2)

1.6 SIGNIFICADO PRACTICO DE LA DESVIACION ESTANDAR

Se ha subrayado que una desviación estándar pequeña para un conjunto de valores indica que estos se encuentran localizados cerca de la media. Por lo contrario, una desviación de valor grande revela que las observaciones están lejos de la media. El matemático ruso P.L. Chebyshev (1821-1894) desarrolló un teorema que permite determinar la proporción mínima de los valores que se encuentran dentro de un número específico de desviaciones estándar con respecto a la media. En términos generales el **Teorema Chebyshev** establece que:

“Para un conjunto cualquiera de observaciones, la proporción mínima de los valores que se encuentran dentro de k desviaciones estándares desde la media es al menos $1 - 1/k^2$, donde k es una constante mayor o igual que 1” (MENDENHALL, 1990, cap. 2).

Nótese que el teorema es válido para cualquier valor de k que se elija, con tal que sea por lo menos igual a 1. Entonces, al calcular la fracción $1 - 1/k^2$, se observa que el teorema afirma que por lo menos esta fracción del número total de n mediciones caerá dentro del intervalo determinado. En la tabla 1.8, se presentan algunos valores ilustrativos para $1 - 1/k^2$ y los intervalos $\mu \pm k\sigma$ (MENDENHALL, 1990, cap. 2).

TABLA 1.8 VALORES ILUSTRATIVOS PARA $1 - 1/k^2$

k	$1 - 1/k^2$	$\mu \pm k\sigma$	Significado
1	0	$\mu \pm 1\sigma$	Poco significado práctico
2	3/4	$\mu \pm 2\sigma$	Al menos 3/4 de las mediciones caerán en este intervalo.
3	8/9	$\mu \pm 3\sigma$	Al menos 8/9 de las mediciones caerán en este intervalo.

El teorema de Chebyshev es muy conservador. Se puede aplicar a cualquier distribución de mediciones y proporciona una cota inferior para la fracción de mediciones que caen dentro de un intervalo específico. En la mayoría de los casos, la verdadera fracción será mucho mayor que $1 - 1/k^2$ (MENDENHALL, 1990, cap. 2).

1.7 LA DISTRIBUCIÓN NORMAL

La *distribución normal* es sin duda la más importante de la *estadística*, debido no sólo al gran número de variables que la siguen, sino al importante teorema central del límite que posibilita la mayor parte de sus aplicaciones (PEÑA y ROSELL, 1996, cap. 7).

Se trata como ya se expuso en la sección 1.5.3.2, de una distribución de probabilidad continua cuya fórmula es (PEÑA y ROSELL, 1996, cap. 7):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Donde: μ = media de la distribución y σ = desviación típica de la distribución.

Una gran parte de las variables en los problemas de análisis químico siguen distribuciones normales o casi normales, por ejemplo, las mediciones de concentraciones de muchos analitos, mediciones de volúmenes, mediciones de pesos, etc.

La distribución probabilística normal y su respectiva curva normal tiene las siguientes características:

1. La curva normal es **acampanada** y presenta un solo pico en el centro de la distribución. La media (aritmética), la mediana y la moda de la distribución son iguales y están localizadas en el pico. De esta forma, la mitad del área bajo la curva se encuentra por arriba de este punto central, y la otra mitad por debajo.
2. La distribución probabilística normal es **simétrica** con respecto a su media. Si se corta la curva normal verticalmente en este valor central, las dos mitades se reflejarán como imágenes en el espejo.
3. La curva normal decrece uniformemente en ambas direcciones a partir del valor central. Es **asintótica**, lo cual significa que la curva se acerca cada vez más al eje X, pero en realidad nunca llega a tocarlo. Esto es, los puntos extremos de la curva se extienden indefinidamente en uno y otro sentido.

4. Si una variable aleatoria x se distribuye normalmente cualquier transformación lineal $y = a + bx$, se distribuye también normalmente (ROMERO y ZUNICA, 1993, cap.5).
5. Si x y y son dos variables aleatorias independientes que se distribuyen normalmente su suma $w = x + y$, se distribuye normalmente (ROMERO y ZUNICA, 1993, cap.5).

Según la función de densidad normal, la distribución queda completamente determinada por su media y su desviación típica. De esta forma conociendo las μ y σ , de alguna distribución, puede calcularse la probabilidad (área bajo la curva) de que la variable x tome un valor entre dos cotas cualesquiera. En la figura 1.11 se reflejan curvas normales con diferentes medias y desviaciones típicas, lógicamente, el área bajo la curva entre dos valores de x , "a" y "b", no es la misma.

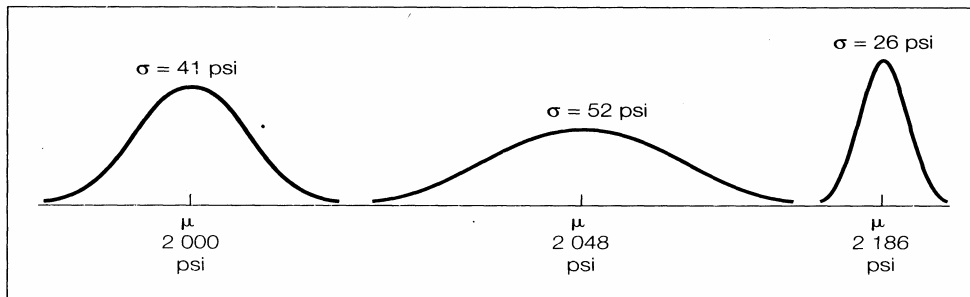


FIGURA 1.11 DISTRIBUCIONES NORMALES CON DIFERENTES MEDIAS Y DESVIACIONES ESTÁNDAR

El cálculo de la probabilidad de que el valor de x sea menor o igual a un valor a , viene dado a partir de:

$$P(x \leq a) = F(a) = \int_{-\infty}^a \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

La dificultad matemática de la fórmula queda solventada mediante la tabulación de sus valores; no obstante, al ser diferentes para cada μ y σ , se tendrían infinitas tablas (PEÑA y ROSELL, 1996, cap. 7).

1.7.1 DISTRIBUCION PROBABILISTICA NORMAL ESTANDAR

Como se puntualizo anteriormente, podría decirse que existe una "familia de distribuciones normales". Por lo tanto el número de distribuciones normales es ilimitado. Resultaría físicamente imposible proporcionar una tabla de probabilidades para cada combinación de μ y σ . Por fortuna, puede utilizarse un elemento de la familia de distribuciones normales para todos los problemas donde tal distribución resulte aplicable. Tiene una media igual a 0 y una desviación estándar igual a 1, y se denomina **distribución normal estándar** (MASON et al., 2001, cap. 7), ver figura 1.12. Cualquier distribución normal puede convertirse en una "distribución normal estándar", para relacionar a dichas distribuciones es necesario hacer un cambio de variable o de escala de la siguiente forma (PEÑA y ROSELL, 1996, cap. 7):

$$z = \frac{x - \mu}{\sigma}$$

Lo que se hace con este cambio de escala es transformar una variable "x" con distribución $N(\mu, \sigma)$ en una variable z con distribución $N(0,1)$, a este proceso se le denomina "**tipificar**" o "**estandarizar**" la variable (PEÑA y ROSELL, 1996, cap. 7).

El valor de la curva normal estándar ($N(0,1)$) viene definida por (PEÑA y ROSELL, 1996, cap. 7):

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

La probabilidad de que una variable normal tipificada sea menor o igual que un valor dado, vendrá dada por la integral de su función de densidad desde $-\infty$ hasta z, es decir:

$$P(x < a) = P\left(\frac{(x-\mu)}{\sigma} < \frac{(a-\mu)}{\sigma}\right) = P(N(0,1) < z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

Los valores de esta integral no pueden calcularse directamente por no existir primitiva de $f(x)$ (ROMERO y ZUNICA, 1993, cap.5). Sin embargo, para valores de z entre 0 y 4 esta

probabilidad viene recogida en tablas como la que se presenta en el anexo A1. En la figura 1.12 se presenta el área de probabilidad de una distribución normal estándar.

En la actualidad con el avance de las computadoras y calculadoras, el uso de tablas se va volviendo menos necesario.

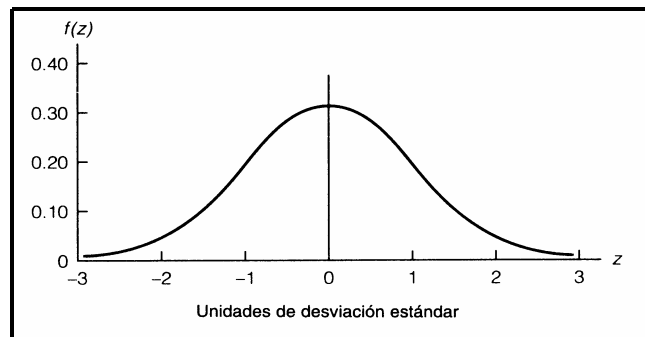


FIGURA 1.12 DISTRIBUCION NORMAL ESTANDAR

1.7.2 REGLA EMPIRICA O REGLA NORMAL

El teorema de Chebyshev se refiere a cualquier conjunto de valores; esto es, la distribución de estos puede tomar cualquier forma (ver sección 1.6). Sin embargo, para una *curva de distribución simétrica de campana*, como la normal, se logra ser más precisos al explicar la dispersión con respecto a la media. Estas relaciones entre la desviación estándar y la media se incluyen en la denominada **regla empírica** que algunas veces se denomina **regla normal** (MENDENHALL, 1990, cap. 2). Esta regla se explica como:

“Para una distribución de frecuencias simétrica de campana, aproximadamente 68% de las observaciones estará a más y menos una desviación estándar desde la media; aproximadamente 95% de tales observaciones se encontrará a más y menos dos desviaciones estándares de la misma; y prácticamente todas las observaciones (99.7%) se hallarán a más y menos tres desviaciones estándares con respecto a la media”.

En el caso de la distribución normal el ajuste a la regla empírica es casi exacto. Las relaciones en la regla empírica se presentan en forma grafica en la figura 1.13.

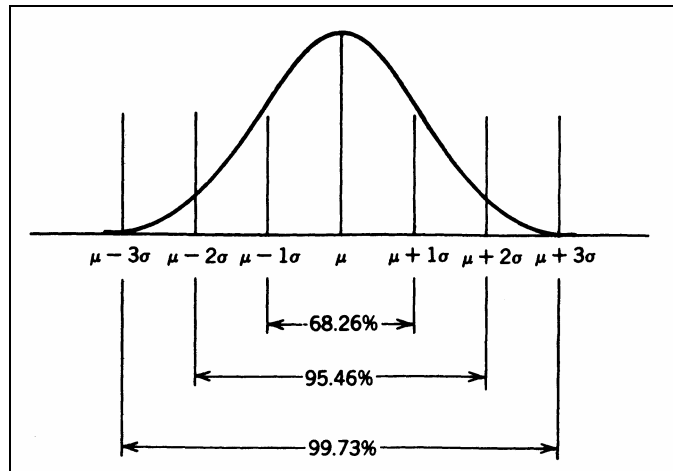


FIGURA 1.13 RELACIONES ENTRE LA MEDIA Y DESVIACIÓN ESTÁNDAR DE ACUERDO A LA REGLA EMPÍRICA

1.8 APROXIMACIONES NORMALES

Si x_1, x_2, \dots, x_n es una sucesión de n variables aleatorias independientes con $E(x_i) = \mu_i$ y $V(x_i) = \sigma_i^2$ (ambas finitas), y $y = x_1 + x_2 + \dots + x_n$, entonces (MONTGOMERY, 1991, cap.3):

$$z_n = \frac{y - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$$

tiene aproximadamente una distribución $N(0, 1)$ en el sentido de que, si $F_n(z)$ es la función de distribución de z_n , y si $\Phi(z)$ es la función de distribución de la variable aleatoria $N(0,1)$, entonces $\lim_{n \rightarrow \infty} [F_n(z) / \Phi(z)] = 1$. A este resultado se le llama el *teorema central del límite* y es uno de los postulados de gran importancia en estadística (MONTGOMERY, 1991, cap.3).

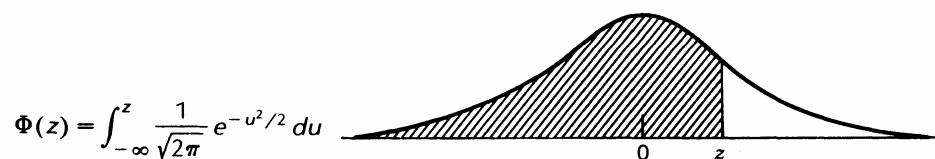
Esencialmente lo que este resultado refleja es que en condiciones muy generales respecto a las distribuciones de los sumandos, la suma de variables aleatorias independientes tiende a distribuirse normalmente a medida que aumenta el número de sumandos. Este resultado teórico justifica, en cierto sentido, la frecuencia con la que se presentan en la realidad variables aleatorias cuya distribución se asemeja a la pauta de variabilidad de la

distribución normal. En efecto muchas variables reales pueden considerarse como el resultado de la asociación de un conjunto de factores independientes (MONTGOMERY, 1991, cap.3). Así, la medición de una concentración de una solución estándar depende de las características del estándar, de la medición de los volúmenes, de las condiciones ambientales del laboratorio, etc. Como consecuencia del *teorema central del límite* cabe esperar que la variabilidad en la medición de esa concentración, que en cierto sentido es la suma de una serie de factores independientes, se distribuirá aproximadamente de forma normal.

REFERENCIAS BIBLIOGRÁFICAS

- **García Álvarez, M.C. (1998).** Análisis, homologación y control de calidad. Curso de maestría en química analítica. Universidad de Valencia.
- **Kinney, T.C. y Taylor, T.R. (1998).** Investigación de mercados. Quinta edición. México: Mc Grawhill Internacional.
- **Mason R.D., Lind, D.A. y Marchal, W.G. (2001).** Estadística para administración y economía. Décima edición. México: Alfaomega grupo editor.
- **Mendenhall, W. (1990).** Estadística para administradores. Segunda edición. México: Grupo Editorial Iberoamerica, S.A. de C.V.
- **Miller, J. C. y Miller, J. N. (1993).** Estadística para química analítica. Segunda edición. Wilmington, Delaware, E.U.A: Addison-Wesley Iberoamericana, S.A.
- **Montgomery, D. C. (1991).** Diseño y análisis de experimentos. México: Grupo Editorial Iberoamerica, S.A. de C.V.
- **Montgomery, D. C. (1991).** Control estadístico de la calidad. México: Grupo Editorial Iberoamerica, S.A. de C.V.
- **Peña Pitarch E. y Rossell J.M. (1996).** Gerencia del control de calidad. Curso de capacitación. El Salvador: "Centre de Cooperació al Desenvolupament"; Universidad Politécnica de Catalunya.
- **Pérez, C. (1999).** Control estadístico de calidad. Mexico: Alfaomega grupo editor S.A. de C.V.
- **Prat, A., Tort-Martorell, X., Grima, P. y Pozueta, L. (2000).** Métodos estadísticos. Control y mejora de la calidad. México: Alfaomega Grupo Editor.
- **Ramis Ramos, G. (1998).** Quimiometría. Curso de maestría en química analítica. Universidad de Valencia.
- **Romero R. y Zúñiga, L.R. (1993).** Estadística (Proyecto de innovación educativa). España: Servicio de publicaciones Universidad Politécnica de Valencia.
- **Skoog, D.A. y West, D.M. (2001).** Química analítica. Séptima edición. México: MC GRAWHILL.

ANEXO A1. DISTRIBUCION ACUMULATIVA NORMAL ESTANDAR
(Tomado de: Montgomery, 1991, p. 430)



z	0.00	0.01	0.02	0.03	0.04	z
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.0
0.1	0.53983	0.54379	0.54776	0.55172	0.55567	0.1
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.2
0.3	0.61791	0.62172	0.62551	0.62930	0.63307	0.3
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.4
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.5
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.6
0.7	0.75803	0.76115	0.76424	0.76730	0.77035	0.7
0.8	0.78814	0.79103	0.79389	0.79673	0.79954	0.8
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.9
1.0	0.84134	0.84375	0.84613	0.84849	0.85083	1.0
1.1	0.86433	0.86650	0.86864	0.87076	0.87285	1.1
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	1.2
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	1.3
1.4	0.91924	0.92073	0.92219	0.92364	0.92506	1.4
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	1.5
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	1.6
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	1.7
1.8	0.96407	0.96485	0.96562	0.96637	0.96711	1.8
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	1.9
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	2.0
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	2.1
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	2.2
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	2.3
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	2.4
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	2.5
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	2.6
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	2.7
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	2.8
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	2.9
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	3.0
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	3.1
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	3.2
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	3.3
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	3.4
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	3.5
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	3.6
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	3.7
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	3.8
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	3.9

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

z	0.05	0.06	0.07	0.08	0.09	z
0.0	0.519 94	0.523 92	0.527 90	0.531 88	0.535 86	0.0
0.1	0.559 62	0.563 56	0.567 49	0.571 42	0.575 34	0.1
0.2	0.598 71	0.602 57	0.606 42	0.610 26	0.614 09	0.2
0.3	0.636 83	0.640 58	0.644 31	0.648 03	0.651 73	0.3
0.4	0.673 64	0.677 24	0.680 82	0.684 38	0.687 93	0.4
0.5	0.708 84	0.712 26	0.715 66	0.719 04	0.722 40	0.5
0.6	0.742 15	0.745 37	0.748 57	0.751 75	0.754 90	0.6
0.7	0.773 37	0.776 37	0.779 35	0.782 30	0.785 23	0.7
0.8	0.802 34	0.805 10	0.807 85	0.810 57	0.813 27	0.8
0.9	0.828 94	0.831 47	0.833 97	0.836 46	0.838 91	0.9
1.0	0.853 14	0.855 43	0.857 69	0.859 93	0.862 14	1.0
1.1	0.874 93	0.876 97	0.879 00	0.881 00	0.882 97	1.1
1.2	0.894 35	0.896 16	0.897 96	0.899 73	0.901 47	1.2
1.3	0.911 49	0.913 08	0.914 65	0.916 21	0.917 73	1.3
1.4	0.926 47	0.927 85	0.929 22	0.930 56	0.931 89	1.4
1.5	0.939 43	0.940 62	0.941 79	0.942 95	0.944 08	1.5
1.6	0.950 53	0.951 54	0.952 54	0.953 52	0.954 48	1.6
1.7	0.959 94	0.960 80	0.961 64	0.962 46	0.963 27	1.7
1.8	0.967 84	0.968 56	0.969 26	0.969 95	0.970 62	1.8
1.9	0.974 41	0.975 00	0.975 58	0.976 15	0.976 70	1.9
2.0	0.979 82	0.980 30	0.980 77	0.981 24	0.981 69	2.0
2.1	0.984 22	0.984 61	0.985 00	0.985 37	0.985 74	2.1
2.2	0.987 78	0.988 09	0.988 40	0.988 70	0.988 99	2.2
2.3	0.990 61	0.990 86	0.991 11	0.991 34	0.991 58	2.3
2.4	0.992 86	0.993 05	0.993 24	0.993 43	0.993 61	2.4
2.5	0.994 61	0.994 77	0.994 92	0.995 06	0.995 20	2.5
2.6	0.995 98	0.996 09	0.996 21	0.996 32	0.996 43	2.6
2.7	0.997 02	0.997 11	0.997 20	0.997 28	0.997 36	2.7
2.8	0.997 81	0.997 88	0.997 95	0.998 01	0.998 07	2.8
2.9	0.998 41	0.998 46	0.998 51	0.998 56	0.998 61	2.9
3.0	0.998 86	0.998 89	0.998 93	0.998 97	0.999 00	3.0
3.1	0.999 18	0.999 21	0.999 24	0.999 26	0.999 29	3.1
3.2	0.999 42	0.999 44	0.999 46	0.999 48	0.999 50	3.2
3.3	0.999 60	0.999 61	0.999 62	0.999 64	0.999 65	3.3
3.4	0.999 72	0.999 73	0.999 74	0.999 75	0.999 76	3.4
3.5	0.999 81	0.999 81	0.999 82	0.999 83	0.999 83	3.5
3.6	0.999 87	0.999 87	0.999 88	0.999 88	0.999 89	3.6
3.7	0.999 91	0.999 92	0.999 92	0.999 92	0.999 92	3.7
3.8	0.999 94	0.999 94	0.999 95	0.999 95	0.999 95	3.8
3.9	0.999 96	0.999 96	0.999 96	0.999 97	0.999 97	3.9

CAPITULO II

MUESTRAS ALEATORIAS Y DISTRIBUCIONES MUESTRALES

CONTENIDO

	Pág.
2.0 Muestreo y muestras aleatorias independientes	2-2
2.1 Definición estadística de muestra en análisis químico	2-5
2.2 Noción de estadístico (ó estadígrafo)	2-6
2.3 Distribuciones de muestreo	2-9
2.4 Pruebas de normalidad	2-11
2.5 Ejercicios de aplicación de la descripción de la variabilidad	2-19
REFERENCIAS BIBLIOGRAFICAS	2-34
ANEXO A2. Tabla de distribución t-Student	2-35
ANEXO B2. Tabla de distribución Chi-cuadrada	2-36
ANEXO C2. Tabla de distribución F-Snedecor	2-37



2.0 MUESTREO Y MUESTRAS ALEATORIAS INDEPENDIENTES

En todo problema estadístico existe un conjunto de elementos sobre los que se recoge información. En general, no resulta posible estudiar la totalidad de elementos de la población para obtener información sobre ésta. Incluso cuando esta posibilidad exista técnicamente, como es el caso de poblaciones finitas, dicho procedimiento suele ser impracticable.

En consecuencia para obtener información sobre una población hay que limitarse a analizar sólo un subconjunto de la misma. A este subconjunto se le denomina *muestra* y a los valores que se observan sobre los elementos que la constituyen, se les conoce como *datos estadísticos*.

Usualmente la razón para tomar muestras es una de las siguientes:

- a) Por restricciones de tiempo, dinero y personal.
- b) La población no existe físicamente (como en el caso de la metrología).
- c) El examen de cada elemento requiere de la destrucción del mismo.

La forma de seleccionar o generar los elementos que han de constituir la *muestra*, se conoce en general como *muestreo*, y tiene como es lógico, una importancia capital, pues esto permite garantizar la obtención de conclusiones que puedan extrapolarse válidamente a la población de la que la muestra procede. Lo anterior implica que, con el objeto de poder "*inferir*" conclusiones válidas sobre una población, la muestra debe ser *representativa* de la misma. En realidad, en muchos casos un conocimiento previo sobre la población es indispensable para decidir si una muestra puede considerarse o no representativa.

¡No hay que olvidar nunca que el objeto final del estudio es siempre la población y que la muestra es sólo un medio para obtener información sobre ésta!

En la figura 2.1 se representa el esquema del objetivo de la obtención de muestras.

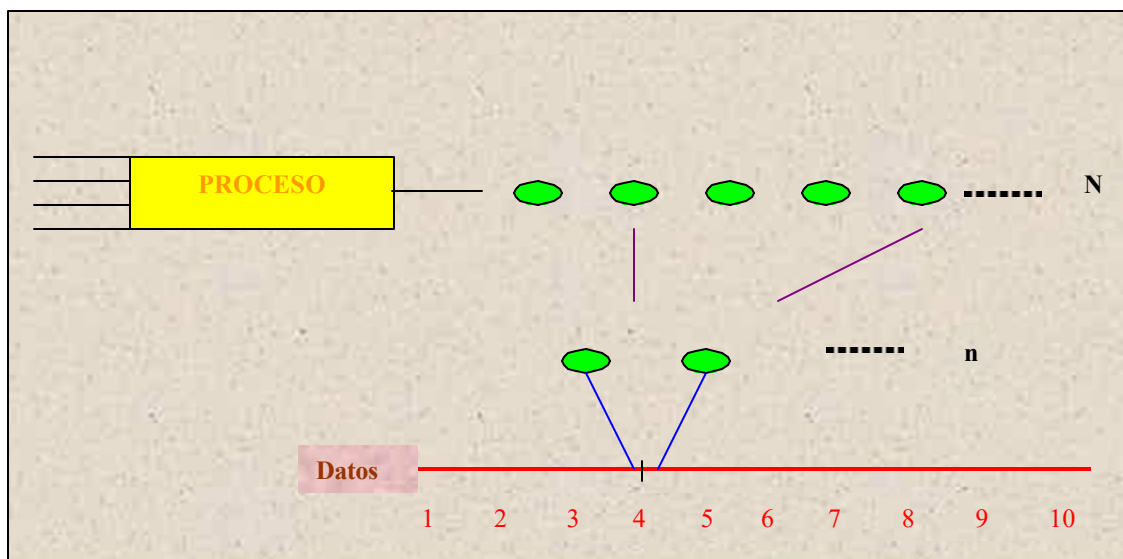


FIGURA 2.1 PROCESO DE OBTENCION DE MUESTRAS

Existen muchos y diferentes procedimientos mediante los cuales los investigadores pueden seleccionar sus muestras, pero inicialmente debe establecerse un concepto fundamental, la diferencia entre (1) una muestra probabilística y (2) una muestra no probabilística (HERNANDEZ et al., 1998, cap.8).

En el *muestreo probabilístico* cada elemento de la población tiene una oportunidad (probabilidad) conocida de ser seleccionado para la muestra, es decir que la selección de la muestra constituye un fenómeno aleatorio probabilizable. Dicha selección se verificará entonces, en condiciones de azar, siendo susceptible de medida de la incertidumbre derivada de la misma. El muestreo se hace mediante reglas de decisión matemáticas que no permiten la discreción al investigador. El muestreo probabilístico permite calcular el grado hasta el cual la información de la muestra puede diferir de la información de la población (OSTLE, 1977, cap. 4).

En el *muestreo no probabilístico* la selección de un elemento de la población para que forme parte de la muestra se basa, en parte, en el criterio del investigador. No existe oportunidad conocida por cualquier elemento particular de la población que se ha seleccionado. Por consiguiente no puede calcularse el grado en que la información muestral difiere de la información poblacional (HERNANDEZ et al., 1998, cap.8).

La elección de uno u otro tipo de muestras dependerá del problema en estudio. En el área de la metrología química (y en cualquier otra área de la metrología), es indispensable

presentar evidencias objetivas del grado de representatividad de la muestra, lo que conduce inevitablemente al uso del muestreo probabilístico o *aleatorio*.

En general, para la obtención de una muestra *probabilística* se necesitan básicamente dos cosas (HERNANDEZ et al., 1998, cap.8):

- a. Determinar el tamaño de la muestra (n).
- b. Seleccionar a los elementos muestrales de forma que todos tengan una probabilidad conocida de ser elegidos, justificando de esta forma la independencia estadística e implicando que el orden en que se han efectuado las mediciones no tiene influencia en los resultados.

Existen dos procedimientos de muestreo para la obtención de *muestras probabilísticas*, que se distinguen de acuerdo al tipo de población de la que provienen los datos. Cuando la población es *real* la muestra se forma aplicando *técnicas de muestreo estadístico*, que se fundamentan en el conocimiento de las características de homogeneidad o heterogeneidad de las unidades o elementos de los que se extraen los datos. Entre estas técnicas de muestreo se encuentran: el muestreo aleatorio simple, el muestreo estratificado, el muestreo por conglomerados y el muestreo sistemático (todos ellos con otras subdivisiones o variantes).

Cuando se muestrea una población *conceptual* (como es la que concierne al área del análisis químico), la forma de "extraer" una muestra aleatoria no es más que efectuar un cierto número de veces el experimento aleatorio (proceso) que genera a los elementos de la población, de forma planificada, lo que implica no sólo el conocimiento del experimento aleatorio, sino también del entorno del mismo, es decir de los factores controlables o no controlables que lo afectan. Las técnicas de extracción de este tipo de muestras se basan en el "diseño de experimentos".

En análisis químico se presentan ambos procedimientos para obtención de muestras probabilísticas. Si se efectúa el *muestreo de los materiales a analizar*, se aplica el primer procedimiento (ver capítulo 6.0). Por otra parte, si se realiza el *muestreo del proceso de medición* para investigación y verificación del mismo, se aplica el segundo procedimiento (ver capítulos 4, 5 y 7).

2.1 DEFINICION ESTADISTICA DE MUESTRA EN ANALISIS QUIMICO

En general, las muestras para análisis químico pueden definirse matemáticamente de la siguiente forma (MEYER, 1973, cap. 13):

“Sea X una variable aleatoria con cierta distribución de probabilidad. Sean X_1, X_2, \dots, X_n n variables aleatorias independientes que tienen cada una la misma distribución que X . Se llama entonces a (X_1, X_2, \dots, X_n) una muestra aleatoria de la variable X ”.

De la definición anterior pueden notarse las observaciones siguientes:

- a) Una muestra aleatoria de tamaño n de una variable aleatoria X corresponde a n mediciones repetidas de X , hechas básicamente bajo las mismas condiciones de repetitividad, con lo que se permite asegurar que la distribución de probabilidad de la variable aleatoria X permanezca la misma prueba tras prueba. En general, la noción matemáticamente idealizada de una muestra aleatoria se puede aproximar mejor sólo por las condiciones experimentales reales. A fin de que X_1, X_2, \dots, X_n tengan la misma distribución, todas las condiciones “relevantes” bajo las cuales se realiza el experimento, deben ser las mismas. Las condiciones experimentales nunca se pueden duplicar idénticamente, por supuesto, sin embargo, lo importante es que esas condiciones, que son diferentes, deben tener poco o ningún efecto en el resultado del experimento. Por ejemplo, si se quiere determinar la concentración de metales pesados en una porción de agua de desecho, la concordancia entre los resultados de mediciones del mesurando dependerá de que se traten de mantener condiciones como: principio de medición, método de medición, instrumento de medición, patrón de referencia, condiciones temperatura y humedad del laboratorio, etc.
- b) Si X es una variable aleatoria continua o discreta, (X_1, X_2, \dots, X_n) deben de ser independientes. Lo que depende de que las pruebas sucesivas para obtener una muestra sean independientes una de otra.

2.2 NOCION DE ESTADISTICO (O ESTADIGRAFO)

Si se elige una muestra formada por n elementos a partir de una población modelizada por la variable X cuya ley de probabilidad sea F , la distribución de cada X_i de la muestra será la misma que la de X , esto es, $F(X_i) = F(X)$ para todo $i = 1, 2, \dots, n$. Además las variables X_i serán independientes entre sí con lo que $F(X_1, X_2, \dots, X_n) = F(X_1).F(X_2).....F(X_n)$ (PEREZ, 1999, cap. 5).

En general, a aquella cantidad directamente asociada a la población se le conoce como *parámetro* (como por ejemplo la media), y como *estadístico o estadígrafo*, a aquella cantidad calculada partiendo de un conjunto de datos muestrales. Los *parámetros* usualmente se representan con letras griegas, mientras que los *estadísticos o estadígrafos* con letras latinas (BOX et al., 1989, cap. 2).

Un *estadístico o estadígrafo* se define como una función medible T de (X_1, X_2, \dots, X_n) . Al tratarse de una función de variables aleatorias es también una variable aleatoria cuya distribución se llama ***distribución de muestreo del estadístico***, que dependerá en general de los parámetros desconocidos de la función de distribución de la variable considerada para la población y para cuya determinación desempeña un papel fundamental la relación $F(X_1, X_2, \dots, X_n) = F(X_1).F(X_2).....F(X_n)$ (PEREZ, 1999, cap. 5).

En la figura 2.2, se esquematiza el carácter aleatorio de los estadísticos, tomando como ejemplo la media y la varianza muestral.

Esa dependencia de la distribución del estadístico, de los parámetros desconocidos y de la ley de probabilidad de la variable considerada para la población será utilizada para realizar la estimación de los parámetros poblacionales a partir del estadístico, utilizando los métodos que proporciona la inferencia estadística (ver figura 2.1) (PEREZ, 1999, cap. 5).

Los *estadísticos*, permiten describir la variabilidad de un conjunto de datos muestrales mediante una función de los mismos que no contiene a los parámetros desconocidos. Las funciones de cálculo para los estadísticos de los distintos tipos se refieren a los resúmenes numéricos que se presentan en las tablas 1.2 y 1.3 de la sección 1.5.2, referidos en este caso a datos muestrales, por lo que es importante tomar en cuenta que, el promedio de los datos se representará por \bar{X} y la varianza por S^2 (por tanto la desviación típica por S) y que las formulas de

calculo se verán modificadas de la forma en la que se presenta en las tablas 2.1 y 2.2.

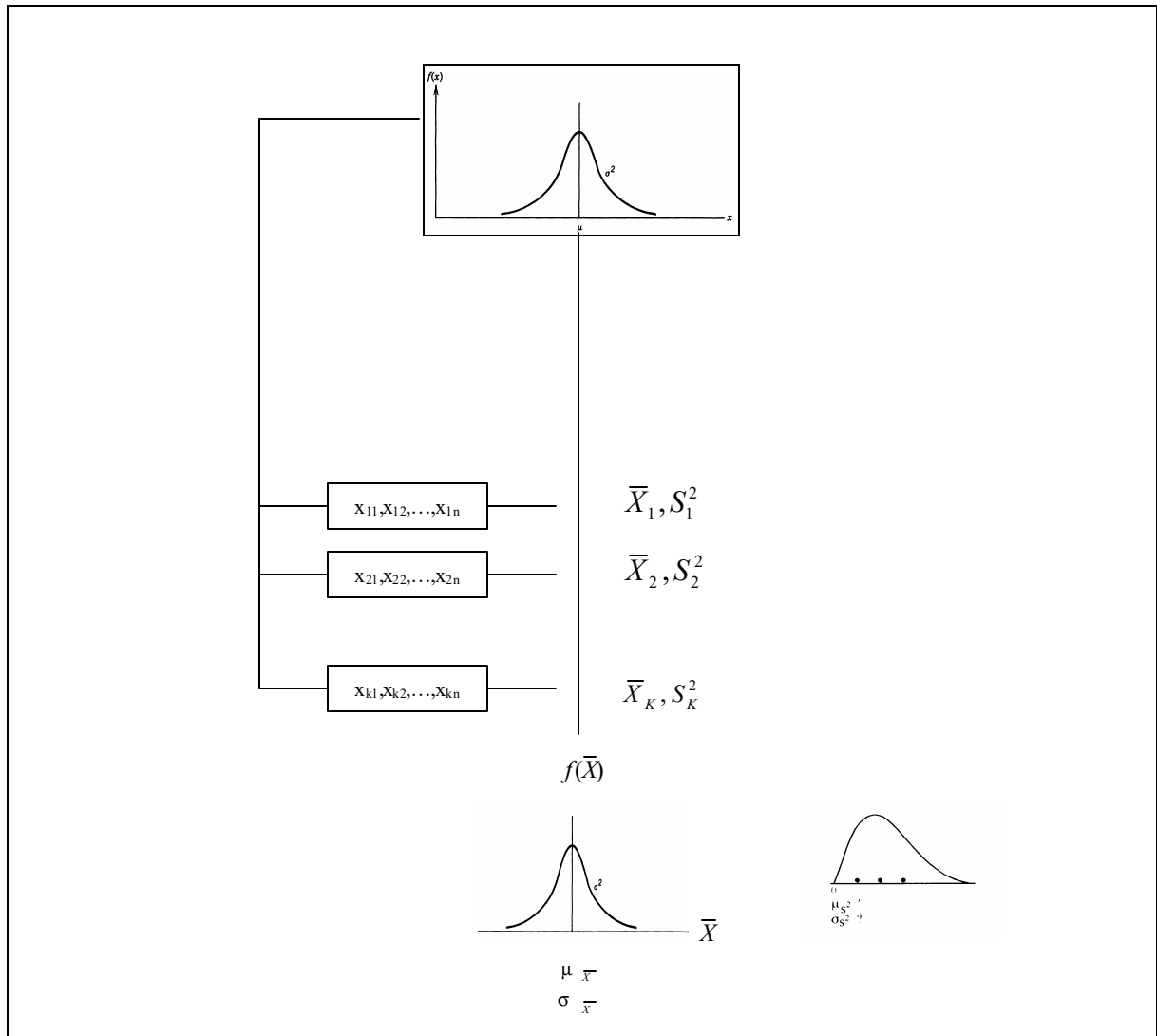


FIGURA 2.2 LOS ESTADÍSTICOS SON VARIABLES ALEATORIAS (PRAT et al. 2000, cap. 4)

TABLA 2.1 RESUMEN DE FORMULAS DE MEDIDAS DE POSICIÓN (PEREZ, 1999, cap. 3)

MEDIDAS DE TENDENCIA CENTRAL	FORMULA DE CALCULO
Media aritmética	$\bar{X} = \frac{\sum x_i}{n}$ <p> \bar{X} = promedio x_i = observaciones n = número de observaciones de la muestra </p>
mediana	$\tilde{x} = x_{[n+1/2]}$ para n impar $\tilde{x} = \frac{x_{(n/2)} + x_{((n/2)+1)}}{2}$ para n par
Cuartiles	$Q_y = \frac{n+1}{4} y$ <p>y = puede ser 1, 2 o 3</p>
Rango intercuartílico	$IQR = Q_3 - Q_1$

TABLA 2.2 FORMULAS PARA LAS MEDIDAS DE DISPERSION Y POSICION (PEREZ, 1999, cap. 3)

MEDIDAS DE DISPERSION	ABSOLUTAS	RELATIVAS
Varianza	$S^2 = \frac{\sum_{i=1}^N (x_i - \bar{X})^2}{n-1}$	
Desviación estándar	$S = \sqrt{S^2}$	Coeficiente de variación $CV = \frac{S}{\bar{X}}(100)$
Recorrido o rango	$R = \max(x_i) - \min(x_i)$	Índice de dispersión respecto a la mediana $R_{rel} = \frac{R}{\tilde{x}}$
MEDIDAS DE FORMA	COEFICIENTES	COEFICIENTES ESTANDARIZADOS
Asimetría	$CA = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^3}{S^3}$	$CA_{est} = \frac{CA}{\sqrt{6/n}}$
Curtosis	$CC = \left[\frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^4}{S^4} \right] - 3$	$CC_{est} = \frac{CC}{\sqrt{6/n}}$

Dos consideraciones respecto a los estadísticos y sus distribuciones son las siguientes (MEYER, 1973, cap. 13):

- Dado que los *estadísticos* son variables aleatorias, dos características importantes para la distribución de estos, serán la *esperanza* y la *varianza* de los estadísticos, que se conocerán como media y varianza muestral respectivamente. La *desviación estándar* de la distribución muestral de un estadístico se conoce, a veces, como *error estándar de la variable estadística*.
- La propiedad más importante de los estadísticos es *el teorema central del límite*. Este teorema se aplica tanto a la media muestral \bar{X} como a la suma muestral $(\sum_{i=1}^n X_i)$ y establece que cuando es grande el tamaño de la muestra n , la desviación muestral de la media (o de la suma) muestral tendrá aproximadamente la distribución normal.

2.3 DISTRIBUCIONES DE MUESTREO

Si se conoce la distribución de probabilidad de la población de la que se obtuvo la muestra, entonces muchas veces es posible determinar la distribución de probabilidad de diversos estadísticos calculados a partir de los datos muestrales. Como se mencionó anteriormente, la distribución de probabilidad de un estadístico se conoce como *distribución de muestreo*. Entre las distribuciones de muestreo más importantes se tienen:

- a) Distribución normal.
- b) Distribución t de Student o distribución t.
- c) Distribución Chi-cuadrado, o Gi-dos (χ^2).
- d) Distribución F de Snedecor o distribución F.

En la tabla 2.3, se presenta un resumen sobre la utilidad, propiedades y formas de las distribuciones de muestreo más importantes. En los anexos A2, B2 y C2 se presentan las tablas para las distribuciones t, Chi-cuadrado y F-Snedecor.

TABLA 2.3 RESUMEN DE LAS DISTRIBUCIONES DE MUESTREO MAS IMPORTANTES

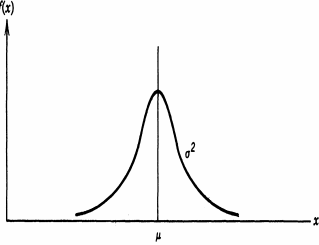
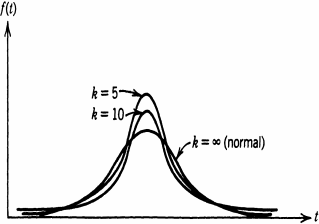
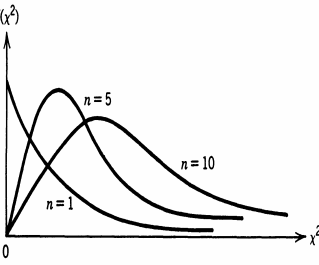
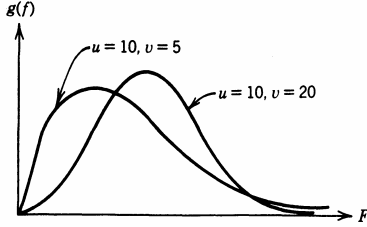
Tipo de distribución	Distribución del estadístico	Descripción	Forma
Normal	\bar{X}	<p>-Sea X una variable aleatoria distribuida normalmente, con media μ y σ^2. Sea x_1, \dots, x_n una muestra aleatoria de tamaño n tomada del proceso. Entonces la distribución de la media muestral \bar{X} es $N(\mu, \sigma^2/n)$. Esto se deduce directamente de los resultados respecto a la distribución de combinaciones lineales de variables aleatorias normales.</p> <p>-Del teorema central del límite se sabe que la distribución de $\sum_{i=1}^n X_i$ es aproximadamente normal, a medida que aumenta n (a partir de $n > 25$), con media $n\mu$ y varianza $n\sigma^2$, sin importar la distribución de la población. Por lo tanto, la distribución de muestreo de la media muestral \bar{X} es aproximadamente es $N(\mu, \sigma^2/n)$.</p> <p>-La función de distribución de la normal y la normal estándar se presentan en la tabla 1.7.</p>	
Distribución t-Student	\bar{X}	<p>-Tiene forma acampanada o de montículo y es perfectamente simétrica respecto de $t = 0$, a diferencia de la normal estándar z, es muy sensible, alargándose con rapidez hacia la izquierda o derecha.</p> <p>-El divisor $(n-1)$ de la suma de cuadrados de las desviaciones que aparece en la fórmula para S^2, es el número de grados de libertad⁽¹⁾ asociada a S^2 y a la estadística t.</p> <p>-La varianza de la t Student es algo mayor que 1, aumenta a medida que aumentan los grados de libertad.</p> <p>-El estadístico \bar{X} de una muestra de tamaño n, extraída de una población normal de media μ sigue una distribución t-Student con $n-1$ grados de libertad. Es útil para $n < 25$.</p> <p>-La función de distribución de la t-Student se presenta en la tabla 1.6.</p>	
Chi-cuadrado	S^2	<p>-Una distribución de muestreo importante, definida en términos de la distribución normal, es la distribución Chi-cuadrado, o χ^2. Si x_1, x_2, \dots, x_n son variables aleatorias normal e independientemente distribuidas con media cero y varianza uno, entonces la variable aleatoria $\chi^2 = x_1^2 + x_2^2 + \dots + x_n^2$, se distribuirá como un Chi-cuadrado, con n grados de libertad. La distribución de probabilidad de la χ^2, se presenta en la tabla 1.6.</p> <p>-La distribución es sesgada, con media n y varianza $2n$.</p> <p>-La distribución de muestreo de $(n-1)S^2/\sigma^2$ es χ_{n-1}^2, cuando se realiza el muestreo a partir de una $N(\mu, \sigma)$</p>	

TABLA 2.3 continuación...

Tipo de distribución	Distribución del estadístico	Descripción	Forma
F de Snedecor (F)	razón de varianzas	<p>-Si X_1 y X_2 son dos variables Chi-cuadrado independientes con n_1 y n_2 grados de libertad. La razón entre ellas divididas cada una por sus grados de libertad se distribuye como una F de Snedecor.</p> <p>-La función de densidad (ver tabla 1.6) es asimétrica y el coeficiente de asimetría decrece cuando aumentan los grados de libertad del numerador y denominador.</p> <p>-La distribución F se usa con el objeto de comparar la variabilidad de diferentes fuentes. Supóngase que se tienen dos procesos normales independientes, sea $X_{11}, X_{12}, \dots, X_{1n_1}$ una muestra aleatoria de tamaño n_1 observaciones a partir del primer proceso normal, y sea $X_{21}, X_{22}, \dots, X_{2n_2}$ una muestra aleatoria de tamaño n_2 tomada del segundo proceso. Si S_1^2 y S_2^2 son las varianzas muestrales, entonces el cociente</p> $\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{n_1-1, n_2-1}$	

(1) Se define el número de grados de libertad de una suma de cuadrados ($SS = \sum(x_i - \bar{x})^2$) como el número de términos independientes en $\sum(x_i - \bar{x}) = 0$, ya que solo $n-1$ elementos en la regla anterior ($x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_{n-1} - \bar{x}$), son independientes.

2.4 PRUEBAS DE NORMALIDAD

Un problema frecuente al estudiar datos reales continuos es analizar hasta qué punto la *distribución normal* resulta un modelo conveniente para describir la pauta de variabilidad de una serie de datos, puesto que pautas de variabilidad que se alejen sensiblemente de la normal pueden exigir el recurso a tratamientos estadísticos especiales o ser el síntoma de anomalías en los datos (ROMERO y ZUNICA, 1993, cap. 5).

Con el fin de estudiar si la distribución normal resulta un modelo adecuado para describir la variabilidad de una serie de datos, (lo cual es importante para la aplicación de técnicas de inferencia), puede acudirse a dos tipos de pruebas, que son (ROMERO y ZUNICA, 1993, cap. 5):

- a) *Pruebas cualitativas*: como el *análisis de histogramas*, que exige al menos de 40 a 50 datos, *análisis de gráficos en escalas probabilísticas*, aplicable a un grupo de

cinco datos a más. Es conveniente acompañar ambos análisis gráficos con el análisis de los coeficientes de curtosis y asimetría de la distribución.

- b) *Pruebas estadísticas formales*: Poco útiles en la práctica (respuesta correcta a la pregunta equivocada), como por ejemplo pueden mencionarse: la prueba Chi-cuadrado, la prueba de Kolmogorov-Smirnoff, etc.

a) *ANÁLISIS DE HISTOGRAMAS* (KUME, 1992, cap. 5): Casi siempre existe variabilidad en un proceso y, por lo general, presenta determinado comportamiento. Este comportamiento o patrón se puede representar en forma de histograma. Un histograma es una representación gráfica de la variación en un conjunto de datos. Muestra la frecuencia o número de observaciones de determinado valor, o dentro de un grupo especificado. Los histogramas proporcionan pistas acerca de las características de la población primitiva de la que se toma la muestra. Al usar un histograma, se puede ver con claridad la forma de la distribución y se pueden hacer inferencias de la población.

Como muchos procesos producen resultados que siguen razonablemente una distribución normal, cuando se combina el concepto de histograma y el concepto de curva normal, se obtiene una herramienta de trabajo práctica conocida como *análisis de histogramas*.

Se selecciona una muestra aleatoria de al menos 50 datos y se toman las medidas de la característica de calidad elegida. Se prepara el histograma y se analiza su comportamiento. El conocimiento del proceso se combina con la información del histograma para obtener conclusiones.

La interpretación consiste en responder a las siguientes preguntas:

- ¿Tiene una forma adecuada la distribución de los datos?
- ¿Está esta distribución en la posición adecuada?
- ¿Que acción si existe es la adecuada?

En la figura 2.3, se ilustran histogramas representativos que con frecuencia se encuentran en el análisis de variabilidad de procesos. El literal a) representa el caso de una distribución normal. Cualquier desviación de este comportamiento por lo general, es el resultado de una influencia externa al proceso, que se debe investigar. Un patrón como el

b) con valores altos y bajos de forma alternada, es causado con frecuencia por error sistemático de la medición, en la manera de agrupar los datos, o sesgo debido al redondeo de los valores de los datos. Un patrón cargado hacia un lado como el de c) es como de forma de campana, pero no es simétrico; la distribución se desvanece en una dirección. Las figuras sesgadas surgen cuando hay un límite natural de los valores de los datos o cuando la distribución de variabilidad de los datos no es normal. Un patrón como el de d) se presenta cuando las asimetrías (positivas o negativas) se vuelven extremas. Un comportamiento uniforme como el de e) muestra una mayor variabilidad que en a), sin tendencia central. A menudo este es el resultado al combinar datos de muchos procesos en forma de campana con diversos centros entre los límites de los datos. Un patrón bimodal como en los casos f) y g) sugiere que se combinan dos grupos de mediciones en forma de campana. Por lo general se desea aislar los procesos o condiciones individuales que provocan este comportamiento.

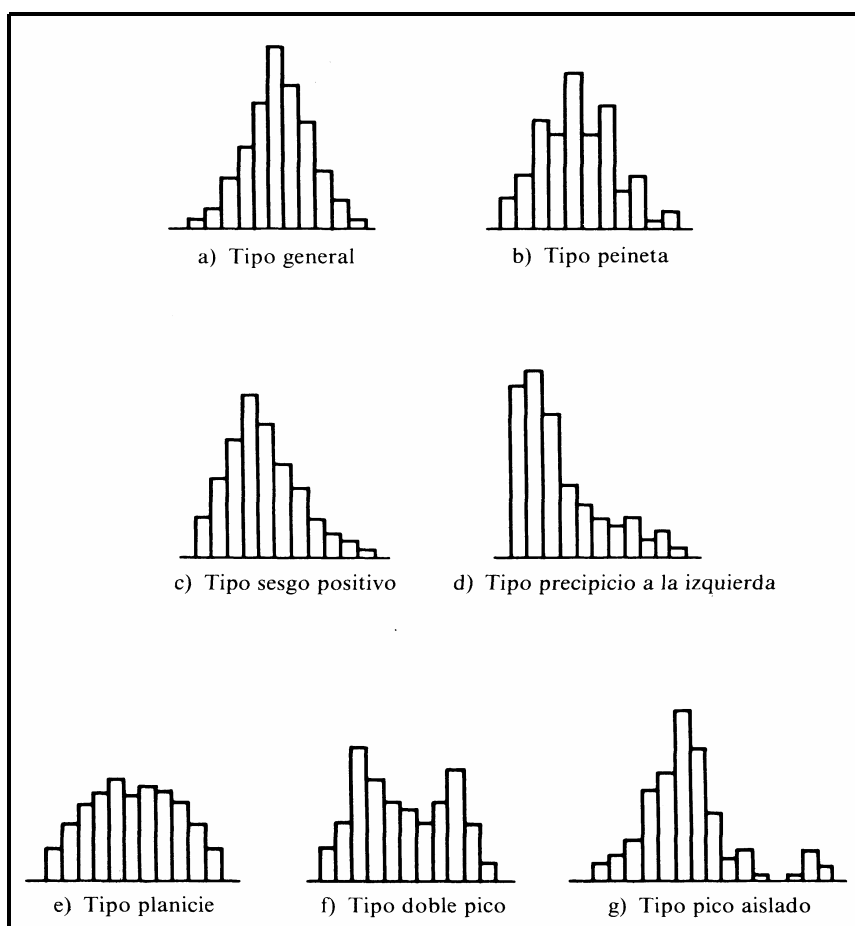


FIGURA 2.3 TIPOS DE HISTOGRAMAS (KUME, 1992, cap. 5)

b) GRAFICAS EN ESCALA PROBABILISTICA NORMAL: La escala probabilística (o papel probabilístico) constituye una herramienta extremadamente práctica de análisis estadístico, utilizándose en el estudio de distintos tipos de distribuciones. En este apartado se tratará solo de la escala probabilística para distribuciones normales. Una representación de un conjunto de datos en escala probabilística hace corresponder un punto a cada observación. La abcisa del punto no es más que el valor observado, mientras que la ordenada corresponde al porcentaje de valores en la muestra que son menores o iguales que el considerado. Los puntos se representan en una escala como la de la figura 2.4. La idea básica para la utilización de la escala probabilística es la siguiente: *cuando datos procedentes de una distribución normal se representan en esta escala, los puntos correspondientes se sitúan aproximadamente a lo largo de una recta*. Si la representación de los datos difiere claramente de una línea recta ello es una prueba de que la distribución de variabilidad no se describe adecuadamente con el modelo normal (ROMERO y ZUNICA, 1993, cap. 5).

Las gráficas de probabilidad normal requieren menos datos que un histograma para identificar normalidad, por lo que es necesario tener cierta experiencia para interpretar los resultados adecuadamente. En la figura 2.5, se observan aspectos típicos de representaciones correspondientes a datos no normales. Las distribuciones sesgadas se pueden identificar por la curvatura que presentan en las gráficas de probabilidad normal (ver figura 2.5 a y b). Es importante señalar que, bajo algunas circunstancias, es posible observar curvaturas en las distribuciones normales, pero no tan pronunciadas como las mencionadas anteriormente. Por otra parte, las observaciones aberrantes se presentan como puntos que están muy alejados del comportamiento lineal (ver figura 2.5c). Y finalmente, comportamientos como el de la figura 2.5d, pueden representar la mezcla de dos poblaciones con diferentes medias.

c) ANALISIS DE LOS COEFICIENTES DE ASIMETRIA Y CURTOSIS: Si unos datos son simétricos lo son respecto a su media, y la suma de los cubos de sus desviaciones respecto a dicha media $\sum(x-\bar{X})^3$ será nula. Por el contrario, dicha suma será positiva, si los datos presentan una cola alargada hacia la derecha y negativa si la presentan hacia la izquierda. Para representar el grado de asimetría de un conjunto de datos se emplea el *coeficiente de asimetría (CA)* (ver sección 1.5.2), si $CA = 0$, la distribución es simétrica, si

CA > 0 la distribución es asimétrica positiva, y si CA < 0 la distribución es asimétrica negativa (ROMERO y ZUNICA, 1993, cap. 5).

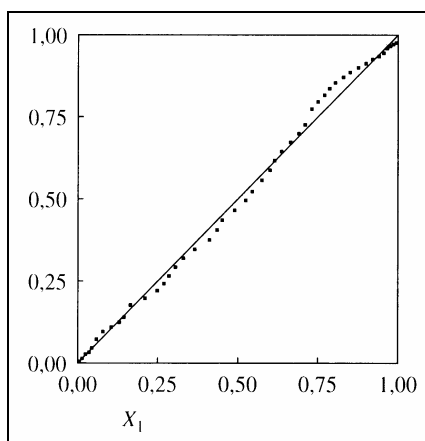


FIGURA 2.4 DISTRIBUCION NORMAL

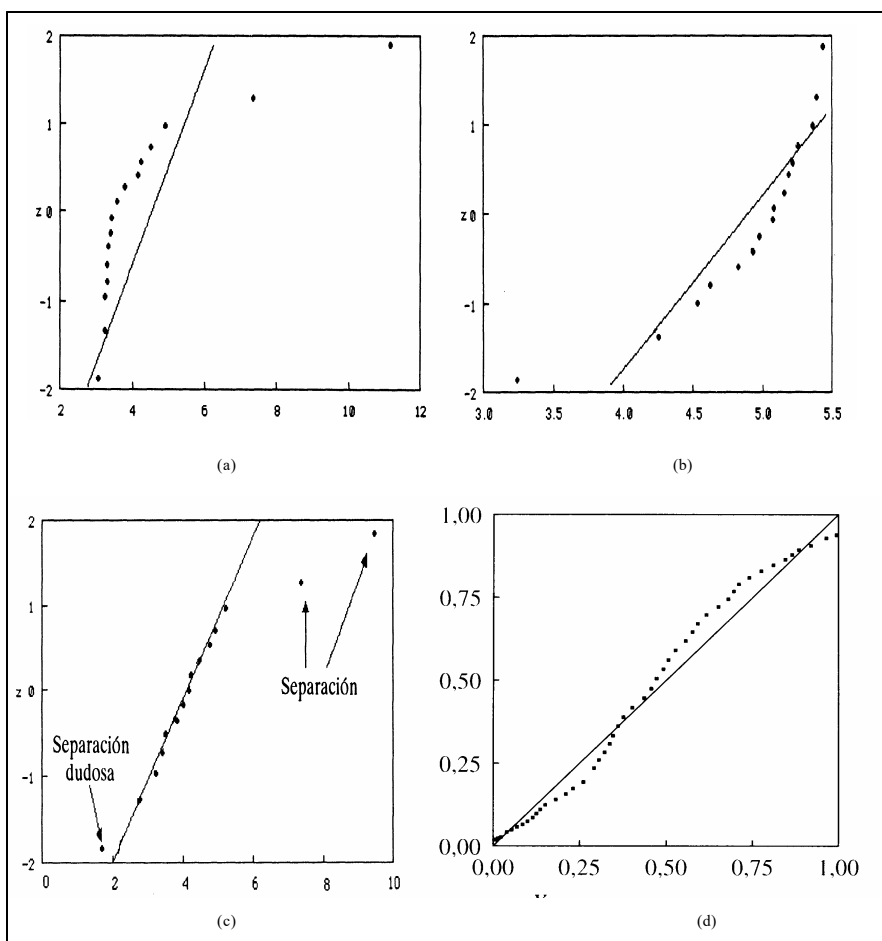


FIGURA 2.5 REPRESENTACION TIPICA DE DATOS NO NORMALES

Dado que el CA de forma poco frecuente será exactamente cero, es necesario contar con otro estadístico que permita considerar adecuado al modelo de la distribución normal para describir la simetría de una serie de datos, este indicador es el *coeficiente de asimetría estandarizado* (CA_{est}) (este coeficiente es asintóticamente normal (0, 1)), el criterio establece que si el CA_{est} se encuentra en el intervalo de [-2, 2] puede considerarse que la distribución no tiene desviaciones significativas en cuanto a la simetría de la *distribución normal* (PEREZ, 1999, cap. 3). Las fórmulas para el cálculo de los coeficientes de asimetría antes mencionados, son las siguientes:

$$CA = \frac{\frac{1}{n-1} \sum_{i=1}^N (x_i - \bar{x})^3}{s^3} \quad \left| \quad CA_{est} = \frac{CA}{\sqrt{6/n}} \right.$$

En la figura 2.6 se reflejan los histogramas posibles para los tipos de asimetría (facilitando su representación usando curvas continuas):

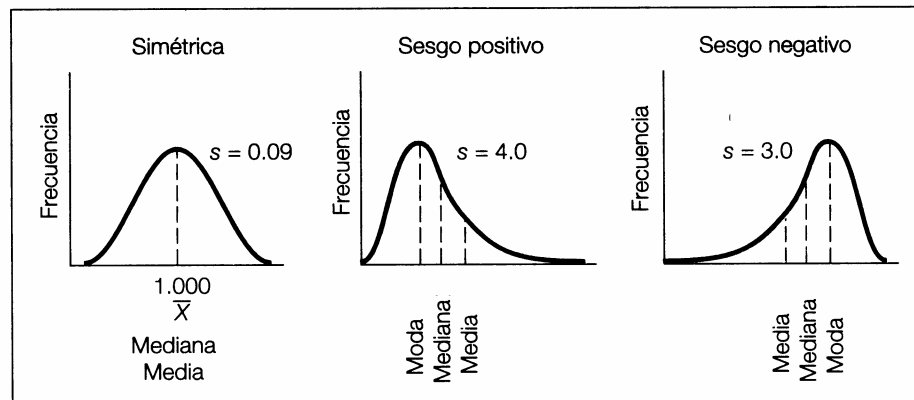


FIGURA 2.6 POLIGONOS DE FRECUENCIA QUE ILUSTRAN LA ASIMETRÍA (MASON et al., 2001, cap. 7)

Por otra parte, se dice que un conjunto de datos es leptocúrtico si presenta valores muy alejados de la media con mayor frecuencia de la que cabría esperar para unos datos normales que tuvieran la misma desviación típica. Obviamente, para compensar estos valores extremos un histograma de datos leptocúrticos es más apuntado en las cercanías de la media de lo que sería el de unos datos normales con la misma desviación típica. Frecuentemente, valores elevados de curtosis de un conjunto de datos suele ser síntoma de que entre los mismos se incluyen observaciones o resultados anómalos (ROMERO y ZUNICA, 1993, cap. 5).

En el otro sentido unos datos se denominan planicúrticos si los valores alejados de la media aparecen con una frecuencia menor que la que cabría esperar si los datos siguieran una distribución normal con la misma desviación típica. Para compensar este hecho, el histograma de unos datos planicúrticos aparece más plano en el entorno de la media de lo que sería el de unos datos normales con idéntica desviación típica. Una planicurtosis excesiva puede revelar que los datos han sido artificialmente censurados para eliminar los valores considerados extremos. El grado de curtosis de un conjunto de datos se mide mediante el *coeficiente de curtosis (CC)*, en datos que siguen una distribución exactamente normal CC es igual a 3, en datos leptocúrticos $CC > 3$, y en datos planicúrticos $CC < 3$ (ROMERO y ZUNICA, 1993, cap. 5).

Al igual que en el caso del coeficiente de asimetría, existe un *coeficiente de curtosis estandarizado (CC_{est})*, el criterio establece que si el CC_{est} se encuentra en el intervalo de $[-2, 2]$ puede considerarse que la distribución no tiene desviaciones significativas en cuanto a la curtosis de la *distribución normal* (PEREZ, 1999, cap. 3). Las fórmulas para el cálculo de los coeficientes de curtosis normal y estandarizado, son las siguientes:

$$CC = \left[\frac{1}{n-1} \sum_{i=1}^N (x_i - \bar{x})^4 \right] / s^4 - 3$$

(en este caso CC de la distribución normal tiene el valor de 0)

$$CC_{est} = \frac{CC}{\sqrt{6/n}}$$

En la figura 2.7, se presentan histogramas (sustituidos por curvas continuas) con idénticas medias y desviaciones típicas pero difiriendo en curtosis.

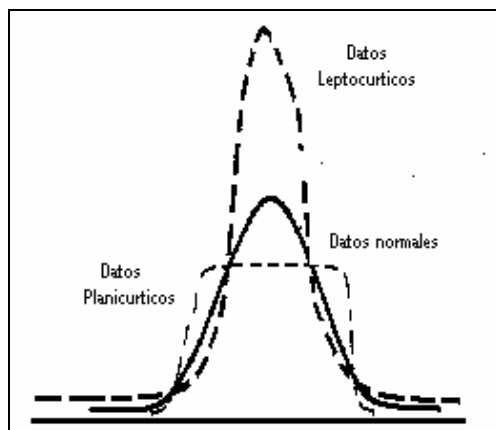


FIGURA 2.7 REPRESENTACION DE LA CURTOSIS (ROMERO y ZUNICA, 1993, cap. 5)

2.4.1 TRANSFORMACION DE DATOS

A veces un conjunto de datos no se ajusta a una de las distribuciones habituales, tal como la distribución normal. Sin embargo, si se aplica una transformación a la característica original esta se transforma en una nueva variable que se distribuye normalmente. En la figura 2.8 se resumen varias de estas transformaciones matemáticas. Estas transformaciones son útiles para: (a) lograr la normalidad de los datos medidos; (b) satisfacer el supuesto de varianzas muestrales iguales, requerido en ciertas pruebas; y (c) satisfacer el supuesto de aditividad de efectos, necesario en determinados ensayos (JURAN y GRANA, 1993, Tomo 2, cap. 23).

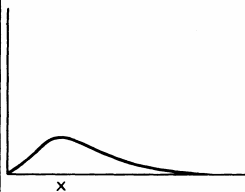
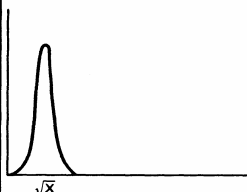
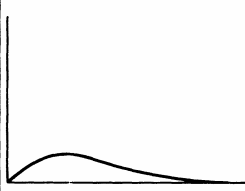
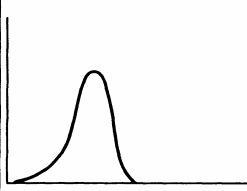
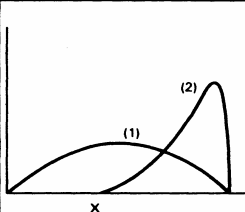
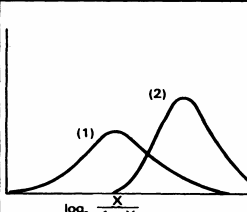
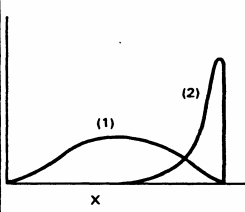
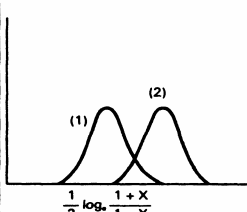
TRANSFORMACION	RECORRIDO DE LA VARIABLE	DISTRIBUCION ORIGINAL	DISTRIBUCION TRANSFORMADA
\sqrt{x}	$0 \leq x < \infty$		
$\log_e x$ o $\log_{10} x$	$0 < x < \infty$		
$\log_e \frac{x}{1-x}$ o $\log_{10} \frac{x}{1-x}$	$0 < x < 1$		
$\frac{1}{2} \log_e \frac{1+x}{1-x}$ o $\log_{10} \frac{1+x}{1-x}$	$-1 < x < 1$		

FIGURA 2.8 RESUMEN DE ALGUNAS DISTRIBUCIONES (JURAN y GRANA, 1993, Tomo 2, cap. 23)

Entre las transformaciones más frecuentes se encuentran (JURAN y GRZYNA, 1993, Tomo 2, cap. 23):

$$\begin{aligned}\xi_1(X_1) &= \sqrt{X_1 - a} \\ \xi_2(X_1) &= X_1^{1/3} \\ \xi_3(X_1) &= \log_{10}(X_1) \\ \xi_4(X_1) &= \arcsin \sqrt{X_1} \\ \xi_5(X_1) &= \sinh^{-1} \sqrt{X_1}\end{aligned}$$

Si una de ellas, digamos $\xi(x_1)$, esta normalmente distribuida, las estimaciones de la media y la varianza de $Y_i = \xi(x_i)$, vendrán dadas por:

$$\bar{Y} = \sum_{j=1}^n \frac{Y_j}{n} \quad s_Y^2 = \sum_{j=1}^n \frac{(Y_j - \bar{Y})^2}{n-1}$$

2.5 EJERCICIOS DE APLICACIÓN DE DESCRIPCIÓN DE LA VARIABILIDAD Y MUESTRAS

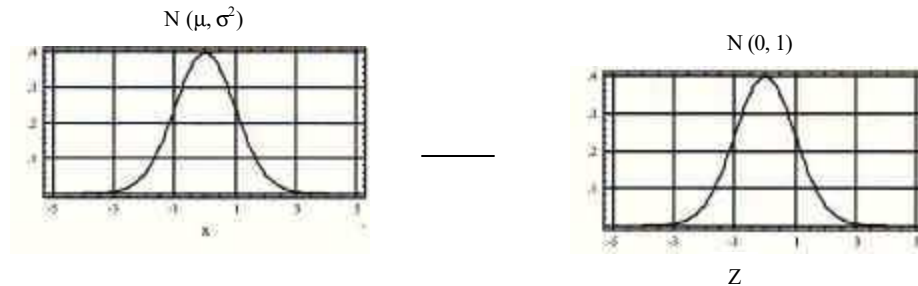
Los problemas que se presentan a continuación, se han resuelto a través de cálculos manuales, cálculos asistidos mediante el modo stat de calculadoras científicas y mediante el uso de hojas de cálculo y funciones estadísticas incluidas en dichas hojas de cálculo.

1. La resistencia a la tensión de un papel utilizado en la fabricación de bolsas para el transporte de cierto material es una característica de calidad importante. Se sabe que la resistencia (X) tiene una distribución normal, con media $\mu = 40 \text{ lb/plg}^2$ y desviación estándar $\sigma = 2 \text{ lb/plg}^2$, lo que se denota por $x \sim N(\mu, \sigma^2)$. El comprador de las bolsas necesita una resistencia de por lo menos 35 lb/plg^2 . Determine la probabilidad de que una bolsa producida con este papel satisfaga o exceda esta especificación es decir $P\{x \geq 35\}$.

OBJETIVO DEL PROBLEMA: Determinar la probabilidad de que una bolsa producida con este papel satisfaga la especificación es decir $P(x \geq 35)$.

SOLUCIÓN: Conociendo que la variable aleatoria, resistencia a la tensión, se distribuye

como una normal de media 40 lb/plg² y varianza 4 lb/plg², ($x \sim N(40,4)$), es posible encontrar la probabilidad pedida al desarrollar la integración de la función de densidad para la normal en el intervalo correspondiente ¡LO CUAL ES MUY COMPLICADO! Sin embargo, se cuenta con el concepto de la distribución normal estándar que nos permitirá evaluar dicha probabilidad a partir de tablas, ver figura. Es decir,

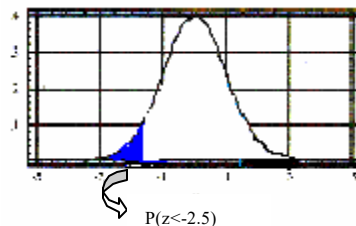


Según las tablas con las que se cuenta esta probabilidad deberá evaluarse como:

$$P(x \geq 35) = 1 - P(x \leq 35)$$

Ahora, estandarizando la distribución para buscar en la tabla, se tiene:

$$P(x \leq 35) = P\{Z \leq (35-40)/2\} = P(Z \leq -2.5) \leftarrow \text{buscando en tablas}$$



es decir, $P(Z \leq -2.5) = 0.0062$, por tanto:

$$P(x \geq 35) = 1 - P(x \leq 35) = 1 - 0.0062 = 0.9938$$

De esta forma se tiene una probabilidad del 99.38% para cumplir con las especificaciones del comprador, ó, en otras palabras sólo el 0.62% de las bolsas que se adquieran podría no cumplir con la especificación.

2. Se realizó la calibración de una pipeta de 10 ml, a través del siguiente procedimiento: Se pesaron un matraz volumétrico pequeño y su tapón. Con la pipeta que se quería calibrar, se transfirieron 10 ml de agua al matraz, y nuevamente se pesó el matraz y su tapón pero ahora conteniendo el agua. También se midió la temperatura del agua para saber su densidad. La masa del agua se calculó como la diferencia entre las dos masas, y esta diferencia se dividió entre la densidad del agua para conocer el volumen vertido por la pipeta. El experimento se repitió 50 veces. En la tabla E22.1 se presentan los resultados obtenidos. Para estos resultados: a) Construya la tabla de distribución de frecuencias y elabore el histograma. ¿Qué opinión tiene respecto a la variabilidad de los datos? (respecto a la pauta de variabilidad normal). b) Determine los estadísticos siguientes: media, moda, varianza, desviación típica, coeficiente de variación, rango. c) ¿Cuáles son las posibles causas de variabilidad de los resultados?

OBJETIVO DEL PROBLEMA: Describir la variabilidad del volumen desalojado por una pipeta de 10ml, gráficamente mediante un histograma y numéricamente calculando los estadísticos de mayor relevancia.

TABLA E22.1 MEDICIONES REPETIDAS EN LA CALIBRACIÓN DE UNA PIPETA DE 10 ml

Ensayo	Volumen, ml	Ensayo	Volumen, ml	Ensayo	Volumen, ml
1	9.988	18	9.975	35	9.976
2	9.973	19	9.980	36	9.990
3	9.986	20	9.994	37	9.988
4	9.980	21	9.992	38	9.971
5	9.975	22	9.984	39	9.986
6	9.982	23	9.981	40	9.978
7	9.986	24	9.987	41	9.986
8	9.982	25	9.978	42	9.982
9	9.981	26	9.983	43	9.977
10	9.990	27	9.982	44	9.977
11	9.980	28	9.991	45	9.986
12	9.989	29	9.981	46	9.978
13	9.978	30	9.969	47	9.983
14	9.971	31	9.985	48	9.980
15	9.982	32	9.977	49	9.983
16	9.983	33	9.976	50	9.979
17	9.988	34	9.983		

SOLUCIÓN:

a) Para la construcción del histograma, hay que partir de la elaboración de una tabla de distribución de frecuencias. Una de las observaciones que deben tomarse al construir el histograma es comparar la pauta de variabilidad de la variable en cuestión con la pauta de variabilidad descrita para una variable que se distribuye como una normal, lo cual en este caso puede hacerse porque se cuenta con cincuenta datos (algunos autores consideran suficientes 40 datos para hacer esta comparación). A continuación se presenta el desarrollo del problema:

(i) *Tabla de distribución de frecuencias*

Para construir la tabla de distribución de frecuencias, debe calcularse primero el rango de los datos, el cual viene dado por:

$$\begin{aligned} \text{Rango o dispersión: } R &= \text{valor máximo} - \text{valor mínimo} = 0.994 - 0.969 \\ R &= 0.025 \text{ ml} \end{aligned}$$

Optar por un número de intervalos (k): Para mejores resultados el número mínimo de intervalos de clase debe ser cercano a la raíz cuadrada del número de observaciones. Para el caso se construirán 9 intervalos.

La amplitud de los intervalos puede calcularse a partir de:

$$h = R/k = 0.025/9 = 0.0028$$

A continuación se fijan los límites de los intervalos, y se indica el número de veces que aparecen datos dentro de cada uno de los intervalos definidos. Posteriormente, se determina la frecuencia para cada intervalo como:

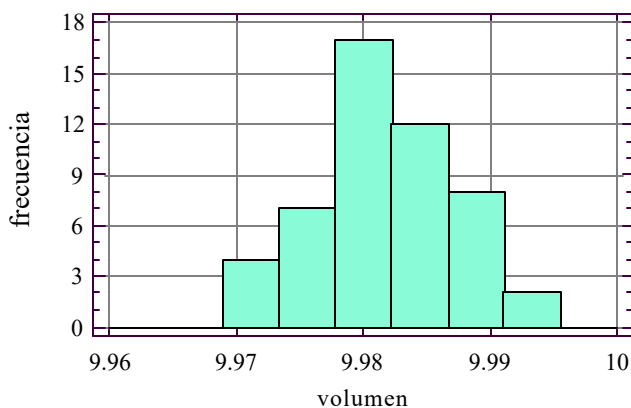
$$\% \text{ frecuencia} = \frac{\# \text{ de conteos por intervalo}}{n} \times 100$$

En la tabla E22.2, se presenta el resumen de todos estos cálculos:

TABLA E2.2 TABLA DE DISTRIBUCION DE FRACUENCIAS

Intervalo de volumen, ml	Frecuencia en el intervalo	% de frecuencia
9.969 a 9.971	3	6
9.972 a 9.974	1	2
9.975 a 9.977	7	14
9.978 a 9.980	9	18
9.981 a 9.983	13	26
9.984 a 9.986	7	14
9.987 a 9.989	5	10
9.990 a 9.992	4	8
9.993 a 9.995	1	2

- (ii) *Histograma:* a continuación se presenta el histograma construido ¿Qué opina respecto a la pauta de variabilidad del volumen desalojado? ¿Cómo interpreta al histograma?



Como puede observarse la distribución de variabilidad de los volúmenes desalojados tiene una tendencia del tipo general, tal como el gráfico del literal a) de la figura 2.3.

b) *Cálculo de estadísticos:*

$$\text{Volumen Promedio: } \bar{V} = \left(\sum_{i=1}^n V_i \right) / n = (9.988 + \dots + 9.979) / 50 = 9.982 \text{ ml}$$

Mediana de los volúmenes: esta se obtiene al ordenar los datos de menor a mayor y luego sacar el promedio entre los dos datos centrales, el ordenamiento

de los datos se puede efectuar usando una hoja de calculo como Excel, usando el comando "sort ascending".

$$\text{Mediana} = V_{\text{mediana}} = 9.982\text{ml}$$

Moda de los volúmenes: esta se obtiene contando el número de datos que más se repite, para el ejemplo esta tiene el valor de:

$$\text{Moda} = V_{\text{moda}} = 9.986 \text{ ml}$$

Rango o dispersión: $R = 0.994 - 0.969 = 0.025 \text{ ml}$

Desviación típica: $s = 0.0056 \text{ ml}$

Coefficiente de variación (CV): $CV = 0.000565$

En general, la determinación de los estadísticos como la media, la desviación típica, el rango, la mediana, los coeficientes de curtosis y asimetría puede facilitarse usando una hoja de calculo. A continuación se presenta un hipervínculo con Excel que permite el cálculo de estadísticos para esta variable:



c) Las fuentes de incertidumbre en la calibración de la pipeta pueden ser:

1. La apreciación visual en el nivel del agua con respecto a la marca en la pipeta.
2. Las fluctuaciones en la temperatura que afectan, al volumen de la pipeta, la viscosidad del líquido y el funcionamiento de la balanza.
3. Vibraciones y falta de nitidez que ocasionan pequeñas variaciones en la lectura de la balanza.

¡Es difícil determinar cuantos más efectos influyen y como influyen!

3. En el desarrollo de sistemas de gestión de aguas residuales suele ser necesario determinar las características estadísticas de los caudales. Se pide determinar las características estadísticas de los siguientes datos trimestrales de caudales obtenidos de los vertidos correspondientes a ciertas explotaciones industriales. En la tabla E23.1, se presentan los datos.

TABLA E23.1

Semana número	Caudal (m ³ /semana)	Semana número	Caudal (m ³ /semana)
1	768	8	971
2	803	9	1007
3	985	10	912
4	888	11	863
5	996	12	840
6	1078	13	828
7	1061		

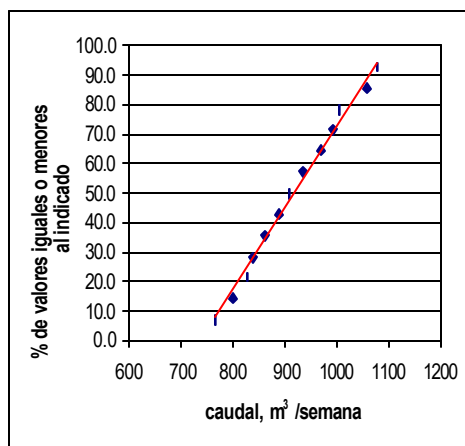
OBJETIVO DEL PROBLEMA: Caracterizar estadísticamente a la variable aleatoria caudal de agua residual semanal.

- a. Determinar gráficamente mediante la representación de los datos en papel de probabilidad, si la distribución de los mismos corresponde a una distribución normal o sesgada (log-normal).
- i. Confeccionar una tabla de análisis de los datos con tres columnas como la descrita a continuación:
1. En la primera columna, la numeración de los datos.
 2. En la segunda columna, los datos ordenados en orden creciente.
 3. En la tercera columna, su probabilidad asociada, $p = [m/(n+1)]100$, con n = número total de datos y m = número de orden del dato correspondiente. Se usa el valor de $(n + 1)$ en lugar de n para considerar la posibilidad de que se den caudales o bien superiores al máximo de los datos, o bien inferiores al mínimo de ellos.

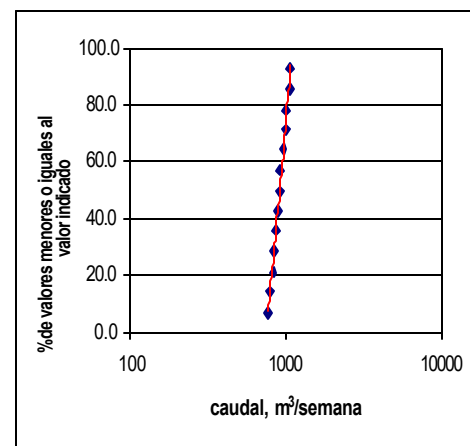
Número de orden	Caudal m ³ /semana	Probabilidad
1	768	7.1
2	803	14.3
3	828	21.4
4	840	28.6
5	863	35.7
6	888	42.9
7	912	50.0
8	936	57.1
9	971	64.3
10	996	71.4
11	1007	78.6
12	1061	85.7
13	1078	92.9

- b. Representar los caudales semanales y sus probabilidades tanto en papel aritmético como logarítmico. A continuación se adjuntan las gráficas E23.1 y E23.2 correspondientes. Debido a que ambas gráficas se ajustan bien con una recta, se pueden adoptar para los datos ambos tipos de distribución, lo que indica que la distribución no es marcadamente sesgada, y que se pueden aplicar los métodos estadísticos habituales para la distribución normal.

GRAFI CA E23.1



GRAFI CA E23.2



A continuación se presenta la determinación de las características estadísticas de los datos de los caudales, mediante la confección de una hoja de cálculo:

caudal	(x-media)	(x-media) ²	(x-media) ³	(x-media) ⁴
768	-151	22801	-3442951	519885601
803	-116	13456	-1560896	181063936
828	-91	8281	-753571	68574961
840	-79	6241	-493039	38950081
863	-56	3136	-175616	9834496
888	-31	961	-29791	923521
912	-7	49	-343	2401
936	17	289	4913	83521
971	52	2704	140608	7311616
996	77	5929	456533	35153041
1007	88	7744	681472	59969536
1061	142	20164	2863288	406586896
1078	159	25281	4019679	639128961
11951		117036	1710286	1967468568

1. Promedio:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{11951}{13} = 919 \text{ m}^3/\text{semana}$$

2. Mediana: al ser 13 datos, la mediana será el dato número 7, es decir:

$$\text{mediana} = 912 \text{ m}^3/\text{semana}$$

3. Moda:

$$\text{Moda} = 3(\text{mediana}) - 2(\text{media}) = 3(912) - 2(919) = 898 \text{ m}^3/\text{semana}$$

4. Desviación típica:

$$s = \sqrt{\frac{\sum (x - \bar{X})^2}{n-1}} = \sqrt{\frac{117036}{13-1}} = 98.76 \text{ m}^3/\text{semana}$$

5. %CV (desviación estándar relativa)

$$\%CV = \frac{100s}{\bar{X}} = \frac{100(98.76)}{919} = 10.75\%$$

6. Coeficiente de asimetría:

$CA = \frac{1}{n-1} \sum_{i=1}^N (x_i - \bar{x})^3 / s^3$ $CA = \frac{1710286/12}{98.76^3}$ $CA = 0.14$	$CA_{est} = \frac{CA}{\sqrt{6/n}}$ $CA_{est} = \frac{0.14}{\sqrt{6/13}}$ $CA_{est} = 0.206$
---	---

7. Coeficiente de curtosis:

$CC = \left[\frac{1}{n-1} \sum_{i=1}^N (x_i - \bar{x})^4 / s^4 \right] - 3$ $CA = \frac{1967468568/12}{98.76^4} - 3$ $CA = -1.28$	$CC_{est} = \frac{CC}{\sqrt{6/n}}$ $CA_{est} = \frac{-1.28}{\sqrt{6/13}}$ $CA_{est} = -1.88$
--	--

Observando estos datos se puede ver que esta distribución es ligeramente asimétrica (coeficiente de asimetría 0.14 frente a 0 de la normal), y que es bastante más aplanada que la normal (coeficiente de curtosis -1.28, frente a 0 de la normal). Si se analizan los coeficientes de asimetría y curtosis estandarizados, ambos caen dentro del intervalo de -2 a 2, por lo tanto puede considerarse que las desviaciones no son muy severas respecto a la normal.

4. Se analizaron en seis botellas el azúcar residual, obteniéndose los resultados de la tabla E24.1. a) Evaluar la desviación estándar s , para cada serie de datos; b) Agrupe los datos para establecer una desviación estándar combinada para el método.

TABLA E24.1

Frasco	Número de muestras medidas	% p/ v de azúcar residual
1	3	0.99, 1.04, 1.02
2	4	1.14, 1.13, 1.17, 1.14
3	5	1.25, 1.32, 1.27, 1.2, 1.28
4	4	0.72, 0.77, 0.73, 0.76
5	3	0.9, 0.92, 0.93
6	4	0.82, 0.88, 0.8, 0.79

OBJETIVO DEL PROBLEMA: Obtener un estimado ponderado de la desviación estándar, lo cual es útil cuando los análisis toman mucho tiempo. Esta desviación estándar ponderada se obtiene a partir de datos de una serie de muestras similares que se agrupan, permitiendo un valor de s para la característica que es mejor que el valor estimado para cualquier subserie individual.

SOLUCION: La desviación estándar ponderada se calcula a partir de la siguiente fórmula.

$$S_{PONDERADA} = \sqrt{\frac{\sum (X_i - \bar{X}_1)^2 + \sum (X_i - \bar{X}_2)^2 + \dots}{N_1 + N_2 + \dots - N_s}}$$

Además,

$$\sum_{i=1}^n (x_i - \bar{x})^2 = s^2 (n-1)$$

DONDE: N_1 , es el número de datos en la serie 1, N_2 , el número de datos en la serie 2, y así sucesivamente. El término N_s es el número de series de datos que se agrupan.

Para las diferentes muestras se tienen:

Frasco	% p/ v de azúcar residual promedio	Suma de los cuadrados de las desviaciones respecto a la media
1	1.016	0.00130
2	1.145	0.00090
3	1.264	0.00772
4	0.745	0.00170
5	0.917	0.00047
6	0.823	0.00488

Los valores de las columnas dos y tres para la muestra 1 se calcularon de la siguiente forma:

	X_i	$(X_i - \bar{X}_1)$	$(X_i - \bar{X}_1)^2$
	0.99	0.03	0.0009
	1.04	0.02	0.0004
	1.02	0.00	0.0000
Suma	3.05		0.0013
Promedio	$3.05 / 3 = 1.02$		

Así para los datos del problema tenemos:

$$S_{PONDERADA} = \sqrt{\frac{0.0013 + 0.0009 + 0.00772 + 0.0017 + 0.00047 + 0.0048}{23 - 6}}$$

$$S_{PONDERADA} = 0.032 \% \text{ p/ v}$$

La desviación estándar ponderada puede calcularse además valiéndose del uso de una hoja de cálculo como la que se presenta a continuación:

serie	1	2	3	4	5	6		
x1	0.99	1.14	1.25	0.72	0.9	0.82		
x2	1.04	1.13	1.32	0.77	0.92	0.88		
x3	1.02	1.17	1.27	0.73	0.93	0.8		
x4		1.14	1.2	0.76		0.79		
x5			1.28					
x6								
s ²	0.000633	0.0003	0.00193	0.000567	0.000233	0.001625	Conteo	Suma
n	3	4	5	4	3	4	6	23
n-1	2	3	4	3	2	3		
s ² (n-1)	0.001267	0.0009	0.00772	0.0017	0.000467	0.004875		0.016928
							sp=	0.032

5. Calcular la desviación estándar del resultado:

$$y = \frac{1.43(\pm 0.02)x10^{-2} - 4.76(\pm 0.06)x10^{-3}}{24.3(\pm 0.7) + 8.06(\pm 0.08)}$$

$$y = 2.948 \times 10^{-4}$$

(Los números entre paréntesis son las desviaciones estándar absolutas)

OBJETIVO DEL PROBLEMA: Estudiar las propiedades de la desviación estándar cuando se calculan resultados.

SOLUCION: Con frecuencia es necesario estimar la desviación estándar de un resultado que se ha calculado a partir de dos o más datos experimentales (ver capítulo 8), donde cada uno tiene una desviación estándar conocida. La forma de hacer estos estimados depende del tipo de operaciones aritméticas empleadas, así para distintos tipos de cálculos se tiene:

Tipo de calculo	Ejemplo	Desviación estándar de x
Suma o resta	$X = p+q-r$	$S_x = \sqrt{s_p^2 + s_q^2 + s_r^2}$
Multiplicación o división	$X = pq/r$	$S_x = \sqrt{(s_p/p)^2 + (s_q/q)^2 + (s_r/r)^2}$
Elevar a una potencia	$X = p^y$	$\frac{S_x}{x} = y \frac{S_p}{p}$

Logaritmo	$X = \log_{10} p$	$S_x = 0.434 \frac{S_p}{p}$
Antilogaritmo	$X = \text{antilog}_{10} p$	$\frac{S_x}{x} = 2.303 S_p$

Para el ejemplo, se tienen combinaciones aditivas y multiplicativas, por lo que se recomienda trabajar por separado las combinaciones aditivas en el numerador y denominador, y posteriormente trabajar con las combinaciones multiplicativas.

$$\text{Para el numerador: } S_{\text{numerador}} = ((0.02E-2)^2 + (0.06E-3)^2)^{0.5} = 4.36E-8$$

$$\text{Para el denominador: } S_{\text{denominador}} = ((0.7)^2 + (0.08)^2)^{0.5} = 0.70$$

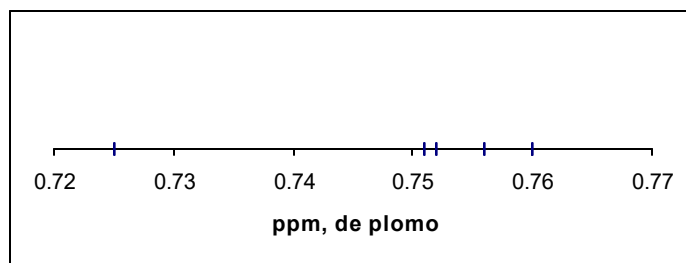
$$\text{COMBINANDO: } S_x/x = ((4.36E-8/9.5E-3)^2 + (0.7/32.46)^2)^{0.5} = 0.02$$

$$\mathbf{S_x = 2.948E-4 * 0.02 = 5.9E-6}$$

6. Se hicieron varios análisis repetidos de una muestra de sangre para determinar su contenido de plomo, y se obtuvieron los siguientes resultados: 0.725, 0.756, 0.752, 0.751, y 0.760 ppm de Pb. a) Analice la variabilidad de los datos; b) calcule el promedio, la varianza, la desviación estándar, el coeficiente de variación, la desviación relativa en partes por mil ($DER = s/\bar{X} * 1000$).

SOLUCION:

- a) Dado que el número de datos es de cinco, un recurso adecuado para analizar descriptivamente la variabilidad de los resultados es la construcción de un diagrama de puntos como el de la siguiente figura:



Del diagrama, puede observarse que existe un punto alejado de la serie como es el correspondiente a 0.725 ppm. Por otra parte, la mayor parte de puntos se localiza en el intervalo de 0.75 a 0.76 ppm.

b) Los estadísticos de la serie de datos son los siguientes.

Promedio:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{0.725 + 0.756 + 0.752 + 0.751 + 0.76}{5} = 0.749$$

Varianza:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1}$$

$$s^2 = \frac{(0.725-0.749)^2 + (0.756-0.749)^2 + (0.752-0.749)^2 + (0.751-0.749)^2 + (0.76-0.749)^2}{5-1}$$

$$s^2 = 1.897E-4$$

Desviación estándar:

$$s = \sqrt{s^2} = 0.0138$$

Coefficiente de variación (CV) y DER

$$CV = \frac{s}{\bar{X}} = \frac{0.0138}{0.749} = 0.018 \quad DER = CV * 1000 = 18.42$$

REFERENCIAS BIBLIOGRÁFICAS

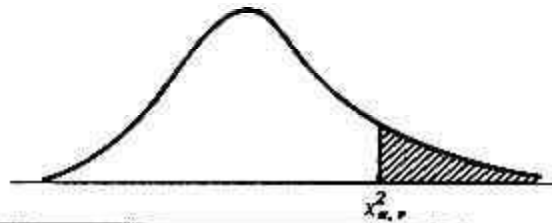
- **Box, G.E.P.; Hunter, W.P. y Hunter, J.S. (1989).** Estadística para investigadores. Barcelona España: Editorial Reverté S.A.
- **García Álvarez, M.C. (1998).** Análisis, homologación y control de calidad. Curso de maestría en química analítica. Universidad de Valencia.
- **Hernández, R.S.; Fernández, C.C. y Baptista, L.P. (1998).** Metodología de la investigación. Segunda edición. México: McGRAW-HILL INTERAMERICANA editores.
- **Juran, J.M. y Gryma, F.M. (1993).** Manual de control de calidad. Tomo II. México: McGRAW-HILL.
- **Mendenhall, W., Beaver, R.J. y Beaver, B.M. (2002).** Introducción a la probabilidad y estadística. México: Thomson editores S.A.
- **Meyer, P.L. (1973).** Probabilidad y aplicaciones estadísticas. Segunda edición. Unites States of america: Fondo Educativo Interamericano.
- **Miller, J. C. y Miller, J. N. (1993).** Estadística para química analítica. Segunda edición. Wilmington, Delaware, E.U.A: Addison-Wesley Iberoamericana, S.A.
- **Ostle, B. (1977).** Estadística aplicada. México: Editora López.
- **Pérez, C. (1999).** Control estadístico de calidad. Mexico: Alfaomega grupo editor S.A. de C.V.
- **Pérez, C. (2002).** Estadística aplicada a través de Excel. México: Alfaomega grupo editor S.A. de C.V.
- **Prat, A., Tort-Martorell, X., Grima, P. y Pozueta, L. (2000).** Métodos estadísticos. Control y mejora de la calidad. México: Alfaomega Grupo Editor.
- **Ramis Ramos, G. (1998).** Quimiometría. Curso de maestría en química analítica. Universidad de Valencia.
- **Romero, R. y Zúnica, L.R. (1993).** Estadística (Proyecto de innovación educativa). España: Servicio de publicaciones de la Universidad Politécnica de Valencia.
- **Skoog, D.A. y West, D.M. (2001).** Química analítica. Séptima edición. México: McGRAW-HILL.

ANEXO A2: TABLA DE LA DISTRIBUCION t-STUDENT

(http://www.usace.army.mil/inet/usace-docs/eng-manuals/em_200_1-7/toc.pdf.)

<i>df</i>	1 • •								
	0.70	0.75	0.80	0.85	0.90	0.95	0.975	0.99	0.995
1	0.727	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657
2	0.617	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.584	0.765	0.978	1.250	1.638	2.353	3.182	4.451	5.841
4	0.569	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.559	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.553	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.549	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.546	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.543	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.542	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.540	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.539	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055
13	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012
14	0.537	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977
15	0.536	0.691	0.866	1.074	1.340	1.753	2.131	2.602	2.947
16	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921
17	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898
18	0.534	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878
19	0.533	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861
20	0.533	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845
21	0.532	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831
22	0.532	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819
23	0.532	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807
24	0.531	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797
25	0.531	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787
26	0.531	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779
27	0.531	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771
28	0.530	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763
29	0.530	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756
30	0.530	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750
40	0.529	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704
60	0.527	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660
120	0.526	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617
•	0.524	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576

ANEXO B2: TABLA DE LA DISTRIBUCION CHI-CUADRADA (χ^2)
(MONTGOMERY, 1991, pág. 432)



ν	α								
	0.995	0.990	0.975	0.950	0.500	0.050	0.025	0.010	0.005
1	0.00 +	0.00 +	0.00 +	0.00 +	0.45	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	1.39	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	2.37	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	3.36	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	4.35	11.07	12.38	15.09	16.75
6	0.68	0.87	1.24	1.64	5.35	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	6.35	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	7.34	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	8.34	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	9.34	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	10.34	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	11.34	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	12.34	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	13.34	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	14.34	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	15.34	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	16.34	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	17.34	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	18.34	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	19.34	31.41	34.17	37.57	40.00
25	10.52	11.52	13.12	14.61	24.34	37.65	40.65	44.31	46.93
30	13.79	14.95	16.79	18.49	29.34	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	39.34	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	49.33	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	59.33	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	69.33	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	79.33	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	89.33	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	99.33	124.34	129.56	135.81	140.17

ANEXO C2: TABLA DE LA DISTRIBUCION F-SNEDECOR
 (http://www.usace.army.mil/inet/usace-docs/eng-manuals/em_200_1-7/toc.pdf.)

Denominator <i>df</i>	1*	Numerator <i>df</i>															
		1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	∞
1	0.90	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2	61.2	61.7	62.3	62.8	63.1	63.3
	0.95	161	200	216	225	230	234	237	239	241	242	246	248	250	252	253	254
	0.975	648	800	864	900	922	937	948	957	963	969	985	993	1001	1010	1014	1018
	0.99	4052	5000	5403	5625	5764	5859	5928	5981	6022	6056	6157	6209	6261	6313	6339	6366
2	0.90	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.42	9.44	9.46	9.47	9.48	9.49
	0.95	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5
	0.975	38.5	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5	39.5
	0.99	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5
3	0.90	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.20	5.18	5.17	5.15	5.14	5.13
	0.95	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.66	8.62	8.57	8.55	8.53
	0.975	17.4	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.5	14.4	14.3	14.2	14.1	14.0	13.9	13.9
	0.99	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	26.9	26.7	26.5	26.3	26.2	26.1
4	0.90	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.87	3.84	3.82	3.79	3.78	3.76
	0.95	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.80	5.75	5.69	5.66	5.63
	0.975	12.2	10.6	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.66	8.56	8.46	8.36	8.31	8.26
	0.99	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.2	14.0	13.8	13.7	13.6	13.5
5	0.90	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.39	3.24	3.21	3.17	3.14	3.12	3.11
	0.95	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.62	4.56	4.50	4.43	4.40	4.37
	0.975	10.0	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.43	6.33	6.23	6.12	6.07	6.02
	0.99	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.72	9.55	9.38	9.20	9.11	9.02
6	0.90	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.87	2.84	2.80	2.76	2.74	2.72
	0.95	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.94	3.87	3.81	3.74	3.70	3.67
	0.975	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.27	5.17	5.07	4.96	4.90	4.85
	0.99	22.8	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.56	7.40	7.23	7.06	6.97	6.88

Denominator <i>df</i>	1*	Numerator <i>df</i>															
		1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	∞
7	0.90	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.63	2.59	2.56	2.51	2.49	2.47
	0.95	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.51	3.44	3.38	3.30	3.27	3.23
	0.975	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.57	4.47	4.36	4.25	4.20	4.14
	0.99	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.31	6.16	5.99	5.82	5.74	5.65
8	0.90	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.46	2.42	2.38	2.34	2.32	2.29
	0.95	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.22	3.15	3.08	3.01	2.97	2.93
	0.975	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.10	4.00	3.89	3.78	3.73	3.67
	0.99	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.52	5.36	5.20	5.03	4.95	4.86
9	0.90	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.34	2.30	2.25	2.21	2.18	2.16
	0.95	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.01	2.94	2.86	2.79	2.75	2.71
	0.975	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.77	3.67	3.56	3.45	3.39	3.33
	0.99	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	4.96	4.81	4.65	4.48	4.40	4.31
10	0.90	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.24	2.20	2.16	2.11	2.08	2.06
	0.95	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.84	2.77	2.70	2.62	2.58	2.54
	0.975	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.52	3.42	3.31	3.20	3.14	3.08
	0.99	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.56	4.41	4.25	4.08	4.00	3.91
12	0.90	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.10	2.06	2.01	1.96	1.93	1.90
	0.95	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.62	2.54	2.47	2.38	2.34	2.30
	0.975	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.18	3.07	2.96	2.85	2.79	2.72
	0.99	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.01	3.86	3.70	3.54	3.45	3.36
15	0.90	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	1.97	1.92	1.87	1.82	1.79	1.76
	0.95	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.40	2.33	2.25	2.16	2.11	2.07
	0.975	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.86	2.76	2.64	2.52	2.46	2.40
	0.99	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.52	3.37	3.21	3.05	2.96	2.87

ANEXO C2: (CONTINUACION) TABLA DE LA DISTRIBUCION F-SNEDECOR

(<http://www.usace.army.mil/inet/usace-docs/eng-manuals/em2001-7/toc.pdf>.)

Denominator <i>df</i>	Numerator <i>df</i>																
	1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	∞	
20	0.90	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.84	1.79	1.74	1.68	1.64	1.61
	0.95	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.20	2.12	2.04	1.95	1.90	1.84
	0.975	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.57	2.46	2.35	2.22	2.16	2.09
	0.99	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.09	2.94	2.78	2.61	2.52	2.42
24	0.90	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.78	1.73	1.67	1.61	1.57	1.53
	0.95	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.11	2.03	1.94	1.84	1.79	1.73
	0.975	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.44	2.33	2.21	2.08	2.01	1.94
	0.99	7.82	6.66	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	2.89	2.74	2.58	2.40	2.31	2.21
30	0.90	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.72	1.67	1.61	1.54	1.50	1.46
	0.95	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.01	1.93	1.84	1.74	1.68	1.62
	0.975	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.31	2.20	2.07	1.94	1.87	1.79
	0.99	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.70	2.55	2.39	2.21	2.11	2.01
60	0.90	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.60	1.54	1.48	1.40	1.35	1.29
	0.95	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.84	1.75	1.65	1.53	1.47	1.39
	0.975	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.06	1.94	1.82	1.67	1.58	1.48
	0.99	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.35	2.20	2.03	1.84	1.73	1.60
120	0.90	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.55	1.48	1.41	1.32	1.26	1.19
	0.95	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.75	1.66	1.55	1.43	1.35	1.25
	0.975	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	1.95	1.82	1.69	1.53	1.43	1.31
	0.99	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.19	2.03	1.86	1.66	1.53	1.38
∞	0.90	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.49	1.42	1.34	1.24	1.17	1.00
	0.95	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.67	1.57	1.46	1.32	1.22	1.00
	0.975	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.83	1.71	1.57	1.39	1.27	1.00
	0.99	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.04	1.88	1.70	1.47	1.32	1.00

CAPITULO III

INFERENCIA ESTADISTICA APLICADA AL ANALISIS QUIMICO

CONTENIDO

	Pág.
3.0 Inferencia estadística	3-2
3.1 División del estudio de la inferencia estadística	3-4
3.2 Estimación	3-5
3.3 Pruebas o contrastes de hipótesis	3-16
3.4 Determinación de datos anómalos	3-41
3.5 Problemas de Inferencia estadística	3-46
REFERENCIAS BIBLIOGRAFICAS	3-68
ANEXO A3. Curvas de potencia para contrastes de hipótesis	3-70
ANEXO B3. Valores críticos para la prueba Q	3-71
ANEXO C3. Valores críticos para la prueba t	3-72
ANEXO D3. Valores críticos para la prueba de Grubbs	3-73

3.0 INFERENCIA ESTADÍSTICA

En general, el propósito de la gran parte de estudios analíticos va más allá de describir las distribuciones de las variables; y se pretende generalizar los resultados obtenidos en la muestra a la población.

Las herramientas estadísticas que se describieron en el capítulo 1, tales como tablas de distribución de frecuencia, gráficos y cálculo de medidas descriptivas, se podrían englobar en el término *estadística descriptiva*, puesto que ellas esencialmente permiten describir, presentar y resumir información que ha sido recolectada de alguna forma. Sin embargo, las técnicas de *estadística descriptiva* no permiten responder interrogantes que pueden surgir cuando no se dispone de la información sobre todos los elementos de la población de interés sino sólo de una parte de ella, es decir, que los datos provienen de una *muestra de elementos* de la población bajo estudio. Por ejemplo, si se conoce que la media o el promedio de 15 mediciones de concentración de ión nitrato, en una muestra concreta de agua, es de 0.5 µg/ml, ¿qué se puede decir sobre el verdadero valor de la media de todas las posibles mediciones de la concentración de nitrato en el río?

La estadística dispone de una gran cantidad de métodos que se engloban dentro de la llamada *estadística inferencial*, los cuales se usan esencialmente para *determinar* la probabilidad de que una conclusión sacada a partir de los datos de una muestra sea admisible en la población muestreada.

El proceso conocido como *inferencia estadística* requiere consideraciones de cómo fue seleccionada la muestra y de cuánto varían las observaciones de una muestra a otra. De manera que los métodos de selección de los elementos que se usarán en las investigaciones analíticas son de considerable importancia para la obtención de resultados y conclusiones válidas (ver capítulo 2.0).

El requisito fundamental de una buena muestra es que sea *representativa* de la población que se trata de describir. Sin embargo, aún cuando se esté seguro que la muestra se obtiene de la apropiada población, otra fuente potencial de error en el muestreo, es el procedimiento en la toma de las mediciones que puede llevar a **resultados sesgados**. El grado de confianza que se puede asignar al valor experimental, método de medición o la

calibración del instrumental utilizado puede dar lugar a resultados que no reflejan la realidad que se quiere analizar¹ (Ver capítulo 6.0).

Por otra parte, en muchas ocasiones no es posible obtener la muestra a partir de todos los elementos que definen la población objetivo, sino sólo a partir de una subpoblación que es accesible al investigador analista en el momento de hacer la selección de los elementos de la muestra y ella recibe el nombre de **población muestreada**.

Los métodos de la inferencia estadística permiten generalizar los resultados de la muestra sólo a los elementos que componen la población muestreada y la generalización hacia la población objetivo está fuera del alcance de la Estadística. Sin embargo, si es posible suponer que la población muestreada es similar a la población objetivo no se cometería un error grande en generalizar los resultados hacia la población objetivo (DAWSON et al., 1994, cap. 5).

En la Figura 3.1, se presentan los alcances de las inferencias realizadas de una muestra.

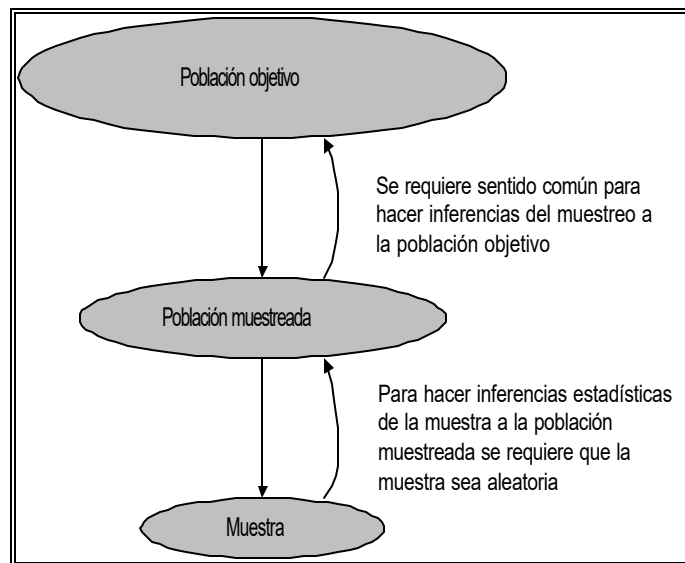


FIGURA 3.1 ALCANCES DE LAS INFERENCIAS REALIZADAS DE UNA MUESTRA (DAWSON et al., 1994, cap. 5)

¹ Aún cuando se esté seguro de que la muestra se obtiene de la población apropiada, es igualmente importante que la muestra se saque de una manera objetiva e insesgada. El problema que surge si el esquema de muestreo no es aleatorio y los datos luego, se analizan como si provienen de una muestra aleatoria, es que se pierde la vital vinculación entre un esquema de muestreo aleatorio y el apropiado método de inferencia estadística, el cual supone siempre que hubo una selección aleatoria de la muestra.

3.1 DIVISION DEL ESTUDIO DE LA INFERENCIA ESTADISTICA

Como se mencionó anteriormente, la *inferencia estadística*, es la metodología consistente en inferir resultados, predicciones y generalizaciones sobre la población estadística, basándose en la información contenida en las muestras representativas previamente elegidas por métodos de muestreo formales (PEREZ, 1999, cap. 5). En la figura 3.2, se presenta un esquema para el proceso general de inferencia.



FIGURA 3.2 PROCESO GENERAL DE INFERENCIA

Existen dos formas generales de efectuar una inferencia acerca de la población que son:

- Estimar valores de parámetros poblacionales, y
- contrastar hipótesis respecto a los valores de parámetros poblacionales o bien del comportamiento de la población.

De ahí que la *inferencia estadística* se divida en dos áreas básicas que son:

1. Problemas de estimación
2. Pruebas o contrastes de hipótesis

La diferencia principal entre ambas áreas radica en que en los problemas de *estimación* se debe elegir un valor de una posible continuidad de alternativas, mientras que en las *pruebas de hipótesis* se debe decidir si aceptar o rechazar un valor especificado o un conjunto de valores especificados de un parámetro.

La *inferencia estadística* está basada en la teoría de la probabilidad, pero tiene un carácter diferente. En *inferencia estadística* se consideraran fenómenos en los que se manifiesta regularidad estadística y se construyen modelos probabilísticos para describirlos.

El problema que se presenta, a la vista de lo expuesto, consiste en determinar el valor de ciertos parámetros poblacionales y/o determinar hipótesis respecto a dichos valores (por ejemplo, la media y la varianza) a partir de la información muestral. Debe advertirse que cuando se trate de poblaciones no caracterizadas por sus parámetros, poblaciones que se denominan *no paramétricas*, los criterios esbozados en orden a la especificación de la población permanecen válidos, aún cuando, en tal caso, los objetivos que satisface el muestreo no sean lógicamente la estimación de parámetros, sino la verificación de otras características poblacionales mediante la denominada *estadística no paramétrica*.

3.2 ESTIMACION

Los procedimientos de estimación pueden dividirse en dos tipos: *estimación puntual* y *estimación por intervalo*. Supóngase que se quiere estimar el contenido medio de nitrito (mg/l), en una muestra de agua de río. La estimación podría darse mediante un solo valor o número, por ejemplo 0.41mg/l, o bien puede reportarse que el contenido de nitrito en el agua, oscila entre 0.405 y 0.415 mg/l. En el ejemplo anterior, a la primera forma de determinar el contenido medio de nitrito en el agua se le conoce como *estimación puntual*, ya que se utiliza un único valor o punto para dar una estimación del parámetro. La segunda forma de dar tal estimación se conoce como *estimación por intervalo*, por que se estiman dos valores o puntos de la recta real que definen un intervalo de valores continuos no numerables.

En esta sección se darán algunas propiedades que debería cumplir un estimador para conseguir estimaciones confiables del parámetro de interés. Se considerarán diferentes formas de estimación y se estudiará una forma de medir la confiabilidad de la estimación.

3.2.1 ESTIMACION PUNTUAL

El procedimiento de *estimación puntual* utiliza la información de una muestra para llegar a un solo número o punto, que estima al parámetro de interés. La estimación real se realiza mediante un *estimador*. Un estimador es una regla que expresa como calcular la estimación, basándose en la información de la muestra y se enuncia en general mediante una fórmula. Los *estadísticos* son en general *estimadores* de los parámetros poblacionales (ver capítulo 2). Por ejemplo, la media muestral (\bar{X}), puede ser un *estimador puntual* de la media poblacional (μ).

Matemáticamente un estimador se define como (MEYER, 1973, cap.14):

“Sea X una variable aleatoria con función de probabilidad $f(x, \theta)$, caracterizada por el parámetro desconocido θ y sea $\underline{x} = \{X_1, X_2, \dots, X_n\}$ una muestra aleatoria, entonces el estadístico $\hat{\theta} = h(X_1, X_2, \dots, X_n)$ correspondiente a θ se denomina estimador puntual de θ ”.

Por ejemplo, suponiendo que x es una variable aleatoria con media $\mu = 4$ y se selecciona una muestra $x = \{2, 5, 11\}$, entonces su media es $\bar{x} = 6$, y su mediana $\tilde{x} = 5$. En este caso, el estimador \tilde{x} produce una estimación más cercana al parámetro μ que el estimador \bar{x} .

En general, se pueden establecer muchos estadísticos para estimar al mismo parámetro y como es de suponer cualquier estimador que se elija evaluará con error al parámetro poblacional. A la diferencia entre un estimador y el correspondiente parámetro de la población se le denomina *error en la estimación del parámetro* θ y se expresa como:

$$\xi = |\hat{\theta} - \theta|$$

Se denomina *sesgo de un estimador* a la diferencia $\theta - E(\hat{\theta})$, donde $E()$ es el operador esperanza matemática (MENDENHALL, 1990, cap. 9).

Otro concepto de interés, es el del *error cuadrático medio (ECM) de un estimador* el cual matemáticamente viene dado como (FREUD y WALPOLE, 1990, cap. 10):

$$ECM(\hat{\theta}) = E((\hat{\theta} - \theta)^2)$$

que mediante desarrollo algebraico y basándose en las propiedades del operador esperanza matemática ($E()$), puede expresarse como:

$$ECM(\bar{\Theta}) = \sigma^2(\Theta) + [\Theta - E(\bar{\Theta})]^2 = \sigma^2(\Theta) + (\text{sesgo})^2$$

3.2.1.1 PROPI EDADES DE LOS ESTI MADORES

El análisis del concepto de *error en la estimación del parámetro* Θ puede conducir a pensar que lo que debe buscarse al elegir un estimador es que sus valores no estén demasiado alejados del valor exacto del parámetro poblacional que se quiere estimar. Sin embargo, es claro que al obtener una estimación puntual su valor dependerá de la muestra que se haya seleccionado y que el valor encontrado puede cambiar de muestra en muestra. De esta forma, las propiedades deseables no serían solamente que cada estimación no se encuentre muy alejada del verdadero valor del parámetro, sino también que no haya demasiada variabilidad entre los valores del estadístico, obtenidos de muestra en muestra. Esto se puede formalizar definiendo las siguientes propiedades a tener en cuenta de los estimadores y que son deseables para una buena estimación puntual, permitiendo a la vez diferenciar y elegir a un estimador de otro (FREUD y WALPOLE, 1990, cap. 10):

1. **Insesgado:** Informalmente esto significa ausencia de error sistemático. De una manera más formal, un estimador cualquiera $\bar{\Theta}$ se dice insesgado si la media o el valor esperado de su correspondiente distribución muestral es igual al parámetro poblacional de interés Θ , es decir el sesgo es cero. De la definición de sesgo, en el ECM se tiene:

$$\begin{aligned} [\Theta - E(\bar{\Theta})]^2 &= 0 \\ \Theta - E(\bar{\Theta}) &= 0, \text{ entonces } E(\bar{\Theta}) = \Theta \end{aligned}$$

2. **Eficiente:** Significa que las estimaciones obtenidas para distintas muestras varían poco entre ellas. De una manera más formal, el estimador más eficiente dentro de

un conjunto de estimadores insesgados será aquel que tiene la menor varianza². Si la variabilidad de las estimaciones se mide a través del desvío estándar, este desvío estándar recibe el nombre de **error estándar del estimador (SE)**. Si $\hat{\Theta}$ es un estimador insesgado de Θ y

$$Var(\hat{\Theta}) = \frac{1}{n \cdot E \left[\left(\frac{\partial \ln f(x)}{\partial \Theta} \right)^2 \right]}$$

entonces $\hat{\Theta}$ es un *Estimador Insesgado de Varianza Mínima (EIVM)*.

3. **Consistente:** Informalmente, un estimador $\hat{\Theta}$ se dice consistente si su variabilidad disminuye a medida que aumenta el tamaño de la muestra (n), es decir que el estimador tiende a acercarse al valor del parámetro. Este hecho se expresa simbólicamente como:

$$\text{Si } n \rightarrow \infty, \text{ entonces } \sigma_{\hat{\Theta}} \rightarrow 0 \text{ o sea que } \hat{\Theta} \rightarrow \Theta$$

Lo que también indica que:

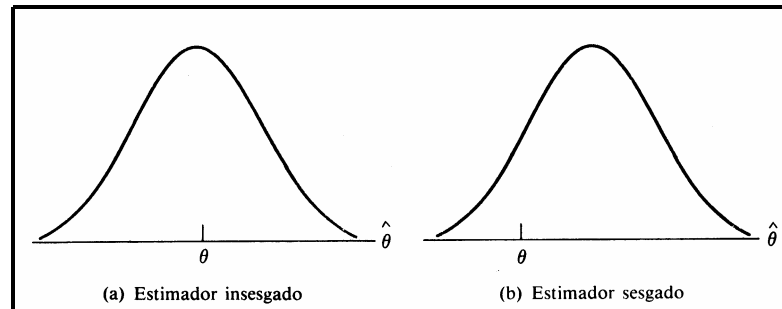
$$\lim_{n \rightarrow N} |\hat{\Theta} - \Theta| \rightarrow 0 \text{ donde } N \text{ es el tamaño de la población}$$

Las distribuciones muestrales para un *estimador insesgado y sesgado*, se presentan en la figura 3.3(a) y las distribuciones muestrales para dos estimadores uno con varianza pequeña y uno con varianza mayor se presenta en la figura 3.3(b).

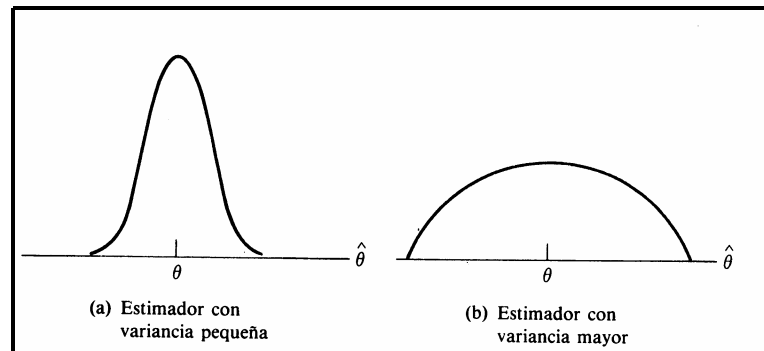
Varios principios de estimación han sido propuestos para la obtención de “buenos” estimadores. Todos ellos conducen a procedimientos matemáticos rutinarios. Tales principios incluyen (OSTLE, 1977, cap. 6):

- 1) El principio de momentos
- 2) Chi-cuadrado mínima
- 3) El método de los mínimos cuadrados
- 4) El principio de la máxima verosimilitud

² No debe confundirse “desvío estándar” de la distribución de la población (variabilidad entre los elementos) con “error estándar” del estimador que es el desvío estándar de la distribución muestral (variabilidad entre las estimaciones de las muestras).



(a)



(b)

FIGURA 3.3 PROPIEDADES DE LOS ESTIMADORES (a) DISTRIBUCION DE ESTIMADORES INSESGADOS Y SESGADOS; (b) COMPARACION DE VARIABILIDAD DE ESTIMADORES (MENDENHALL, 1990, cap. 6)

La aplicación de estos principios a casos particulares conduce a estimadores que pueden diferir y, por lo mismo, tener diferentes atributos de "bondad". Un principio muy de uso que conduce a estimadores con muchos atributos deseables de "bondad", por procedimientos matemáticos rutinarios fácilmente aplicables, es el de la *máxima verosimilitud* establecido por R.A. Fisher (1922) (OSTLE, 1977, cap. 6).

El procedimiento para determinar la estimación de máxima verosimilitud de un parámetro Θ de una población es como sigue (OSTLE, 1977, cap. 6):

- 1) Determine la función de densidad de la muestra $g(X_1, X_2, \dots, X_n; \Theta)$, con *función de probabilidad*:

$$g(X_1, X_2, \dots, X_n; \Theta) = f(X_1; \Theta) f(X_2; \Theta) \dots f(X_n; \Theta) = \prod_{i=1}^n f(X_i; \Theta)$$

- 2) Determine:

$$L = \log \prod_{i=1}^n f(X_i; \Theta)$$

Este paso no es esencial. Sin embargo, ya que las funciones de probabilidad son productos y ya que las sumas son más fáciles de manejar que los productos, se acostumbra aumentar al máximo el logaritmo de la probabilidad, más que la probabilidad misma.

- 3) Determine el valor de Θ que aumente al máximo a L , mediante la solución de la ecuación $\partial L / \partial \Theta = 0$.

En el cuadro 3.1 se presentan los estimadores de máxima verosimilitud para los principales parámetros de distintas distribuciones poblacionales.

TABLA 3.1 ESTIMADORES DE MAXIMA VEROSIMILITUD ASOCIADOS A CIERTAS DISTRIBUCIONES

DI STRI BUCI ON	PARAMETRO (Θ)	ESTI MADOR PUNTUAL ($\hat{\Theta}$)
Binomial	P	$\hat{p} = f / n = \text{frecuencia relativa}$
Poisson	$\lambda (= \mu)$	$\hat{\lambda} = \hat{\mu} = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$
Normal	μ	$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$
	σ	$\hat{\sigma}^2 = s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$
Exponencial	θ	$\hat{\theta} = 1 / \hat{\mu} = 1 / \bar{x} = \frac{n}{\sum_{i=1}^n x_i}$

3.2.2 ESTI MACI ON POR I NTERVALOS

Dado que no puede esperarse que las estimaciones puntuales realmente coincidan con las cantidades que intentan estimar y que estas no garantizan la precisión de dicha estimación

(pues aun el estimador insesgado más eficiente es improbable que estime con exactitud al valor del parámetro poblacional) es preferible reemplazarlas con estimaciones por intervalos. Esto es, con intervalos en los cuales se puede esperar con un grado razonable de certeza que contengan al parámetro en cuestión.

Una estimación por intervalo de un parámetro Θ es un intervalo de la forma $L_1 \leq \Theta \leq L_2$, donde L_1 y L_2 dependen del valor puntual que tome el estimador $\hat{\Theta}$ en una muestra dada y también de la distribución muestral de $\hat{\Theta}$ (FREUD y WALPOLE, 1990, cap. 11).

Como diferentes muestras generalmente producirán valores diferentes de $\hat{\Theta}$, y por consiguiente, diferentes valores de L_1 y L_2 , estos puntos extremos del intervalo son valores de variables aleatorias. Por tanto, con base a la distribución muestral de $\hat{\Theta}$ se puede calcular el intervalo al cual pertenecería la proporción $(1-\alpha)$ de las estimaciones (FREUD y WALPOLE, 1990, cap. 11).

Este intervalo $L_1 \leq \Theta \leq L_2$ determinado en relación con una muestra particular recibe el nombre de intervalo de confianza del $(1-\alpha)100\%$. La proporción $(1-\alpha)$ se conoce como coeficiente o grado de confianza, α es el nivel de significación, y los extremos L_1 y L_2 reciben el nombre de límites de confianza inferior y superior. Por ejemplo, cuando $\alpha = 0.05$ el grado de confianza es de 0.95 y se obtiene un intervalo de confianza del 95% (FREUD y WALPOLE, 1990, cap. 11).

Al interpretarse un intervalo de confianza se tiende a decir: "hay un $(1-\alpha)100\%$ de confianza de que el intervalo contenga al parámetro". La interpretación correcta, si el nivel de confianza es del $(1-\alpha)100\%$, es que si se toman cien muestras y se calcula para cada una de ellas el intervalo de confianza, se espera que el $(1-\alpha)100\%$ de esos intervalos contengan al parámetro Θ . O se puede decir que la probabilidad de que un intervalo de confianza contenga al valor exacto del parámetro Θ es de $(1-\alpha)/100$ ($(1-\alpha)$ de 100 intervalos posibles).

En el grado de confianza $(1-\alpha)100\%$, α se considera a menudo como la probabilidad de cometer un error, ya que indica la proporción de intervalos creados que no contienen al parámetro Θ . De esta forma, $(1-\alpha)$ será la proporción intervalos de confianza creados que contienen al parámetro. Por ejemplo, si el valor $(1-\alpha)100\%$ es de 95%, se llegaría a establecer que en promedio sólo en el 5% de los casos se cometería error al suponer que el intervalo contiene al valor exacto del parámetro.

Para encontrar estos intervalos se debe conocer la distribución muestral de cada estimador. Como ya se estableció en la sección 3.2.1.1, esto depende del parámetro de interés y del estadístico que se elija para estimar dicho parámetro. Sin embargo, es posible dar la forma general que adopta un intervalo de confianza en cualquier caso.

En general, si se quiere estimar un parámetro Θ de la población a través del estadístico $\hat{\Theta}$, y si el error estándar de la distribución de $\hat{\Theta}$ es $\sigma_{\hat{\Theta}}$, que se denotará como $SE(\hat{\Theta})$, entonces un *intervalo de confianza para Θ con un nivel de confianza del $(1-\alpha)100\%$ ($IC_{(1-\alpha)100\%}$)* viene dado por la expresión:

$$IC_{(1-\alpha)100\%} = [\hat{\Theta} - k_1 SE(\hat{\Theta}), \hat{\Theta} + k_2 SE(\hat{\Theta})]$$

Donde k_1 y k_2 dependen de la forma de la distribución muestral de $\hat{\Theta}$.

En síntesis, en el proceso de construcción de intervalos de confianza ($IC_{(1-\alpha)100\%}$) se pueden distinguir dos etapas (PEREZ, 1999, cap.5):

1. **Etapa teórica:** En esta etapa se establece formalmente el intervalo aleatorio $[L_1, L_2]$ de tal forma que $P[L_1 \leq \Theta \leq L_2] = (1-\alpha)$. Se puede hablar de la probabilidad de que el intervalo contenga el valor del parámetro Θ .
2. **Etapa práctica:** Se toma una muestra de la población a la que pertenece el parámetro, y sustituida en L_1 y L_2 se obtienen dos números. Como no se sabe si el parámetro estará o no contenido en el intervalo, se tiene un nivel de confianza del $(1-\alpha)100\%$ de que lo esté. También se dice que el nivel de significación es α . Es deseable lograr que la longitud del intervalo de confianza sea lo más corta posible. Otra propiedad deseable es que la longitud esperada, $E[L_2 - L_1]$, sea lo más pequeña posible.

En las tablas 3.2 a 3.4 se presentan los intervalos de confianza para la estimación de diversos parámetros en poblaciones normales y no normales.

TABLA 3.2 CONSTRUCCION DE INTERVALOS DE CONFIANZA PARA LOS PARAMETROS DE POBLACIONES NORMALES

	Parámetro	σ conocido	σ desconocido
$n \geq 30$ distribuciones normales o no normales	μ	$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ <p>\bar{X}: promedio muestral $z_{\alpha/2}$: valor de la distribución $N(0,1)$, para el nivel de significación α σ: desviación típica de la población n: tamaño de la muestra</p>	$\bar{x} \pm t_{(\alpha/2, n-1)} \frac{S}{\sqrt{n}}$ <p>\bar{X}: promedio muestral $t_{(\alpha/2, n-1)}$: valor de la distribución t para el nivel significación α. S: desviación típica de la muestra n: tamaño de la muestra n-1: grados de libertad</p>
$n < 30$ distribuciones normales	μ	$\bar{x} \pm t_{(\alpha/2, n-1)} \frac{S}{\sqrt{n}}$ <p>\bar{X}: promedio muestral $t_{(\alpha/2, n-1)}$: valor de la distribución t para el nivel significación α. S: desviación típica de la muestra n: tamaño de la muestra n-1: grados de libertad</p>	$\bar{x} \pm t_{(\alpha/2, n-1)} \frac{S}{\sqrt{n}}$ <p>\bar{X}: promedio muestral $t_{(\alpha/2, n-1)}$: valor de la distribución t para el nivel significación α. S: desviación típica de la muestra n: tamaño de la muestra n-1: grados de libertad</p>
	σ^2		$\left(\frac{(n-1)S^2}{\chi_{\alpha/2, n-1}^2}, \frac{(n-1)S^2}{\chi_{1-\alpha/2, n-1}^2} \right)$ <p>$\chi_{\alpha/2, n-1}^2$: valor de distribución chi-cuadrado</p>
	σ		$\left(\sqrt{\frac{(n-1)S^2}{\chi_{\alpha/2, n-1}^2}}, \sqrt{\frac{(n-1)S^2}{\chi_{1-\alpha/2, n-1}^2}} \right)$

TABLA 3.3 CONSTRUCCION DE INTERVALOS DE CONFIANZA PARA COMPARACION DE POBLACIONES NORMALES (PEREZ, 1999, cap.5)

Parámetro	Estadístico	Intervalo de confianza
$\mu_1 - \mu_2$ $(\sigma_1, \sigma_2$ <i>dadas</i>)	$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \rightarrow N(0,1)$	$\bar{x}_1 - \bar{x}_2 \pm Z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
$\mu_1 - \mu_2$ $(\sigma_1 \neq \sigma_2$ <i>no dadas</i>)	$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \rightarrow t_f$ $f = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1+1} + \frac{(S_2^2/n_2)^2}{n_2+1}} - 2$	$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2, f} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ $\bar{x}_1 - \bar{x}_2 \pm Z_{\alpha/2} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \quad (n_1, n_2 > 30)$
$\mu_1 - \mu_2$ $(\sigma_1 = \sigma_2$ <i>no dadas</i>)	$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightarrow t_{n_1+n_2-2}$ $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$	$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2, n_1+n_2-2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
σ_1^2 / σ_2^2	$\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \rightarrow F_{n_1-1, n_2-1}$	$\left(\frac{S_1^2 / S_2^2}{F_{\alpha/2, n_1-1, n_2-1}}, \frac{S_1^2 / S_2^2}{F_{1-\alpha/2, n_1-1, n_2-1}} \right)$
σ_1 / σ_2	$\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \rightarrow F_{n_1-1, n_2-1}$	$\left(\sqrt{\frac{S_1^2 / S_2^2}{F_{\alpha/2, n_1-1, n_2-1}}}, \sqrt{\frac{S_1^2 / S_2^2}{F_{1-\alpha/2, n_1-1, n_2-1}}} \right)$

TABLA 3.4 CONSTRUCCION DE INTERVALOS DE CONFIANZA PARA PARAMETROS DE POBLACIONES NO NORMALES

DISTRIBUCION	PARAMETRO	INTERVALO
BINOMIAL	p	$p \pm z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}} \quad \text{ó} \quad p \pm z_{\alpha/2} \sigma_p$ <p>para poblaciones finitas $\sigma_p = \sqrt{\frac{p(1-p)}{n} \cdot \frac{N-n}{N-1}}$</p> <p>$p$: proporción de éxitos n: tamaño de la muestra N: tamaño de la población $z_{\alpha/2}$: coeficiente de la distribución σ_p: desviación típica de la muestra</p>
DIFERENCIA DE PARAMETROS BINOMIALES	$p_1 - p_2$	$(p_1 - p_2) \pm z_{\alpha/2} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$ <p>subíndices 1 y 2: indican muestras 1 y 2</p>
POISSON	λ	$\bar{X} \pm \lambda_{\alpha} \sqrt{\frac{\bar{X}}{n}}$ <p>siendo $\lambda_{\alpha} = F^{-1}(1-\alpha/2)$ que es la función de distribución de una $N(0,1)$</p> <p>\bar{X}: promedio muestral</p>

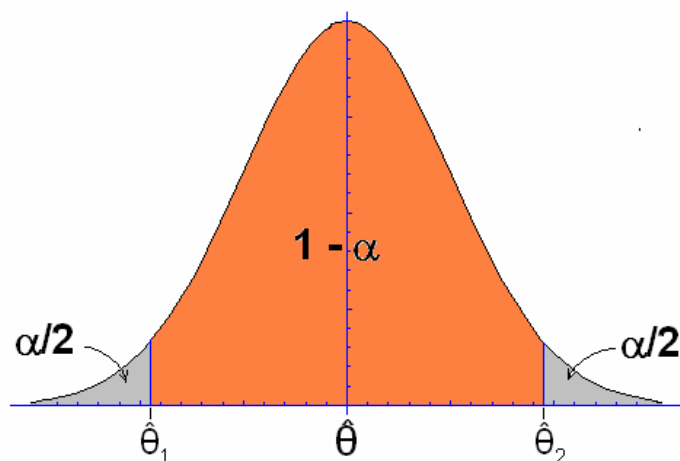


FIGURA 3.4 REPRESENTACION GRAFICA DE UN INTERVALO DE CONFIANZA

3.3 PRUEBAS O CONTRASTES DE HIPÓTESIS

El problema central de la *inferencia estadística* es un problema de toma de decisiones, del cual la *estimación y contraste de hipótesis* son aspectos importantes, diferenciados pero complementarios.

Existen múltiples problemas en los cuales, en lugar de tener la necesidad de hacer una estimación de un parámetro, se debe decidir si una afirmación relativa a un parámetro es verdadera o falsa. Esto es, se debe probar una hipótesis relativa a un parámetro. Por ejemplo, para un análisis EDTA se quiere determinar si la concentración rotulada de una solución estándar de tiosulfato continúa siendo la misma o bien ha cambiado a tal grado que el uso de esa solución para un análisis ya no sea confiable y deba reemplazarse.

Un *contraste de hipótesis*, una *prueba de significación* o una *prueba de hipótesis*, son términos que indican el proceso mediante el cual se decide si una proposición acerca de la población debe ser aceptada o no. Esta proposición es lo que se conoce como *hipótesis estadística*. Una *hipótesis estadística* es por tanto, una proposición acerca de la distribución de probabilidad de una o de varias variables aleatorias. Tal proposición puede referirse bien a la forma de la distribución, bien al valor o valores de los parámetros o bien a ambos.

El contraste de la hipótesis estadística se basará en la información proporcionada por la muestra. De modo que si se rechaza la hipótesis, se quiere indicar que los datos de la muestra ofrecen cierta evidencia sobre su falsedad. Si se acepta simplemente se quiere significar que no se rechaza.

Todo contraste va a implicar la comparación entre dos hipótesis a la luz de la información muestral. Como resultado el investigador tomará una *decisión*.

Por ejemplo, se desea decidir si la utilización de un nuevo método de tratamiento de muestras para la determinación de calcio mejora o permite la obtención de datos tan confiables como los generados al tratar las muestras con el método oficial. La investigación anterior puede ir orientada a tratar de comparar la media muestral del contenido de calcio en un estándar o bien comparar los promedios muestrales entre un grupo de datos en los que se ha utilizado el método oficial de tratamiento de muestras, y un grupo de datos en los que se ha empleado el nuevo método de tratamiento de muestras. El contraste en el primer caso consistirá en comparar la hipótesis: "Existe diferencia entre el promedio muestral y el verdadero", con la hipótesis de que "no existe

diferencia", y en el segundo caso contrastar las hipótesis sobre la igualdad o diferencia entre los promedios de calcio utilizando ambos métodos a partir de las observaciones de las dos muestras. En general, los resultados favorecerán más a una de las dos hipótesis que a la otra, y por tanto se podrá tomar una decisión.

3.3.1 ELEMENTOS DE UNA PRUEBA DE HIPÓTESIS

A continuación se describen los elementos más importantes para cualquier contraste de hipótesis (MENDENHALL, 1990, cap.7):

- a. HIPOTESIS NULA: Es aquella hipótesis que se desea contrastar, se simboliza por H_0 . Esta suele ser una estrategia o medio del que se sirve el investigador para probar la alternativa. El planteamiento de H_0 permite elaborar un modelo probabilístico a partir del cual se puede llegar a una decisión final.
- b. HIPOTESIS ALTERNATIVA: También se conoce como experimental y se representa por H_1 . Esta es la hipótesis de investigación. De modo que se espera que hay un argumento para la hipótesis de investigación (o alternativa) H_1 , demostrando que no lo hay para su contraria, la hipótesis nula.

En general, las hipótesis pueden ser *simples y compuestas*. Una hipótesis es simple si se especifica exactamente el valor del parámetro, ejemplos: $\mu = 10$, $\sigma^2 = 4$ ó $\rho_{xy} = 1$, etc. Una hipótesis es compuesta si contiene dos o más valores para el parámetro, como por ejemplo: $\mu \neq 10$, $\sigma^2 > 4$, etc.

La hipótesis nula, por ser más concreta, suele ser simple, y la alternativa compuesta. No obstante, es frecuente plantearlas como complementarias. Suele ser habitual hacer supuestos que dependen del tipo de contraste o prueba y que se pueden resumir como:

- Características de los datos que se van a manipular: independencia de las observaciones, nivel de medida utilizado, etc.
- Supuestos acerca de la forma de la distribución de partida: normal, binomial, etc.

Algunos de estos supuestos podrán contrastarse, otros serán de difícil verificación. La violación de los supuestos podrá invalidar más o menos el modelo probabilístico y conducir a conclusiones erróneas.

- c. ESTADÍSTICO DE PRUEBA: Los valores muestrales se emplean para calcular un solo número que corresponde a un punto sobre una recta. Tal número sirve para tomar decisiones y se denomina *estadístico de prueba o contraste* y se simboliza en general como T . Tal estadístico, será una variable aleatoria, con una distribución muestral, conocida si la hipótesis hecha es verdadera, que proporcionará las probabilidades asociadas a un valor determinado o a un intervalo de valores del *estadístico de contraste*.
- d. REGLAS DE DECISION: Una regla de decisión es un criterio utilizado para decidir si se acepta o se rechaza la hipótesis nula, a partir del espacio muestral de valores del estadístico de contraste y de las probabilidades asociadas. Este criterio consiste en dividir tal espacio en dos zonas mutuamente excluyentes y exhaustivas: *la zona de rechazo* (o región crítica) y *la zona de aceptación*. La zona de rechazo está constituida por aquellos valores del estadístico de contraste que por ser muy grandes o muy pequeños (o lo que es igual alejarse mucho de H_0), es muy poco probable que ocurran si H_0 es verdadera.

Las reglas de decisión se pueden expresar en términos de probabilidad como: "rechazar H_0 si la probabilidad de que ocurra un valor concreto T (T_k) del estadístico de prueba es menor que α ". Por ejemplo, para el problema de comparar los dos métodos de tratamiento de muestras para la cuantificación de la concentración de calcio, la regla de decisión puede ser: "rechazar el hecho de que ambos métodos de tratamiento de muestras producen resultados equivalentes si T_k toma un valor tan grande (o tan pequeño) que la probabilidad de su ocurrencia sea menor o igual que $\alpha = 0.05$ ". El valor de α se conoce como *nivel de significación o nivel de riesgo* y representa la probabilidad de que un valor concreto del estadístico de contraste caiga en la zona de rechazo o crítica.

Si se establece previamente la zona de rechazo o crítica, se determina un valor único de α . Sin embargo, en la práctica, se suele fijar de antemano un valor de α ,

lo que lleva consigo distintas zonas críticas (en ocasiones infinitas) de las que habrá seleccionar la mejor. En la práctica, los niveles de significación más frecuentemente empleados son de 0.05 y 0.01, que en algunos casos aparecen como 5 por 100 y el 1 por 100.

3.3.2 ERRORES EN LOS CONTRASTES DE HI POTESI S

En general, una prueba de hipótesis estadística no es otra cosa que un procedimiento para tomar una decisión bajo incertidumbre, sobre la validez de la hipótesis nula. Usando la evidencia de los datos de la muestra o de las muestras se llega a dos decisiones:

Decisión 1: rechazar H_0 ; **Decisión 2:** aceptar H_0

La realidad es también doble: H_0 es falsa o H_0 es verdadera. Por tanto, si H_0 es verdadera y se rechaza se comete un error. Si H_0 es falsa y se acepta se comete otro error. En conclusión, existe un riesgo inherente de llegar a una conclusión incorrecta cuando se prueban hipótesis, a causa de la naturaleza aleatoria del muestreo. Este riesgo se puede clasificar de dos maneras:

- *ERROR TIPO I:* Es el que se comete al rechazar H_0 cuando es verdadera. La probabilidad de cometer este error es α .
- *ERROR TIPO II:* Es el que se comete al aceptar H_0 cuando es falsa. La probabilidad de cometer este error es β .

Lo antes expuesto se puede resumir esquemáticamente como sigue:

DECISION	SITUACIÓN DE H_0	
	Verdadera	Falsa
Aceptar H_0	<i>Decisión correcta</i> ($1-\alpha$)	<i>Error tipo II</i> (β)
Rechazar H_0	<i>Error tipo I</i> (α)	<i>Decisión correcta</i> ($1-\beta$)

En general, se puede decir que un buen contraste o una buena regla de decisión tenderá a minimizar los errores inherentes a toda decisión. Pero esto no es fácil, pues para una muestra de tamaño n , al intentar disminuir uno de los errores aumenta el otro. El investigador analista debe saber en cada caso qué error tiene mayor importancia, para controlarlo y tratar de disminuirlo.

Se llama *potencia de un contraste*, a la probabilidad de rechazar la hipótesis nula cuando es falsa (decisión correcta), su probabilidad es $1-\beta$. Más estrictamente debería llamarse *potencia de una región crítica*, y no es más que la probabilidad de que esta detecte una H_0 falsa, dado un valor para H_1 (MONTGOMERY, 1991, cap.2). En resumen, las probabilidades de importancia en los contrastes de hipótesis son (PEREZ, 1999, cap. 5):

- $\alpha = P\{\text{error tipo I}\} = P\{\text{rechazar } H_0/H_0 \text{ es verdadera}\}$
- $\beta = P\{\text{error tipo II}\} = P\{\text{no rechazar } H_0/H_0 \text{ es falsa}\}$
- $1-\beta = \text{Poder de la prueba} = P\{\text{rechazar } H_0/H_0 \text{ es falsa}\}$

En una situación práctica, en la mayoría de las pruebas estadísticas el interés generalmente es "rechazar H_0 ", por lo que la mayor parte de las veces la probabilidad de error que se controla usualmente durante este procedimiento es, justamente, el error asociado a esta decisión (probabilidad de error tipo I) es decir la probabilidad de rechazar H_0 cuando es cierta.

Puede ser muy difícil calcular los valores de β para los diferentes parámetros de prueba ya que dependerá de la hipótesis alternativa que se escoja³. Es más fácil detectar diferencias grandes con un valor de H_0 de un parámetro, que detectar diferencias pequeñas. Si se quieren detectar diferencias pequeñas con el valor de H_0 de un parámetro, β será grande. Si se desea detectar diferencias grandes, β será pequeño. Para un tamaño muestral fijo (n) al aumentar la región de rechazo (y por lo tanto α), β disminuye. Si por el contrario α decrece, β aumentará. Por otra parte, al aumentar el tamaño de la muestra (n), α y β decrecen a la vez (MENDENHALL, 1990, cap. 7).

³ Que pueden ser de tipo compuestas

3.3.3 PROCEDIMIENTOS DE UN CONTRASTE O PRUEBA DE HIPÓTESIS

En resumen, el procedimiento general para realizar un contraste de hipótesis es especificar el valor de la probabilidad de cometer un error tipo I (α), y después diseñar un procedimiento de prueba para obtener un pequeño valor de la probabilidad de cometer un error tipo II (β) ($\leq 10\%$). Por regla general, se elige directamente el riesgo α y se intenta minimizar el riesgo β . El riesgo β como se vio en la sección 3.3.2, es una función del tamaño muestral, y se controla indirectamente. Cuanto más grande es el tamaño muestral para la prueba, tanto menor será el riesgo β .

Existen tres procedimientos generales para el contraste de hipótesis que son:

- a. Procedimiento general de un contraste de hipótesis basado en la región de rechazo.
- b. Procedimiento general de un contraste de hipótesis basado en el p-valor.
- c. Procedimiento general de un contraste de hipótesis basado en la creación de intervalos de confianza.

En las secciones de la 3.3.3.1 a la 3.3.3.3 se detallan estos procedimientos.

3.3.3.1 PROCEDIMIENTO GENERAL DE UN CONTRASTE DE HIPÓTESIS BASADO EN LA REGIÓN DE RECHAZO

A continuación se presenta una secuencia de pasos que puede usarse de forma sistemática para efectuar cualquier contraste de hipótesis (PEREZ, 1999, cap.5):

- 1) Formular, de forma clara y explícita, cual es la hipótesis nula (H_0), y cuál es la hipótesis alternativa (H_1) en términos de parámetros.
- 2) Determinar el tamaño de la muestra (n), y fijar el nivel de significación deseado α , que es el máximo error aceptable cuando se rechaza H_0 siendo cierta, y usar este valor para construir la *región crítica*.
- 3) Usar la teoría estadística para construir un indicador de concordancia entre los datos y la hipótesis nula. Este indicador denominado *estadístico de prueba* será usado para juzgar la hipótesis H_0 .

- 4) Establecer el criterio o regla de decisión, es decir la zona de la distribución del estadístico de prueba donde se rechaza H_0 , la cual se denomina *región crítica*.
- 5) Calcular el valor del estadístico de prueba a partir de la muestra.
- 6) Aplicar la regla de decisión: si el valor del estadístico de prueba pertenece a la región crítica, entonces rechazar H_0 . En caso contrario, lo que se puede afirmar es que no existe suficiente evidencia para rechazar H_0 , y se acepta H_0 .
- 7) Si se dispone de una hipótesis alternativa y de la distribución del estadístico de contraste bajo la suposición de que es válida la hipótesis alternativa, es posible calcular la probabilidad de error tipo II.

3.3.3.2 PROCEDIMIENTO GENERAL DE UN CONTRASTE DE HIPOTESIS BASADO EN EL P-VALOR

Cuando se interpretan los resultados de un contraste, las conclusiones están basadas en una regla de decisión. Esta, como se ha visto anteriormente, se establece teniendo en cuenta el riesgo que asume el investigador analista de cometer un error tipo I, siendo la probabilidad de este error el nivel de significación α . Otras personas que quieran utilizar los resultados del experimento pueden tener en mente, sin embargo, la decisión a tomar para un nivel de significación diferente, con lo cual será útil conocer qué tipo de decisión se puede adoptar según el nivel de significación real de un contraste basándose en los datos observados. Este concepto actuará como contrapuesto al nivel de significación elegido antes de realizar el contraste (PEREZ, 1999, cap.5).

Otro procedimiento general para efectuar un contraste de hipótesis, más usado en la actualidad debido a la potencialidad de los ordenadores y del software estadístico existente, consiste en tomar una decisión a partir del cálculo de la probabilidad del estadístico muestral que se utiliza para contrastar la hipótesis nula, y que es proporcionada por las salidas de tales programas. Esta probabilidad se denomina p -valor o simplemente p . En otras palabras se denomina p -valor o nivel de significación observado o valor p , a la probabilidad del estadístico muestral en su distribución y se compara con el nivel de significación α .

El procedimiento de contraste de hipótesis usando el concepto del p -valor se puede resumir en los siguientes pasos (PEREZ, 1999, cap.5):

- 1) Suponer que H_0 es cierta.
- 2) Para confrontar esta suposición con la información (parcial) que proveen los datos sobre la realidad de H_1 , se forma una especie de "indicador" de concordancia, denominado *estadístico de la prueba*, el cual es función de los datos.
- 3) Como el estadístico depende de la información de los datos, con cada muestra posible hay asociado un valor de este estadístico y en consecuencia se genera una nueva variable aleatoria. Asociada a esta variable hay una cierta distribución de probabilidad, a partir de la cual se determina la probabilidad de que la información de los datos concuerde con la hipótesis nula, denominado *p*-valor.
- 4) Es claro que si de antemano se fija la máxima probabilidad de error al rechazar H_0 cuando es cierta, esta debe ser igual a α , otra forma de tomar la decisión es comparar el valor del *p*-valor con α . Así:
 - Si $p \leq \alpha$ entonces la decisión es **rechazar H_0** .
 - Si $p > \alpha$ la decisión es que **no existe evidencia suficiente para rechazar H_0** .

3.3.3.3 PROCEDIMIENTO GENERAL DE UN CONTRASTE DE HIPÓTESIS BASADO EN LA CREACION DE INTERVALOS DE CONFIANZA

Los intervalos de confianza tienen la ventaja frente a los contrastes de hipótesis de que siempre dan una idea de la zona en que se va a encontrar el verdadero valor del parámetro poblacional, mientras que en el caso de los contrastes de hipótesis, cuando se rechaza una hipótesis nula, no se conoce el valor del parámetro de la población en cuestión. Todo lo que se sabe es que es más verosímil que el valor del parámetro sea mayor o menor que un valor concreto (PEREZ, 1999, cap. 5).

En algunas ocasiones, en lugar de llevar a cabo un procedimiento de contraste de hipótesis sobre un parámetro, se puede construir un intervalo de confianza para el parámetro y extraer conclusiones del mismo. Supóngase que se desea contrastar $H_0: \mu = \mu_0$ frente a $H_1: \mu \neq \mu_0$. Si el nivel de significación es α , se pueden utilizar los datos de la muestra para construir un intervalo de confianza para μ al $(1-\alpha)100\%$. Si el valor de μ_0 no está incluido en el intervalo hay que rechazar H_0 ; pero si el valor de μ_0 se encuentra en el intervalo calculado, no hay evidencia para rechazar a H_0 . Una ventaja de este

procedimiento es que se pueden contrastar muchas hipótesis nulas de una vez, ya que se puede rechazar cualquier hipótesis nula de la forma $H_0: \mu = k$ para cualquier valor de k que no pertenezca al intervalo (PEREZ, 1999, cap.5).

3.3.4 USO DE CURVAS DE POTENCIA O CURVAS OC

Un aspecto de mucha importancia en los problemas de contrastes de hipótesis es la selección del tamaño muestral. Dicha selección y la probabilidad del error tipo II (β) están muy relacionados. Supóngase que se contrasta la hipótesis: $H_0: \mu = \mu_0$ frente a $H_1: \mu \neq \mu_0$, y que la media no es igual al valor especificado, de forma que $\Delta = \mu - \mu_0$. Como $H_0: \mu = \mu_0$ no es verdadera el problema está en decidir erróneamente no rechazar H_0 . La probabilidad de error tipo II depende de la diferencia verdadera que hay entre la media y el valor especificado (Δ). La gráfica de β contra Δ , en el caso particular de un tamaño muestral n de la muestra, se conoce como **Curva Característica de Operación** o **Curva OC** de la prueba. Generalmente, para un valor dado de Δ , el valor de β disminuye a medida que aumenta el tamaño de la muestra n . Esto significa que la diferencia específica del nivel medio es más fácil de detectar en muestras grandes que en muestras pequeñas. Todo contraste de hipótesis tiene una curva característica de operación (OC) (JURAN y GRAYMA, 1993, cap. 23).

En muchas situaciones prácticas, las probabilidades de error tipo II (β) pueden determinarse a partir de diagramas semejantes a los dados en el anexo A3. El diagrama *a*, del anexo A3, permite localizar los valores de β que corresponden a los de μ , μ_0 , σ , n y α para la hipótesis propuesta, y se fundamenta en la suposición de que la distribución muestral de la media es una normal; en consecuencia puede utilizarse cuando n es relativamente grande o si la población que se está muestreando tiene forma aproximada a la distribución normal. Para el caso, la figura *a*, se aplica a pruebas de dos colas con $\alpha = 0.05$ y la cantidad que se grafica no es μ sino (MILLER y FREUND, 1986, cap. 7):

$$d = \frac{|\Delta|}{\sigma} = \frac{|\mu - \mu_0|}{\sigma}$$

En la figura *b*, del anexo A3, se presenta la curva de potencia para una diferencia de medias de dos poblaciones en la que la cantidad que se grafica es (MONTGOMERY, 1991, cap.2):

$$d = \frac{|\Delta|}{2\sigma} = \frac{|\mu_1 - \mu_2|}{2\sigma}$$

3.3.4.1 DETERMINACION DEL TAMAÑO MUESTRAL PARA CONTRASTAR HIPOTESIS.

En los apartados anteriores, se daba por supuesto que el tamaño muestral se fijaba por razones estadísticas y que el error tipo I solo quedaba predefinido en la prueba. El procedimiento ideal es definir los errores tipo I y II deseados y calcular el tamaño muestral n requerido para cubrir ambos tipos de errores. El tamaño requerido dependerá de (JURAN y GRYMA, 1993, cap. 23):

- i) Los riesgos de muestreo deseados α y β .
- ii) El tamaño de la menor diferencia verdadera que ha de ser detectada.
- iii) La variación de la característica que ha de ser medida.

El tamaño muestral n puede determinarse haciendo uso de la curva OC correspondiente a la prueba.

Supóngase que el tiempo medio de secado de un precipitado es aproximadamente de 35 minutos, específicamente se tiene un 90% de seguridad de detectar ese cambio ($\beta = 0.1$). Además si la media verdadera fuera de 30 minutos y $\sigma = 10$, debía haber sólo un 5% de riesgo de rechazar H_0 ($\alpha = 0.05$). Del gráfico a) del anexo A3 se define "d" como:

$$d = \frac{|\mu - \mu_0|}{\sigma} = \frac{|35 - 30|}{10} = 0.5$$

Localizando el punto correspondiente a $d = 0.5$ y $\beta = 0.10$ en el diagrama *a* del anexo A3, se encuentra que la muestra n ha de ser igual a 30, aproximadamente.

En la práctica del análisis químico, es necesario el análisis de las curvas OC correspondientes a muestras de varios tamaños para ayudar a adoptar una decisión acerca del tamaño requerido de los riesgos I y II. En general, los valores de α y β deberán

mantener un equilibrio, α debe regularse de tal forma que la probabilidad de β no sea superior al 10%.

3.3.4.2 CONTRASTES DE HIPOTESIS UNILATERALES Y BILATERALES

Supóngase, como ilustración que se utiliza el símbolo θ para denotar el parámetro poblacional de interés, por ejemplo, θ puede ser μ , $(\mu_1 - \mu_2)$, π ó $(\pi_1 - \pi_2)$, y el símbolo $\hat{\theta}$ para denotar el estimador puntual insesgado correspondiente. Desde el punto de vista práctico se puede tener interés en contrastar la hipótesis nula $H_0: \theta = \theta_0$, contra la alternativa de que el parámetro poblacional es mayor que θ_0 , o sea $H_1: \theta > \theta_0$. En esta situación, se rechazará H_0 cuando $\hat{\theta}$ sea grande, o sea cuando el estadístico de prueba sea mayor que un cierto valor llamado valor crítico, que separa las regiones de rechazo y no rechazo de la hipótesis. La probabilidad de rechazar la hipótesis nula cuando es cierta será igual al área bajo la curva de la distribución muestral del estadístico de prueba sobre la región de rechazo. Por ejemplo, en el caso de que se esté trabajando con una distribución normal y un $\alpha = 0.05$, se rechaza la hipótesis nula cuando $\hat{\theta}$ se encuentre a más de $1,645\sigma_{\hat{\theta}}$ a la derecha de θ_0 . En general, para detectar $\theta > \theta_0$, se sitúa la región de rechazo en la extremidad de valores superiores a $\hat{\theta}$. Para detectar $\theta < \theta_0$ se ubica la región de rechazo en la extremidad izquierda de la distribución de $\hat{\theta}$, o sea para valores inferiores a $\hat{\theta}$. A este tipo de contraste se le denomina prueba estadística de una cola o unilateral, es decir que un contraste unilateral es aquel tipo de prueba en la que la región de rechazo se localiza solamente en una cola o extremo de la distribución muestral del estadístico de contraste (JURAN y GRYMA, 1993, cap. 23).

Si hay que detectar diferencias mayores o menores de θ_0 , la hipótesis alternativa será $H_1: \theta \neq \theta_0$, es decir: $\theta > \theta_0$ o bien $\theta < \theta_0$. En este caso la probabilidad de error Tipo I (α) se repartirá entre las dos colas de la distribución muestral del estadístico, y se rechazará H_0 para valores de $\hat{\theta}$ mayores que un valor crítico ($\theta_0 + C$) o menor que ($\theta_0 - C$). Esta prueba se llama prueba estadística bilateral o de dos colas (JURAN y GRYMA, 1993, cap. 23).

En los contrastes unilaterales, utilizando la creación de un intervalo de confianza, H_0 se rechaza si θ_0 es mayor que el extremo superior del intervalo de confianza (caso de $\theta > \theta_0$),

mientras que si θ_0 menor o igual que el extremo superior del intervalo de confianza, no hay evidencia para rechazar H_0 . Si se quiere contrastar $H_0: \theta = \theta_0$ frente a $H_1: \theta < \theta_0$, a un nivel de significación α se puede razonar de manera similar, si se construye un intervalo de confianza para θ al $100(1-2\alpha)\%$. Si el valor de θ_0 es menor que el extremo inferior del intervalo, hay que rechazar H_0 ; si θ_0 es mayor o igual que el extremo inferior del intervalo, no hay evidencia para rechazar H_0 . Para el caso de contrastes bilaterales, se rechaza H_0 si θ_0 cae fuera del intervalo creado del $100(1-\alpha)\%$ (JURAN y GRYMA, 1993, cap. 23).

A continuación se detallan algunas directrices para la realización de un ensayo unilateral o bilateral basado en el análisis de las curvas características de operación (JURAN y GRYMA, 1993, cap. 23):

- Realícese un ensayo unilateral (una cola) con todo el riesgo a la derecha si: (1) se sabe que (si θ_0 no es verdadero) el parámetro real (o de comparación) es mayor que θ_0 ; (2) los valores del parámetro poblacional menores que θ_0 (o el valor especificado) son aceptables y solo se está interesado en detectar un parámetro poblacional mayor que θ_0 .
- Realícese un ensayo bilateral (dos colas) si: (1) no hay un conocimiento previo acerca de la posición del parámetro real o, (2) se está interesado en detectar un parámetro poblacional real menor o mayor que el θ_0 enunciado en la hipótesis original.

3.3.5 MODO DE SACAR CONCLUSIONES DE LOS CONTRASTES DE HIPÓTESIS

Con los contrastes de hipótesis se pueden sacar conclusiones útiles. Cuando se rechaza una hipótesis, la conclusión práctica es "el valor del parámetro especificado en la hipótesis es erróneo". Se llega a esta conclusión con profunda convicción, sencillamente hablando, con un nivel de confianza del $100(1-\alpha)\%$. La pregunta clave es entonces: ¿Cuál es, pues, una buena estimación del valor del parámetro correspondiente a la población? Se puede ayudar a responder esta pregunta calculando los "límites de confianza" del parámetro (JURAN y GRYMA, 1993, cap. 23).

Cuando se acepta una hipótesis, el valor numérico del parámetro contrastado en la hipótesis no se ha probado, pero no se ha refutado. No es válido decir que ha quedado probado que la hipótesis es correcta, con un nivel de confianza del $100(1-\alpha)\%$. Se podrían aceptar otras muchas hipótesis para la misma muestra de observaciones y, sin embargo, sólo una hipótesis puede ser verdadera. Por lo tanto, una aceptación *no* significa haber probado, con una elevada probabilidad, que una hipótesis determinada es correcta. (Si permanecen invariables todos los demás factores, cuanto más pequeña sea la muestra, más probable es que se acepte la hipótesis. La menor evidencia no implica ciertamente una prueba.) Por esta razón, la expresión más frecuente utilizada actualmente es la de que "la hipótesis no ha sido rechazada, con un nivel de significación α " en vez de "la hipótesis ha sido aceptada, con un nivel α " (JURAN y GRYMA, 1993, cap. 23).

Cuando se acepta una hipótesis, la pregunta clave es: ¿Qué conclusión puede extraerse, si es que puede extraerse alguna, acerca del valor del parámetro de la hipótesis? Se sugieren dos enfoques (JURAN y GRYMA, 1993, cap. 23):

1. Calcular los límites de confianza del resultado de la muestra. Estos límites de confianza definen un intervalo dentro del cual está el verdadero parámetro de la población. Si este intervalo es pequeño, una decisión de aceptación en el ensayo de hipótesis significa que el verdadero valor de la población es igual o muy próximo al indicado en la hipótesis. Luego, es razonable actuar como si el valor del parámetro especificado fuera de hecho correcto. Si el intervalo de confianza es relativamente amplio, implica una seria advertencia de que el valor enunciado en la hipótesis no ha quedado probado y que el valor verdadero de la población podría ser muy diferente al especificado en la hipótesis.
2. Construir y analizar la curva característica para el contraste de hipótesis. Esta define la probabilidad de que puedan aceptarse con la prueba otros valores posibles del parámetro poblacional. El conocimiento de estas probabilidades para valores relativamente cercanos a la hipótesis original puede ayudar a extraer nuevas conclusiones acerca de la aceptación de la hipótesis original.

Hay que tener siempre cuidado de extraer conclusiones técnicas de conclusiones estadísticas, en particular cuando se acepta la hipótesis (JURAN y GRYMA, 1993, cap. 23).

3.3.6 CASOS PARTICULARES DE CONTRASTES DE HIPOTESIS

En los contrastes de hipótesis es frecuente hacer supuestos que dependen del tipo de contraste o prueba, como son las características acerca de los datos que se van a manipular (independencia de las observaciones, nivel de medida utilizada, etc.) y los supuestos de la distribución de partida (normal, binomial, etc.). Algunos de estos supuestos podrían controlarse, otros serán de difícil verificación. La violación de dichos supuestos podrá invalidar más o menos el modelo probabilístico y conducir a conclusiones erróneas.

En general, un contraste de hipótesis puede servir para:

- Comparar un promedio muestral con un valor especificado.
- Comparar la variabilidad muestral con un valor específico.
- Comparar varianzas de dos muestras.
- Comparar promedios de dos muestras.
- Comparar parámetros de poblaciones binomiales.
- Comparar parámetros de poblaciones de Poisson.

En las secciones de la 3.3.6.1 a la 3.3.6.3, se presentan los resúmenes de los elementos principales para los distintos tipos de contrastes de hipótesis, tanto para el caso unilateral como bilateral.

3.3.6.1 COMPARACION DE PROMEDIOS MUESTRALES

En la comparación de promedios muestrales pueden presentarse dos casos generales:

1. Comparación de un promedio muestral con un valor especificado (μ_0). Que presenta dos casos particulares: (i) Comparación de un promedio muestral con un valor especificado (μ_0) cuando se conoce la desviación estándar de la población (σ), y (ii) comparación de un promedio muestral con un valor especificado (μ_0) cuando no se conoce la desviación estándar de la población (σ).
2. Comparación de dos promedios a partir de las medias muestrales (\bar{X}_1 y \bar{X}_2). Se presentan tres casos particulares: (i) Comparación de dos promedios cuando las

desviaciones estándar de las dos poblaciones son conocidas (σ_1 y σ_2); (ii) comparación de dos promedios cuando las desviaciones estándar de las dos poblaciones son desconocidas y consideradas diferentes ($\sigma_1 \neq \sigma_2$); y (iii) comparación de dos promedios cuando las desviaciones estándar de las dos poblaciones son desconocidas y consideradas iguales ($\sigma_1 = \sigma_2$).

El control de algunos supuestos a tomar en cuenta para llevar a cabo estos tipos de contrastes, puede lograrse cuando los datos se recolectan de forma que el único factor que influya, sea aquel cuyo efecto se quiere evaluar, así como también, mediante la aleatorización de la recogida de datos (todo lo que se pueda) para protegerse de posibles sesgos introducidos por factores no identificados.

En las tablas de la 3.5 a la 3.9 se describen los componentes principales de los contrastes de hipótesis antes mencionados en sus distintos casos particulares.

TABLA 3.5 COMPARACION DE LA MEDIA MUESTRAL CON UN VALOR ESPECIFICADO, CASO σ CONOCIDA

1. Hipótesis nula: $H_0: \mu = \mu_0$	
2. Hipótesis alternativa	
Prueba de una cola	Prueba de dos colas
$H_1: \mu > \mu_0$ (o bien $\mu < \mu_0$)	$H_1: \mu \neq \mu_0$
3. Estadístico de contraste: $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	
4. Región de rechazo:	
Prueba de una cola	Prueba de dos colas
$Z > Z_\alpha$ (o sea, $Z < -Z_\alpha$ cuando la hipótesis alternativa es $H_1: \mu < \mu_0$)	$ Z > Z_{\alpha/2}$
Los valores de z , z_α y $z_{\alpha/2}$ se obtienen de las tablas de la distribución $N(0,1)$, como la que se presenta en la tabla A2 del capítulo 1.	
Suposiciones: La muestra se seleccionó aleatoria e independientemente de una población distribuida normalmente o bien para muestras con $n > 30$.	

**TABLA 3.6 COMPARACION DE LA MEDIA MUESTRAL CON UN VALOR ESPECIFICADO,
CASO σ DESCONOCIDA**

1. Hipótesis nula: $H_0: \mu = \mu_0$

2. Hipótesis alternativa

Prueba de una cola

$$H_1: \mu > \mu_0$$

(o bien $\mu < \mu_0$)

Prueba de dos colas

$$H_1: \mu \neq \mu_0$$

3. Estadístico de contraste: $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$

4. Región de rechazo:

Prueba de una cola

$$t > t_{\alpha}$$

(o sea, $t < -t_{\alpha}$ cuando la hipótesis

alternativa es $H_1: \mu < \mu_0$)

Prueba de dos colas

$$t > t_{\alpha/2} \text{ o bien } t < -t_{\alpha/2}$$

Los valores de t , t_{α} y $t_{\alpha/2}$ se basan en $(n-1)$ grados de libertad. Estos valores están tabulados en el anexo A del capítulo 2.

Suposiciones: La muestra se seleccionó aleatoria e independientemente de una población distribuida normalmente con $n < 30$.

TABLA 3.7 COMPARACION DE LAS MEDIAS MUESTRALES DE DOS MUESTRAS, CASO σ_1 Y σ_2 DADAS

1. Hipótesis nula: $H_0: \mu_1 - \mu_2 = D_0$, donde D_0 es una diferencia especificada que se quiere probar. Para muchas pruebas tal vez se desee probar que no hay una diferencia entre μ_1 y μ_2 , es decir $D_0 = 0$.

2. Hipótesis alternativa

Prueba de una cola

$$H_1: (\mu_1 - \mu_2) > D_0$$

(o bien $(\mu_1 - \mu_2) < D_0$)

Prueba de dos colas

$$H_1: (\mu_1 - \mu_2) \neq D_0$$

3. Estadístico de contraste:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

4. Región de rechazo:

Prueba de una cola

$$Z > Z_\alpha$$

(o sea, $Z < -Z_\alpha$ cuando la hipótesis alternativa es $H_1: \mu_1 < \mu_2$)

Prueba de dos colas

$$|z| > Z_{\alpha/2}$$

Los valores de z , z_α y $z_{\alpha/2}$ se obtienen de las tablas de la distribución $N(0,1)$, como la que se presenta en la tabla A2 del capítulo 1.

Suposiciones: Las muestras se seleccionaron aleatoriamente de poblaciones distribuidas normalmente con $n > 30$.

TABLA 3.8 COMPARACION DE LAS MEDIAS MUESTRALES DE DOS MUESTRAS, CASO σ_1 Y σ_2 DESCONOCIDAS Y DIFERENTES

1. Hipótesis nula: $H_0: \mu_1 - \mu_2 = D_0$, donde D_0 es una diferencia especificada que se quiere probar. Para muchas pruebas tal vez se desee probar que no hay una diferencia entre μ_1 y μ_2 , es decir $D_0 = 0$.

2. Hipótesis alternativa

Prueba de una cola

$$H_1: (\mu_1 - \mu_2) > D_0$$

(o bien $(\mu_1 - \mu_2) < D_0$)

Prueba de dos colas

$$H_1: (\mu_1 - \mu_2) \neq D_0$$

3. Estadístico de contraste:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Grados de libertad:

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2 / n_1)^2}{n_1 + 1} + \frac{(s_2^2 / n_2)^2}{n_2 + 1}} - 2$$

4. Región de rechazo:

Prueba de una cola

$$t > t_{\alpha}$$

(o sea, $t < -t_{\alpha}$ cuando la hipótesis alternativa es $H_1: \mu_1 < \mu_2$)

Prueba de dos colas

$$t > t_{\alpha/2} \text{ o bien } t < -t_{\alpha/2}$$

Los valores de t , t_{α} y $t_{\alpha/2}$ se basan en v grados de libertad. Estos valores están tabulados en el anexo A del capítulo 2.

Suposiciones: Las muestras se seleccionaron aleatoria e independientemente de dos poblaciones distribuidas normalmente. Las varianzas poblacionales, σ_1 y σ_2 , son diferentes.

TABLA 3.9 COMPARACION DE LAS MEDIAS MUESTRALES DE DOS MUESTRAS, CASO σ_1 Y σ_2 DESCONOCIDAS E IGUALES

1. Hipótesis nula: $H_0: \mu_1 - \mu_2 = D_0$, donde D_0 es una diferencia especificada que se quiere probar. Para muchas pruebas tal vez se desee probar que no hay una diferencia entre μ_1 y μ_2 , es decir $D_0 = 0$.

2. Hipótesis alternativa

Prueba de una cola

$$H_1: (\mu_1 - \mu_2) > D_0$$

(o bien $(\mu_1 - \mu_2) < D_0$)

Prueba de dos colas

$$H_1: (\mu_1 - \mu_2) \neq D_0$$

3. Estadístico de contraste:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Grados de libertad:

$$v = n_1 + n_2 - 2$$

Desviación estándar ponderada S_p :

$$S_p^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 + \sum_{i=1}^{n_2} (x_i - \bar{x}_2)^2}{n_1 + n_2 - 2} \quad \text{ó}$$

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

4. Región de rechazo:

Prueba de una cola

$$t > t_\alpha$$

(o sea, $t < -t_\alpha$ cuando la hipótesis alternativa es $H_1: \mu_1 < \mu_2$)

Prueba de dos colas

$$t > t_{\alpha/2} \text{ o bien } t < -t_{\alpha/2}$$

Los valores de t , t_α y $t_{\alpha/2}$ se basan en v grados de libertad. Estos valores están tabulados en el anexo A del capítulo 2.

Suposiciones: Las muestras se seleccionaron aleatoria e independientemente de dos poblaciones distribuidas normalmente. Las varianzas poblacionales, σ_1 y σ_2 , son iguales. En el caso de que la suposición de igualdad de varianzas de las dos poblaciones no se cumpla, la prueba t de varianzas iguales es robusta cuando los tamaños de las dos muestras no difieren. Dicha prueba también es robusta frente a desviaciones moderadas de la normalidad cuando el tamaño de las muestras son grandes.

3.3.6.2 COMPARACION DE VARIANZAS MUESTRALES

La comparación de varianzas muestrales puede ser de dos tipos:

1. Comparar una varianza (o desviación típica) con un valor especificado (σ_0).
2. Comparar dos varianzas (o desviaciones típicas) σ_1 y σ_2 .

Las consideraciones a tomar en cuenta en la recolección de datos para asegurar la aproximación al cumplimiento de los supuestos tomados en este tipo de contrastes de hipótesis son las mismas que las sugeridas en la sección 3.3.6.1 para la comparación de promedios. En las tablas 3.10 y 3.11 se describen los componentes principales de los contrastes de hipótesis antes mencionados.

TABLA 3.10 COMPARACION DE LA VARIANZA MUESTRAL CON UN VALOR ESPECIFICADO σ_0

1. Hipótesis nula: $H_0: \sigma^2 = \sigma_0^2$	
2. Hipótesis alternativa	
Prueba de una cola	Prueba de dos colas
$H_1: \sigma^2 > \sigma_0^2$ (o bien $\sigma^2 < \sigma_0^2$)	$H_1: \sigma^2 \neq \sigma_0^2$
3. Estadístico de contraste: $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	
4. Región de rechazo:	
Prueba de una cola	Prueba de dos colas
$\chi^2 > \chi_{\alpha}^2$	$\chi^2 > \chi_{\alpha/2}^2$ o sea $\chi^2 < \chi_{(1-\alpha/2)}^2$
(o bien, $\chi^2 < \chi_{(1-\alpha)}^2$ cuando la hipótesis alternativa es $H_1: \sigma^2 < \sigma_0^2$)	Donde $\chi_{\alpha/2}^2$ y $\chi_{(1-\alpha/2)}^2$ son los valores de la cola superior e inferior, respectivamente, de χ^2 que ponen $\alpha/2$ en las áreas de las colas.
χ_{α}^2 y $\chi_{(1-\alpha)}^2$ son los valores de la cola superior e inferior, respectivamente de χ^2 que ponen α en las áreas de las colas.	Los valores críticos de χ^2 se basan en (n-1) grados de libertad. Estos valores tabulados se dan en el anexo B del capítulo 2.
Suposiciones: La muestra se seleccionó aleatoriamente de una población distribuida normalmente.	

TABLA 3.11 COMPARACION DE DOS VARIANZAS σ_1 Y σ_2

1. Hipótesis nula: $H_0: \sigma_1^2 = \sigma_2^2$	
2. Hipótesis alternativa	
Prueba de una cola	Prueba de dos colas
$H_1: \sigma_1^2 > \sigma_2^2$ (o bien $\sigma_1^2 < \sigma_2^2$)	$H_1: \sigma_1^2 \neq \sigma_2^2$
3. Estadístico de contraste: $F = \frac{S_1^2}{S_2^2}$	
En donde S_1^2 , es la varianza muestral mayor	
4. Región de rechazo:	
Prueba de una cola	Prueba de dos colas
$F > F_{\alpha, n_1-1, n_2-2}$ ó $F < F_{1-\alpha, n_1-1, n_2-2}$	$F > F_{\alpha/2, n_1-1, n_2-1}$ ó $F < F_{1-\alpha/2, n_1-1, n_2-1}$
Los valores críticos de F_α y $F_{\alpha/2}$ se basan en $(n_1 - 1)$ grados de libertad en el numerador y $(n_2 - 1)$ grados de libertad en el denominador. Estos valores se encuentran tabulados, para $\alpha = 0.05, 0.01, 0.025, 0.975$, en el anexo C del capítulo 2.	Los valores críticos de $F_{\alpha/2}$ y $F_{1-\alpha/2}$ se basan en $(n_1 - 1)$ grados de libertad en el numerador y $(n_2 - 1)$ grados de libertad en el denominador. Estos valores se encuentran tabulados, para $\alpha = 0.05, 0.01, 0.025, 0.975$, en el anexo C del capítulo 2.
<i>Suposiciones:</i> Las dos muestras se seleccionaron aleatoria e independientemente de dos poblaciones con distribución normal.	

3.3.6.3 COMPARACION DE DATOS POR PARES

En una comparación del comportamiento de un sistema sometido a dos tratamientos puede considerarse que en los datos obtenidos influyen cuatro tipos de factores:

- I. El factor cuyo efecto se desea estudiar.
- II. Factores identificados que pueden influir en la respuesta, pero que es posible mantener constantes durante la extracción de las dos muestras.
- III. Factores identificados que pueden influir en la respuesta y que resulta imposible mantener constantes para los dos tratamientos. Esto hace necesario bloquear, es

decir analizar diferencias dentro de bloques homogéneos en los que los factores afectan por igual.

- IV. Otros factores no identificados, que también pueden influir en los resultados obtenidos (respuesta). Lo anterior, implica la aleatorización.

Si se considera un único factor del tercer tipo que influye sobre la respuesta y del cual no puede asegurarse que afecte exactamente igual a los dos tratamientos, es necesario "bloquear" (crear bloques), es decir analizar las diferencias dentro de los bloques homogéneos en que los factores afectan por igual. Por ejemplo, con frecuencia ocurre que deben compararse dos métodos de análisis por medio del estudio de muestras que contienen, de manera sustancial, diferentes cantidades de analito. El problema consiste en estudiar la eficiencia de los métodos de análisis independientemente de las concentraciones analito que pueden contener las muestras de diferente procedencia. En general, la consigna en el diseño de la recolección de datos es bloquear lo que sea necesario y aleatorizar el resto.

Lo anterior sugiere que los pares de datos se recolecten por bloque considerando que este afecta por igual a los dos tratamientos. La idea básica es que las diferencias dentro de los distintos bloques *fluctúan menos que los valores individuales*, ello va a hacer posible llevar a cabo comparaciones mucho más precisas, y el análisis se hará sobre los valores de dichas diferencias. Las diferencia entre los tratamientos para los distintos bloques se definen como:

$$d_i = x_{i1} - x_{i2}$$

En la tabla 3.12 se presenta un resumen para la prueba de hipótesis de comparación de tratamientos considerando el efecto de bloque.

TABLA 3.12 PRUEBA DE HIPOTESIS PARA DATOS POR PARES ($\mu_1 - \mu_2 = \mu_d$)

1. Hipótesis nula: $H_0: \mu_d = 0$

2. Hipótesis alternativa

Prueba de una cola

$$H_1: \mu_d > 0$$

(o bien $\mu_d < 0$)

Prueba de dos colas

$$H_1: \mu_d \neq 0$$

3. Estadístico de contraste:

$$t = \frac{\bar{d} - 0}{s_d / \sqrt{n}} = \frac{\bar{d}}{s_d / \sqrt{n}}$$

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i \quad \text{y} \quad s_d = \sqrt{\frac{\sum_{i=1}^n d_i^2 - \frac{\left(\sum_{i=1}^n d_i\right)^2}{n}}{n-1}}$$

$$s_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}}$$

4. Región de rechazo:

Prueba de una cola

$$t > t_{\alpha}$$

(o sea, $t < -t_{\alpha}$ cuando la hipótesis alternativa es $H_1: \mu_d < 0$)

Prueba de dos colas

$$t > t_{\alpha/2} \text{ o bien } t < -t_{\alpha/2}$$

Los valores de t , t_{α} y $t_{\alpha/2}$ se basan en $(n-1)$ grados de libertad. Estos valores están tabulados en el anexo A del capítulo 2.

Suposiciones: Se seleccionan aleatoriamente las n diferencias por parejas de una población distribuida normalmente.

Si se quiere evaluar las diferencias entre los parámetros de más de dos grupos, por ejemplo la durabilidad de distintos tipos de cubiertas, no es correcto realizar tales evaluaciones usando el contraste de diferencias de medias tomando los diferentes pares de medias. Para solucionar este problema si las mediciones resultantes son continuas y se cumplen las siguientes suposiciones: i) los datos son obtenidos de manera aleatoria e independiente o que los individuos sean asignados a los grupos aleatoriamente, ii) los valores de cada grupo están distribuidos normalmente, iii) la varianza dentro de cada

población debería ser igual para todas las poblaciones, existe una metodología conocida como análisis de varianza (ANOVA) para comparar la medias de los grupos y cuya hipótesis nula es $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$. Por otra parte, si los supuestos ii) o iii) no se cumplen existen metodologías que permiten todavía llevar cabo las comparaciones deseadas (métodos no paramétricos).

3.3.6.4 CONTRASTES PARA LOS PARAMETROS BINOMIALES p Y $(p_1 - p_2)$ Y PARA EL PARAMETRO DE LA DISTRIBUCION DE POISSON λ

Sea $p = x/n$ la proporción poblacional estimada del número de veces que aparece el suceso de Bernoulli de entre n repeticiones de un experimento ($x = n^\circ$ de veces que aparece el suceso). Si $np(1-p) > 9$ el estadístico $z = \frac{X + 1/2 - np}{\sqrt{np(1-p)}}$ se aproxima a una distribución Normal(0, 1), siendo $X = n^\circ$ de veces que aparece en la muestra el suceso considerado.

Cuando la población sigue una distribución de Poisson de parámetro λ con $\lambda > 9$, el estadístico $z = \frac{X + 1/2 - \lambda}{\sqrt{\lambda}}$ se aproxima a una distribución Normal(0, 1) (PEREZ, 1999, cap.5).

En la tabla 3.13 se presentan los elementos para el contraste de hipótesis de parámetros binomiales y de Poisson.

TABLA 3.13 CONTRASTES DE HI POTESIS PARA PARÁMETROS BINOMIALES Y DE POISSON

HI POTESIS	ESTADISTICO	REGIONES CRITICAS
$H_0: p = p_0$ $H_1: p \neq p_0$ $H_1: p < p_0$ $H_1: p > p_0$	$z = \frac{X + 1/2 - np_0}{\sqrt{np_0(1-p_0)}}$	$ z > Z_{\alpha/2}$ $z < -Z_{\alpha}$ $z > Z_{\alpha}$
$H_0: \lambda = \lambda_0$ $H_1: \lambda \neq \lambda_0$ $H_1: \lambda < \lambda_0$ $H_1: \lambda > \lambda_0$	$z = \frac{X + 1/2 - \lambda_0}{\sqrt{\lambda_0}}$	$ z > Z_{\alpha/2}$ $z < -Z_{\alpha}$ $z > Z_{\alpha}$
$H_0: p_1 = p_2$ $H_1: p_1 \neq p_2$ $H_1: p_1 < p_2$ $H_1: p_1 > p_2$	$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$	$ z > Z_{\alpha/2}$ $z < -Z_{\alpha}$ $z > Z_{\alpha}$

Los valores de \hat{p}_1 y \hat{p}_2 del cuadro anterior representan las proporciones de éxitos en dos muestras grandes independientes de tamaños n_1 y n_2 respectivamente, tomadas de dos poblaciones binomiales.

3.4 DETERMINACIÓN DE DATOS ANOMALOS

Frecuentemente en los resultados de una serie de mediciones se observa que una o más de ellas parecen ser muy improbablemente grandes o pequeñas respecto de la mayoría de mediciones de la muestra. Tal medición (o mediciones) recibe el nombre de *valor inusitado* o *anómalo* o "outlier". Es decir, los datos anómalos o outliers son las observaciones que

aparentan ser inconsistentes con el resto de los datos colectados. Un valor anómalo puede ser consecuencia de:

- a. *Un evento poco probable pero que ocurre en alguna ocasión.*
- b. *Una medición que provenga de una muestra que realmente no forme parte de la población muestreada (por ejemplo, debido a una contaminación).*
- c. *Una medición errónea.*

El término de dato anómalo o outlier se usa colectivamente tanto para observaciones discordantes (caso c), como para contaminantes (casos a y b).

La presencia de un valor (o valores) anómalos hace surgir en el químico analista la siguiente interrogante: ¿Es conveniente conservar este dato en la muestra o debe rechazarse?

En general, la elección de criterios para el rechazo de valores anómalos tiene sus peligros. Si se establece un criterio poco riguroso que haga difícil eliminar un resultado dudoso, se corre el riesgo de retener resultados falsos que pueden tener un efecto errático sobre las estimaciones e inferencias que deban hacerse con dichos datos. Si por el contrario hay mucha rigurosidad en el criterio elegido, posiblemente se descarten mediciones que son correctas dentro del conjunto, lo que podría llevar a una tendencia en los datos. Por desgracia, no hay una regla universal que resuelva el asunto del rechazo o retención de resultados. Sin embargo, existen dos caminos para administrar la presencia de los datos anómalos. En el laboratorio, es recomendable el mantenimiento de un buen registro de cada experimento (aseguramiento de calidad), en donde todos los datos deberían ser registrados con cualquier posible explicación o información adicional. En el análisis de los datos, existen varios métodos estadísticos robustos recomendados, como contrastes de hipótesis, que permiten detectar con cierta probabilidad la presencia de valores anómalos y en los que puede definirse el siguiente contraste de hipótesis:

H_0 : no hay datos anómalos en el conjunto de datos, y

H_1 : hay al menos un dato anómalo en el conjunto de datos.

Entre estos métodos se encuentran: a) La prueba GRAF/HENNING (ó z-score) ($n > 25$); b) La prueba "Q" de Dixon; c) La prueba t; y d) La prueba de GRUBBS ($3 < n < 150$); y e) La prueba de HUBER.:

A continuación se exponen los procedimientos para las pruebas antes mencionadas:

a. **PRUEBA DE GRAF/ HENNING** (MENDENHALL, 1990, cap. 2)

Esta prueba se aplica cuando se cuenta con un número de más de 25 mediciones, y consiste en comparar si el valor del dato anómalo (x) excede el rango:

$$\bar{X} \pm 4s$$

en donde \bar{X} y s son el promedio y desviación estándar de la muestra sin tomar en cuenta el valor anómalo (x).

b. **PRUEBA "Q" DE DIXON** (BODE, 1999)

La prueba Q es buena para sistemas con datos relativamente dispersos o cuando el control estadístico no está todavía establecido, es bueno para detectar un pequeño número de datos anómalos, y se usa para muestras de tamaños que oscilen entre 3 y 25 datos. El procedimiento para la prueba es el siguiente:

1. Definir el rango de los datos existentes: $RA = X_{MAX} - X_{MIN}$
2. Encontrar la distancia entre el valor anómalo (X_a) y el valor más cercano a él existente (X_c), $DI = \text{abs}(X_a - X_c)$.
3. Calcule el cociente $QU = DI/RA$ y compare con un valor tabulado.
4. Si el cociente es más pequeño que ese valor tabulado en la tabla del anexo B3 para un nivel de confianza dado, no hay razón para creer que el dato sospechoso es un dato anómalo.

La prueba de Dixon puede repetirse para evaluar otros outliers, sin embargo, el poder de la misma disminuye cuando el número de repeticiones incrementa.

c. **PRUEBA "t"** (APHA, 1996, cap. 1)

Este criterio se aplica a un grupo de 3 o más resultados, el procedimiento para la prueba "t" es el siguiente:

1. Se encuentra el promedio (\bar{x}) y la desviación estándar (s) de los datos.
2. Se obtiene la distancia entre el valor sospechoso (X_a) y el promedio (\bar{x}): $D_p = \text{abs}(X_a - \bar{x})$.
3. Se obtiene el cociente $t = D_p/R$.
4. Si el valor de t es menor que el valor tabulado en la tabla del anexo B3 para un nivel confianza dado, no hay razón para creer que el dato debe rechazarse. Esta prueba puede emplearse además, para determinar si un nuevo resultado obtenido pertenece a la distribución establecida por una serie anterior, sólo que en este caso el valor de t tabulado deberá obtenerse de la tabla de distribución t del anexo C3 para el nivel de significación propuesto en la prueba.

d. **PRUEBA DE GRUBBS** (GRUBBS, 1979)

La prueba de Grubbs, también conocida como la prueba del máximo residuo normalizado (esta prueba es la recomendada por la EPA como la prueba estadística para la determinación de outliers (US EPA, 1992)), se utiliza cuando un sistema esta bajo un buen control estadístico, para una muestra de tamaño n en el intervalo de 3 a 150 ($3 < n < 150$). La prueba de Grubbs detecta un dato anómalo cada vez. El dato anómalo es eliminado del set de datos y la prueba es iterada hasta que no se detectan más datos anómalos. Sin embargo, las iteraciones múltiples cambian las probabilidades de detección, y la prueba de esta forma no puede usarse para muestras de tamaño seis o menos.

El procedimiento para la prueba es el siguiente:

1. Calcule la desviación estándar (s) de los datos existentes.
2. Definir la distancia $DM = |\bar{x} - x|$, donde x es el dato sospechoso
3. Calcular el cociente $G = DM/s$, para pruebas iterativas el estadístico de Grubbs se define como $G = \max |\bar{x} - x| / s$

4. Si $G > \frac{(n-1)}{\sqrt{n}} \sqrt{\frac{t_{(\alpha/(2n), n-2)}^2}{n-2 + t_{(\alpha/(2n), n-2)}^2}}$, en donde $t_{(\alpha/(2n), n-2)}^2$ es el valor crítico de la

distribución t con (n-2) grados de libertad y un nivel de significación $\alpha/(2n)$, puede concluirse que el dato es anómalo o outlier. También puede compararse G con los valores críticos ($G_{\alpha, n}$) presentados en la tabla del anexo D3 para diferentes tamaños muestrales.

e. **PRUEBA DE HUBER (z-score modificado)** (BODE, 1999)

Consiste en calcular la mediana de las desviaciones absolutas (MAD) de los datos individuales con respecto a la mediana, es decir:

$$\text{DESVIACIONES ABSOLUTAS} = |x_{\text{obs}} - \text{mediana}|$$

Se dice que un valor es un dato anómalo si no cae dentro del intervalo $\text{MEDIANA} \pm k_{\text{mad}}$, con $k = 3.5$ (que corresponde a un error en la estimación del 5%, con un grado de confianza del 95%). Ó bien, un dato es un anómalo si su residuo absoluto es mayor que $5.06k_{\text{MAD}}$ (a un nivel de confianza del 95%).

Esta es una prueba confiable dado que los parámetros usados para calcular el estadístico están mínimamente afectados por la presencia de datos anómalos outliers.

Como puede observarse se han desarrollado muchas pruebas estadísticas como criterios de rechazo o retención de resultados discordantes. No obstante la mayoría de ellas asume normalidad por lo que es recomendable verificar primero que los datos se distribuyen razonablemente con una distribución normal. Estas pruebas para datos normales son potentes y fáciles de usar, sin embargo, las pruebas para datos no normales son menos potentes y más difíciles (Iglewicz, 1993). Algunas de estas pruebas están incluidas en Barnett y Lewis (1984). En muchas situaciones los datos pueden transformarse en distribuciones aproximadamente normales y analizarse luego las técnicas presentadas anteriormente (Ejemplo, muchas mediciones ambientales son log-normales) (SKOOG et al., 2000, cap. 6).

Por otra parte, estas pruebas deben usarse con extrema precaución cuando se aplican a muestras que sólo tienen unos cuantos datos. Mandel (1978), en su análisis del tratamiento conjuntos pequeños de datos, escribe "Quienes crean que con la sanción estadística pueden descartar observaciones usando pruebas estadísticas para rechazar datos discordantes, simplemente, se engañan a sí mismos". De tal forma, las pruebas estadísticas para el rechazo de datos sólo deben usarse como ayuda al sentido común cuando se tienen pocos datos (SKOOG et al., 2000, cap. 6).

La aplicación indiscriminada de las pruebas estadísticas para retener o descartar una medición dudosa en un conjunto pequeño de datos no puede ser más confiable que una decisión arbitraria. Una mejor estrategia es aplicar un buen criterio basado en una amplia experiencia del método analítico. Finalmente, la única razón válida para descartar un resultado de un conjunto pequeño de datos es la certidumbre de que se cometió un error durante el proceso de medición. Sin este conocimiento, *es recomendable ser muy precavidos al rechazar los datos discordantes*⁴ (SKOOG et al., 2000, cap. 6).

Las siguientes son algunas recomendaciones para el tratamiento de un conjunto pequeño de datos que contenga un valor dudoso (SKOOG et al., 2000, cap. 6):

- Volver a examinar cuidadosamente todos los datos relacionados con el valor dudoso para ver si un error grueso pudo haber alterado su valor.
- De ser posible tratar de estimar la incertidumbre que razonablemente puede esperarse de un procedimiento para asegurarse que el resultado discordante es realmente dudoso.
- Si se dispone de muestra y tiempo suficientes, conviene repetir el análisis. La concordancia entre el nuevo dato obtenido y los que en el conjunto original se consideran como válidos, apoyarán la decisión de rechazar el dato discordante.
- Si no pueden obtenerse más datos, deben aplicarse las pruebas de rechazo de resultados (que se consideren convenientes), para establecer sobre bases

⁴ Si las observaciones son determinadas estadísticamente como anómalas, la EPA sugiere determinar una explicación para estos datos antes de su exclusión de los análisis posteriores (US EPA, 1992). Si no puede encontrarse una explicación, la observación debe tratarse como una medición extrema pero válida que debe mantenerse en los análisis posteriores.

estadísticas si el resultado dudoso debe rechazarse o no. Es preferible aplicar la pruebas paramétricas siempre y cuando se tenga la confianza en la normalidad de los datos, de lo contrario es recomendable aplicar pruebas no paramétricas.

- Si las pruebas indican retener el dato, hay que considerar expresar los resultados en términos de la mediana y no como la media. La mediana tiene la gran ventaja de permitir que se incluyan todos los datos de un conjunto sin que influya mucho un valor discordante (recuerde que la mediana es un estimador "robusto").

3.5 PROBLEMAS DE INFERENCIA ESTADÍSTICA

3.5.1 PROBLEMAS DE ESTIMACIÓN

1. Un método de absorción atómica para la determinación de cobre en combustibles tiene una desviación estándar ponderada de $\sigma = 0.32 \mu\text{g}$ de Cu/ml. El análisis de una muestra de combustible mostró un contenido de cobre de $8.53 \mu\text{g}$ de Cu/ml. Calcule los límites de confianza al 90 y 95% para el resultado, si se basa en: a) Un solo análisis; b) El promedio de tres análisis.

Solución

El problema expuesto se trata de la creación de intervalos de confianza para el valor medio en la determinación de cobre en muestras de combustible. Dado que se tiene el valor de σ , de la tabla 3.2, se obtiene entonces que el intervalo de confianza viene dado como:

$$\bar{x} \pm \frac{Z_{\alpha/2} \sigma}{\sqrt{n}}$$

\bar{x} : promedio muestral; $Z_{\alpha/2}$: valor de la distribución $N(0,1)$, para el nivel de significación $\alpha/2$; σ : desviación típica de la población; n : tamaño de la muestra.

A continuación se presenta la solución de cada literal:

- a. De la tabla del anexo A1, $Z_{(1-\alpha)} = Z_{90\%} = 1.64$ y , $Z_{(1-\alpha)} = Z_{95\%} = 1.96$. De esta forma:

El intervalo de confianza al 90% es:

$$I.C._{(90\%)} = (8.53 \mu\text{g de Cu/ml}) \pm \frac{1.64 \times (0.32 \mu\text{g de Cu / ml})}{\sqrt{1}}$$

$$I.C._{(90\%)} = (8.53 \pm 0.52) \mu\text{g de Cu / ml}$$

Y el intervalo de confianza al 95% es:

$$I.C._{(95\%)} = (8.53 \mu\text{g de Cu/ml}) \pm \frac{1.96 \times (0.32 \mu\text{g de Cu / ml})}{\sqrt{1}}$$

$$I.C._{(95\%)} = (8.53 \pm 0.63) \mu\text{g de Cu / ml}$$

De los cálculos anteriores puede concluirse que existe un 90% de probabilidad que el intervalo de 8.01 y 9.05 $\mu\text{g de Cu/ml}$ contenga el valor medio de la concentración de cobre del tipo de gasolina analizada. Así mismo la probabilidad de que el intervalo 7.9 y 9.16 $\mu\text{g de Cu/ml}$ contenga el promedio de la concentración media de cobre es del 95%.

b) Para las tres mediciones

el intervalo de confianza al 90% viene dado por:

$$I.C._{(90\%)} = (8.53 \mu\text{g de Cu/ml}) \pm \frac{1.64 \times (0.32 \mu\text{g de Cu / ml})}{\sqrt{3}}$$

$$I.C._{(90\%)} = (8.53 \pm 0.30) \mu\text{g de Cu / ml}$$

Mientras que el intervalo de confianza al 95% es:

$$I.C._{(95\%)} = (8.53 \mu\text{g de Cu/ml}) \pm \frac{1.96 \times (0.32 \mu\text{g de Cu / ml})}{\sqrt{3}}$$

$$I.C._{(95\%)} = (8.53 \pm 0.36) \mu\text{g de Cu / ml}$$

De los cálculos anteriores puede concluirse que para esta muestra de tamaño 3, existe un 90% de probabilidad de que el intervalo de 8.23 y 8.83 $\mu\text{g de Cu/ml}$ contenga valor medio de la concentración de cobre en la gasolina del análisis. Asimismo la probabilidad de que el intervalo 8.17 y 8.89 $\mu\text{g de Cu/ml}$ contenga el promedio de la concentración de cobre es del 95%.

Como puede observarse a medida que se aumenta el tamaño muestral se reduce el intervalo creado para ambos niveles de confianza.

2. Teniendo en cuenta una amplia experiencia previa, se ha encontrado que la desviación estándar de un método para el análisis de monóxido de carbono en los gases de escape de los automóviles es de 0.80 ppm. a) Estimar el intervalo de confianza al 95% para un análisis por triplicado; b) ¿Cuántas mediciones se necesitarían para que el intervalo de confianza del 90% sea de 0.50 ppm?

Solución

a) En este caso debe determinarse el ancho del intervalo para un grado de confianza del 95%. Las características de la muestra son las siguientes:

$$n = 3; \sigma = 0.80 \text{ ppm}$$

Dado que se cuenta con el valor de σ , el intervalo de confianza y el ancho de dicho intervalo, a un determinado nivel de confianza, vienen definidos por:

$$\bar{x} \pm \frac{z_{\alpha/2} \sigma}{\sqrt{n}}$$

De la tabla para la distribución normal tipificada (anexo A1) se ve que z para el 95% de probabilidad es 1.96, por lo tanto sustituyendo en la ecuación se tiene:

$$\bar{X} \pm \left(\frac{1.96 * 0.80 \text{ ppm}}{\sqrt{3}} \right)$$

$$\bar{X} \pm \mathbf{0.905 \text{ ppm}}$$

es decir que el ancho del intervalo es de **0.905 ppm**

b) Para un ancho de intervalo de 0.50 ppm, y un nivel de confianza del 90% ($Z_{90\%} = 1.64$), el tamaño de muestra vendrá dado por:

$$z_{\alpha/2} \sigma / \sqrt{n} = \pm 0.5 \text{ ppm}$$

$$\sqrt{n} = (1.64 * 0.80 \text{ ppm}) / 0.5 \text{ ppm} = 2.624$$

$$n = \mathbf{6.88}$$

es decir que deben efectuarse **7 análisis**

3. Un químico obtuvo los siguientes datos para el porcentaje de lindano en el análisis de un insecticida: 7.47, 6.98, 7.27. Calcular el límite de confianza al 90% para la media de los datos suponiendo que la única información que se tiene acerca de la precisión del método es la desviación estándar para los tres datos.

Solución:

Se quiere determinar un intervalo de confianza para el porcentaje medio de lindano en muestras de pesticidas, cuando se hacen tres mediciones ($n = 3$) y no se cuenta con el valor de σ (σ desconocida). En este caso el intervalo de confianza para la media viene dado como:

$$I.C._{(1-\alpha)} = \bar{x} \pm \frac{t_{\alpha/2, n-1} s}{\sqrt{n}}$$

\bar{x} : promedio muestral; $t_{\alpha/2, n-1}$: coeficiente de la distribución t-student para el nivel de significación $\alpha/2$ y con $n-1$ grados de libertad; s : desviación típica de la de la muestra; n : tamaño de la muestra.

Los estadísticos para las tres mediciones vienen dados por:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{7.47 + 6.98 + 7.27}{3} = 7.24\%$$

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(7.47 - 7.24)^2 + (6.98 - 7.24)^2 + (7.27 - 7.24)^2}{3-1}} = 0.25\%$$

El valor de la variable t para $\alpha/2 = 0.05/2 = 0.025$ y $n-1=3-1=2$ grados de libertad es **4.3**. Por tanto, el intervalo de confianza al 95% viene dado como:

$$I.C._{(95\%)} = (7.24\%) \pm \frac{4.3 \times (0.25\%)}{\sqrt{3}}$$

$$I.C._{(95\%)} = (7.24 \pm 0.62) \% \text{ es decir } [6.62\%, 7.86\%]$$

El valor promedio del porcentaje de lindano en las muestras de pesticidas oscila entre 6.62% y 7.86%, con un grado de confianza en la estimación del 95%.

4. En el monitoreo de la contaminación de un río se ha determinado un nivel contaminación considerable con cromo en 10 de los noventa días monitoreados. Construir un intervalo de confianza a un nivel de confianza del 95% para la probabilidad de que se de una elevada concentración de cromo bajo las condiciones de monitoreo establecidas.

Solución:

El intervalo de confianza para una proporción viene dado como (ver tabla 3.4):

$$p \pm z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}} \quad \text{ó} \quad p \pm z_{\alpha/2} \sigma_p$$

La proporción de contaminación es $p = 10/90 = 0.11$, por tanto el valor estimado de s_p es:

$$s_p = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.11(0.89)}{90}} = 0.033$$

De la tabla del anexo A1, el valor de Z al 95% de confianza es 1.96. Los límites del intervalo son por tanto:

$$\text{Limite inferior} = 0.11 - (1.96 * 0.033) = 0.04$$

$$\text{Limite superior} = 0.11 + (1.96 * 0.033) = 0.169$$

O sea que la probabilidad de contaminación con cromo del río oscila entre $0.04 < p < 0.17$ con un nivel de confianza del 95%.

3.5.2 PROBLEMAS DE CONTRASTE DE HI POTESI S

1. El tanto por ciento de níquel de una muestra particular de acero de referencia del NIST, es de un 1.12%. Un nuevo método espectrofotométrico para la determinación de níquel dio los siguientes resultados sobre muestras de dicho acero de referencia: 1.10, 1.08, 1.09, 1.12, 1.109. ¿Hay algún indicio de error sistemático del método para un nivel de significación del 5%?;

Solución:

En general, este problema consiste en determinar si existe evidencia de error sistemático del nuevo método empleado para la determinación de níquel. Esta

determinación se hará mediante un contraste de hipótesis bilateral de comparación de un promedio muestral con un valor especificado, en este caso, no se cuenta con una estimación de σ .

El contraste de hipótesis es por tanto:

$$H_0: \mu = 1.12\% \text{ (no existe evidencia de error sistemático)}$$

$$H_1: \mu \neq 1.12\% \text{ (si existe evidencia de error sistemático)}$$

El estadístico de contraste y el criterio de rechazo para esta prueba vienen dados, según la tabla 3.7, por:

$$t_o = \frac{(\bar{X} - \mu_o)}{s/\sqrt{n}} \quad \text{y} \quad |t_o| > t_{\alpha/2, n-1}$$

La media y la desviación típica para la muestra son los siguientes:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{1.10 + 1.08 + 1.09 + 1.12 + 1.09}{5} = 1.096\%$$

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1}} = 0.0152\%$$

De esta forma el valor del estadístico de contraste para la muestra es:

$$t_o = \frac{(1.096 - 1.12)}{0.0152/\sqrt{5}} = \frac{-0.024}{0.0068} = -3.53$$

de la tabla del anexo A2, se tiene:

$$t_{\alpha/2, n-1} = t_{0.05/2, 5-1} = t_{0.025, 4} = 2.78$$

Conclusión:

Dado que el valor observado de t ($|t_o| = 3.53$) es mayor que el valor crítico ($t_{0.025, 4} = 2.78$), existe evidencia significativa al 5% de error sistemático en el nuevo método espectrofotométrico para la determinación de níquel en muestras de acero.

2. Un método volumétrico para la determinación de calcio en piedra caliza se comprobó analizando una piedra caliza del NIST que contenía un 30.15% de CaO. El resultado de la media de cuatro análisis fue de 30.26% de CaO, con una desviación estándar del 0.085%. A partir de la ponderación de diversos análisis, se estableció que $\sigma = 0.094\%$ de CaO. (a) ¿Indican los datos la presencia de un error determinado (sistemático) para un nivel de confianza del 95%?; (b) ¿indican los datos la presencia de un error determinado para un nivel de confianza del 95%, sino se dispusiera para σ de un valor ponderado?

Solución:

La media de los cuatro datos es de 30.26% de CaO, la desviación estándar estimada a partir de diversos análisis es $\sigma = 0.094\%$ de CaO, y el contraste de hipótesis bilateral es el siguiente:

$$H_0: \mu = 30.15\% \text{ (no existe evidencia de error sistemático)}$$

$$H_1: \mu \neq 30.15\% \text{ (si existe evidencia de error sistemático)}$$

a) Dado que se cuenta con una estimación de σ poblacional, el estadístico de contraste, según la tabla 3.6, vendrá dado por:

$$z_o = \frac{(\bar{X} - \mu_o)}{\sigma/\sqrt{n}}$$

Que de acuerdo a la información proporcionada toma un valor de:

$$z_o = \frac{(30.26 - 30.15)\%}{0.094\%/\sqrt{4}} = 2.34$$

El criterio de rechazo para estas condiciones viene dado por: $|z_o| > z_{\alpha/2}$. El valor del $z_{\text{crítico}}$ para un nivel de significación del 5%, según la tabla del anexo A1, es de 1.96.

Conclusión:

Dado que el valor del estadístico calculado ($z_o = 2.34$) es mayor que el $z_{\text{crítico}}$ ($z_{\alpha/2} = 1.96$), existe evidencia significativa a un 5% de error determinado (sistemático) al aplicar el método volumétrico para la determinación de CaO.

b) Para el caso en que no se cuenta con una estimación de σ , el estadístico de contraste y el criterio de aceptación son los siguientes:

$$t_o = \frac{(\bar{X} - \mu_o)}{s/\sqrt{n}} \quad \text{y} \quad |t_o| > t_{\alpha/2, n-1}$$

Por tanto, para la muestra se tiene:

$$t_o = \frac{(30.26 - 30.15)}{0.085/\sqrt{4}} = 2.59$$

de la tabla del anexo A2, se tiene:

$$t_{\alpha/2, n-1} = t_{0.05/2, 4-1} = t_{0.025, 3} = 3.18$$

Conclusión:

Dado que el valor observado de t ($|t_o| = 2.59$) es menor que el valor crítico ($t_{0.025, 3} = 3.18$), no existe evidencia significativa al 5%, como para concluir sobre la presencia de error sistemático en el método volumétrico para la determinación de CaO.

La probabilidad de que una diferencia como esta se deba únicamente a errores aleatorios se puede obtener con la función de Excel DISTR.T (t_o , grados de libertad, n° de colas). El resultado para el problema se presenta a continuación:

$$\text{DISTR.T}(2.59, 3, 2) = 0.08107407$$

probabilidad que al compararla con el valor de α de 0.05, conduce a la misma conclusión descrita en función de la comparación de valores del estadístico calculado y de tablas.

3. La homogeneidad de una muestra patrón de cloruro se evaluó mediante el análisis de porciones de material tomadas en la superficie y en el fondo del contenedor, los resultados obtenidos fueron:

% de cloruro	
Superficie	Fondo
26.32	26.28
26.33	26.25
26.38	26.38
26.39	

- a) ¿Se puede demostrar la no homogeneidad del material a un nivel de

confianza del 95%?

b) ¿Se puede demostrar la no homogeneidad del material con un nivel de confianza del 95% si se sabe que $\sigma = 0.03\%$ de CI?

Solución:

Los promedios y varianzas calculados de acuerdo a los datos para cada porción son los siguientes:

Estadístico	Superficie	Fondo
\bar{X}	26.355	26.30
S^2	0.00123	0.0046

a) Antes de probar si existe falta de homogeneidad del material es conveniente primero efectuar un contraste de homogeneidad de varianzas entre ambas porciones, ya que esto permitirá seleccionar adecuadamente el tipo de estadístico y el criterio de rechazo que deben emplearse para la prueba.

El contraste para la prueba de igualdad de varianzas viene dado por:

$$H_0: \sigma_{fondo}^2 = \sigma_{superficie}^2$$

$$H_1: \sigma_{fondo}^2 \neq \sigma_{superficie}^2$$

Las hipótesis anteriores sugieren un contraste bilateral de la igualdad de varianzas, según la tabla 3.11, el estadístico de contraste y el criterio de rechazo para esta prueba son los siguientes:

Estadístico: $F_o = \frac{s_1^2}{s_2^2}$ en donde s_1^2 es la mayor de las varianzas.

Criterio de rechazo: $F_o > F_{\alpha/2, n_1-1, n_2-1}$ ó $F_o < F_{1-\alpha/2, n_1-1, n_2-1}$

La razón de varianzas para el problema es:

$$F_o = \frac{s_1^2}{s_2^2} = \frac{0.00463}{0.00123} = 3.76$$

y $F_{\alpha/2, n_1-1, n_2-1} = F_{0.025, 2, 3} = 16.04$, $F_{1-\alpha/2, n_1-1, n_2-1} = F_{0.975, 2, 3} = 0.0255$ (según tabla del anexo B2).

Dado que $F_o = 3.76$ no es mayor que $F_{0.025, 2, 3} = 16.04$ ó bien no es menor que $F_{0.975, 2, 3} = 0.0255$, no existe evidencia como para afirmar que la variabilidad en el

% de Cl difiere en ambas porciones con un nivel de significación del 5%. No hay evidencia significativa al 5% como para rechazar a la hipótesis nula. Usando la función PRUEBA.F (matriz1, matriz2) de Excel, para el cálculo del p-valor de la prueba, se tiene:

SUPERFICIE	FONDO
26.32	26.28
26.33	26.25
26.38	26.38
26.39	
P-VALOR=	0.30485271

Dado que el P-Valor de la prueba (=0.3048) es mayor que el nivel de significación de 0.05, no existe evidencia significativa como para rechazar la hipótesis nula. A continuación se presenta el análisis para el contraste a partir de la creación de un intervalo de confianza para la razón de varianzas:

El intervalo de confianza para la razón de varianzas para un nivel de confianza $(1-\alpha)\%$, viene dado por (ver tabla 3.3):

$$\frac{s_1^2/s_2^2}{F_{\alpha/2, n_1-1, n_2-1}} < \sigma_1^2/\sigma_2^2 < \frac{s_1^2/s_2^2}{F_{1-\alpha/2, n_1-1, n_2-1}}$$

De esta forma el intervalo de confianza al 95% ($\alpha = 0.05$) para la razón de varianzas para ambas porciones es:

$$\frac{3.76}{16.04} < \sigma_1^2/\sigma_2^2 < \frac{3.76}{0.0255}$$

$$0.234 < \sigma_1^2/\sigma_2^2 < 147.45$$

Como puede observarse el intervalo de confianza para la razón de varianzas contempla el valor de uno, por lo que puede concluirse que no existe suficiente evidencia como para afirmar que las varianzas del contenido de %Cl para ambas porciones del material son diferentes, a un nivel de significación del 5%. Sin embargo, el intervalo es bastante amplio se sugeriría aumentar el nivel de significación o bien preferiblemente el tamaño de la muestra.

a.1) El contraste de hipótesis para determinar si los promedios de %Cl encontrado para ambas porciones difieren significativamente, es el siguiente:

$$H_0: \mu_{fondo} = \mu_{superficie}$$

$$H_1: \mu_{fondo} > \mu_{superficie}$$

El estadístico de contraste y el criterio de aceptación según la tabla 3.8 para diferencia de medias cuando las varianzas son desconocidas e iguales, son los siguientes:

$$t_o = \frac{(\bar{x}_1 - \bar{x}_2) - D_o}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}; \quad \text{Criterio para la prueba de una cola } t > t_{\alpha/2, v}$$

s_p se calcula como sigue:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$v = n_1 + n_2 - 2$$

Efectuando cálculos se tiene:

$$t_o = \frac{(26.355 - 26.30) - 0}{0.0508 \sqrt{\frac{1}{4} + \frac{1}{3}}} = 1.418$$

$$s_p^2 = \frac{(4-1)0.00123 + (3-1)0.0046}{4+3-2} = 0.002578$$

$$s_p = \sqrt{s_p^2} = 0.0508$$

$$v = 4 + 3 - 2 = 5$$

$$t_{\alpha/2, v} = t_{0.025, 5} = 2.571$$

Dado que $t_o = 1.418$ es menor que $t_{0.025, 5} = 2.571$, a un nivel de significación del 5%, no existe evidencia como para afirmar que el % de CI encontrado en ambas porciones de la muestra difiere significativamente o bien, que no hay evidencia significativa al 5% de falta de homogeneidad en el material.

Usando la función de Excel PRUEBA.T (matriz1, matriz2, colas, tipo)

(PRUEBA.T(superficie, fondo, 2, 2)) para el cálculo del p-valor de la prueba t, se tiene:

SUPERFICIE	FONDO
26.32	26.28
26.33	26.25
26.38	26.38
26.39	

P-VALOR, Prueba F= 0.30485271

P-VALOR, Prueba T= 0.24145269

Dado que $0.2414 > \alpha = 0.05$, no existe evidencia como para rechazar H_0 , a un nivel de significación de 0.05.

El intervalo de confianza para la diferencia de medias viene dado por:

$$\begin{aligned} \text{I.C.} &= (\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2, \nu} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \\ \text{I.C.} &= (26.355 - 26.30) \pm (2.571)(0.0508) \sqrt{\frac{1}{4} + \frac{1}{3}} = 0.055 \pm 0.099 \\ \text{I.C.} &= [-0.45, 0.154] \end{aligned}$$

Si se usa el criterio del intervalo de confianza para contrastar la hipótesis de la diferencia de medias puede concluirse que: "dado que el intervalo creado contempla el valor de cero no existe evidencia como para rechazar que los promedios de %Cl obtenidos para ambas porciones son iguales, a un nivel de significación del 5%".

b) Dado que en este caso se conoce el valor de la desviación estándar de la población ($\sigma = 0.03\%$), el estadístico de contraste viene dado por (ver tabla 3.7):

$$Z_o = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{26.355 - 26.30}{0.03 \sqrt{\frac{1}{4} + \frac{1}{3}}} = 2.40$$

De la tabla de la distribución normal estándar se obtiene (ver anexo A1) $Z_{\alpha/2} = Z_{0.025} = 1.96$. Como puede observarse $|Z_o| \geq Z_{0.025}$, por lo que se concluye que existe evidencia significativa al 5% de que el material no es homogéneo si $\sigma = 0.03\%$.

Comparando los resultados de los literales a) y b) puede decirse que para obtener resultados más concluyentes, si no se conoce la desviación estándar poblacional, es necesario aumentar el tamaño de cada muestra o bien por lo menos usar igual número de réplicas en cada una. Lo anterior también depende de cuan crítica sea la falta de homogeneidad del material para el análisis químico, así como de los recursos disponibles.

Si se emplearan bs datos anteriores como base para el recalcu del tamaño de muestra, para detectar con alta probabilidad una diferencia mínima (Δ) de 0.1% de Cl y una desviación estándar estimada según el ensayo del literal a.1) de 0.0508%, se tendría que $d = \Delta/2\sigma = 0.05 = 0.098$, por lo tanto, el tamaño de muestra requerido para un $\beta = 0.1$ viene dado según la figura b del anexo A3, como $n^* \cong 13 = 2n - 1$, es decir, se necesitarían $n_1 = n_2 = 7$ ensayos para cada porción del material.

4. En una serie de experimentos para la determinación de estaño en productos alimenticios, las muestras se llevaron al punto de ebullición con HCl a reflujo durante diferentes tiempos. Los resultados fueron los siguientes:

Tiempo de ebullición (min)	Estaño encontrado (mg/kg)
30	57, 57, 55, 56, 56, 57, 56, 55
75	51, 60, 48, 32, 46, 54, 56, 51

Estudiar si: a) la variabilidad de los resultados es más grande para los tiempos largos de ebullición; b) Las medias difieren significativamente.

Solución:

La media y la desviación típica para los dos tiempos de reflujo son:

Tiempo de ebullición (min)	Promedio y varianza del Estaño encontrado (mg/kg)
30	$\bar{X} = 56.125$ $s = 0.8345$
75	$\bar{X} = 49.75$ $s = 8.43$

a) Las hipótesis para probar que la variabilidad de los resultados es más grande a medida que aumenta el tiempo de ebullición, vienen dada como:

$$H_0: \sigma_{30\text{min}}^2 = \sigma_{75\text{min}}^2$$

$$H_1: \sigma_{30\text{min}}^2 > \sigma_{75\text{min}}^2$$

Las hipótesis anteriores sugieren un contraste unilateral de la igualdad de varianzas, según la tabla 3.11, el estadístico de contraste y el criterio de rechazo para esta prueba son los siguientes:

Estadístico: $F_o = \frac{s_1^2}{s_2^2}$ en donde s_1^2 es la mayor de las varianzas.

Criterio de rechazo: $F_o > F_{\alpha, n_1-1, n_2-1}$

La razón de varianzas para el problema es:

$$F_o = \frac{s_1^2}{s_2^2} = \frac{(8.43)^2}{(0.8345)^2} = 102.05$$

$$y F_{\alpha, n_1-1, n_2-1} = F_{0.05, 7, 7} = 3.787$$

Dado que $F_o = 102.05 > F_{0.05, 7, 7} = 3.787$, existe evidencia como para afirmar que la variabilidad de los resultados es más grande a tiempos más largos de ebullición con un nivel de significación del 5%. En otras palabras, hay evidencia significativa al 5% como para rechazar la hipótesis nula de igualdad de varianzas. Usando la función DISTR.F de Excel, para el cálculo del p-valor del F calculado, se tiene:

$$\text{DISTR.T}(102.05, 7, 7) = 1.64521\text{E-}06$$

Dado que $1.645 \times 10^{-6} < \alpha = 0.05$, existe evidencia significativa como para rechazar H_o , a un nivel de significación de 0.05.

A continuación se presenta el análisis del contraste a partir de la creación de un intervalo de confianza.

El intervalo de confianza para la razón de varianzas para un nivel de confianza $(1-\alpha)\%$, viene dado por (ver tabla 3.3):

$$\frac{s_1^2/s_2^2}{F_{\alpha/2, n_1-1, n_2-1}} < \sigma_1^2/\sigma_2^2 < \frac{s_1^2/s_2^2}{F_{1-\alpha/2, n_1-1, n_2-1}}$$

De esta forma el intervalo de confianza al 95% ($\alpha = 0.05$) para la razón de varianzas de las muestras a 30 y 75 minutos de tiempo de reflujo es:

$$\frac{102.05}{4.99} < \sigma_1^2/\sigma_2^2 < \frac{102.05}{0.20}$$

$$20.45 < \sigma_1^2/\sigma_2^2 < 510.25$$

Como puede observarse el intervalo de confianza para la razón de varianzas no contempla el valor de uno y se desplaza a la derecha respecto a este valor, por lo que puede concluirse que existe suficiente evidencia como para afirmar que la varianza al tiempo de reflujo de 75 minutos es mayor que la varianza al tiempo de reflujo de 30 minutos, a un nivel de significación del 5%.

b) El contraste de hipótesis para determinar si los promedios de estaño encontrado para ambos tiempos difieren significativamente, es el siguiente:

$$H_0: \mu_{30\text{min}} = \mu_{75\text{min}}$$

$$H_1: \mu_{30\text{min}} > \mu_{75\text{min}}$$

El estadístico de contraste y el criterio de aceptación según la tabla 3.8 para diferencia de medias cuando las varianzas son desconocidas y diferentes, son los siguientes:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}; \quad \text{Criterio para la prueba de una cola } t > t_{\alpha/2, v}$$

Dado que las varianzas son diferentes (literal a), los grados de libertad para el cálculo de $t_{\alpha, v}$, vienen dados por:

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 + 1} + \frac{(s_2^2/n_2)^2}{n_2 + 1}} - 2$$

Efectuando cálculos se tiene:

$$t = \frac{(56.125 - 49.75) - 0}{\sqrt{\frac{71.06}{8} + \frac{0.696}{8}}} = 2.13; \quad v = \frac{\left(\frac{71.06}{8} + \frac{0.696}{8}\right)^2}{\frac{(71.06/8)^2}{9} + \frac{(0.696/8)^2}{9}} - 2 = 7.2 \sim 7$$

$$t_{\alpha/2, v} = t_{0.025, 7} = 2.365$$

Dado que $t = 2.13 > t_{0.05, 7} = 2.365$, a un nivel de significación del 5%, no existe evidencia como para afirmar que las cantidades de estaño extraídas a los tiempos de reflujo de 30 y 75 minutos difieren significativamente.

Usando la función DISTR.T para el cálculo del p-valor del t calculado, se tiene:

$$\text{DISTR.T}(2.13, 7, 2) = 0.0706702$$

Dado que $0.07067 > \alpha = 0.05$, no existe evidencia como para rechazar H_0 , a un nivel de significación de 0.05.

El intervalo de confianza para la diferencia de medias viene dado por:

$$\text{I.C.} = (\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2, v} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\text{I.C.} = (56.125 - 49.75) \pm 2.365 \sqrt{\frac{71.06}{8} + \frac{0.696}{8}} = 6.37 \pm 7.09$$

$$\text{I.C.} = [-0.72, 13.46]$$

Si se usa el criterio del intervalo de confianza para contrastar la hipótesis de la diferencia de medias puede concluirse que: "dado que el intervalo creado contempla el valor de cero no existe evidencia como para rechazar que los promedios de estaño obtenidos para ambos tiempos de reflujos son iguales, a un nivel de significación del 5%".

5. Se han enviado 12 muestras de agua a dos laboratorios para un estudio de determinación de contenido de Flúor en aguas provenientes de doce pozos ubicados en distintos puntos del país. Los resultados obtenidos por cada laboratorio en $\mu\text{g/l}$ se presentan en la tabla a continuación:

Muestra	Laboratorio A	Laboratorio B
1	0.13	0.08
2	0.14	0.08
3	0.2	0.09
4	0.07	0.06
5	0.11	0.13
6	0.15	0.1
7	0.07	0.07
8	0.09	0.14
9	0.08	0.09
10	0.65	0.83
11	0.76	0.83
12	0.77	0.83

¿Los dos laboratorios proporcionaron valores para las concentraciones medias de fluor que difieren significativamente?

Solución:

Se quiere comparar si hay diferencia en el trabajo de los dos laboratorios de análisis, sin embargo, una fuente significativa de variación puede ser el hecho de que las muestras sean de distintas procedencias. Para concluir sobre el trabajo de los dos laboratorios independientemente de la procedencia de las muestras se usará un análisis de datos por pares, prueba bilateral (Ver sección 3.3.6.3, tabla 3.12).

El primer paso en la solución de este problema es sacar las diferencias entre los dos datos de cada muestra y luego continuar con el análisis de dichas diferencias.

HIPOTESIS DE TRABAJO:

$H_0: \mu_d = 0$ (que significa que no existe diferencia entre los dos métodos)

$H_1: \mu_d \neq 0$ (que significa que sí existe diferencia entre los métodos)

OBTENCION DE LAS DIFERENCIAS

Muestra	Laboratorio A	Laboratorio B	$d_i = LA - LB$
1	0.13	0.08	0.05
2	0.14	0.08	0.06
3	0.2	0.09	0.11
4	0.07	0.06	0.01
5	0.11	0.13	-0.02
6	0.15	0.1	0.05
7	0.07	0.07	0
8	0.09	0.14	-0.05
9	0.08	0.09	-0.01
10	0.65	0.83	-0.18
11	0.76	0.83	-0.07
12	0.77	0.83	-0.06

OBTENCIÓN DE ESTADISTICOS

$$\bar{D} = \frac{\sum_{i=1}^n d_i}{n} = \frac{0.05 + 0.06 + \dots + (-0.06)}{12} = -0.00917$$

$$s_d^2 = \frac{\sum_{i=1}^n (d_i - \bar{D})^2}{n-1} = \frac{(0.05 + 0.0092)^2 + (0.06 + 0.0092)^2 + \dots + (-0.06 + 0.0092)^2}{11}$$

$$s_d^2 = 0.00579$$

$$s_d = 0.0761$$

$$t_o = \bar{D} / s_d \sqrt{n} = (-0.0092) / (0.0761) \sqrt{12} = -0.0349$$

CRITERIO DE DECISION

$t_{0.05/2, 12-1} < |t_o|$ debe rechazarse H_0 . Para el caso, $t_{0.025, 11} = 2.201 > |t_o| = 0.0349$ por lo tanto no existe evidencia significativa al 5% como para decir que los dos laboratorios arrojan diferentes resultados.

Usando la función de Excel PRUEBA.T (matriz1, matriz2, colas, tipo) (PRUEBA.T(Laboratorio A, Laboratorio B, 2, 1)) para el cálculo del p-valor de la prueba t, se tiene:

$$P\text{-VALOR} = 0.684481475$$

Dado que el Pvalor es de $0.6844 > \alpha = 0.05$, no existe evidencia como para rechazar H_0 , a un nivel de significación de 0.05. En el archivo *Fluor.xls* se encuentra el procedimiento de cálculo en la hoja Excel.



6. Un laboratorio ha desarrollado un método nuevo para la determinación de Indio en Zinc y sus aleaciones, usando ICPAES. El método se verificó usando patrones de referencia a diferentes concentraciones, las mediciones hechas con el método y la concentración reportada del patrón se presentan en la tabla a continuación.

Patrón	Observación	Certificado
CRM322	2	3
CRM323	6	5
CRM324	15	16
CRM325	44	46
CRM352	2	3
CRM353	1	3
CRM354	7	10
CRM357	6	3
CRM358	12	7
CRM359	16	16
CRM360	29	30

¿Puede decirse que el método permite obtener resultados confiables?.

Solución:

Se quiere comparar si hay diferencia entre la medición hecha con el método desarrollado en el laboratorio y la concentración certificada del patrón empleado. Sin embargo, una fuente significativa de variación puede ser el hecho de que los patrones sean de diferentes concentraciones. Para concluir sobre la fiabilidad del método desarrollado independientemente de la concentración de los patrones de referencia empleados se usará un análisis de datos por pares a partir de la creación de un intervalo de confianza para las diferencias.

El primer paso en la solución de este problema es sacar las diferencias entre los dos datos de cada muestra y luego continuar con el análisis de dichas diferencias.

HIPOTESIS DE TRABAJO:

$H_0: \mu_d = 0$ (que significa que no existe diferencia entre los dos métodos)

$H_1: \mu_d \neq 0$ (que significa que sí existe diferencia entre los métodos)

Este problema puede resolverse usando el procedimiento establecido en el problema 5. Sin embargo, en este caso se analizará mediante la creación de un intervalo de confianza para la diferencia media. Según se especifica en la tabla 3.2, el intervalo de confianza para un promedio viene dado por:

$$I.C._{(1-\alpha)} = \bar{x} \pm \frac{t_{\alpha/2, n-1} s}{\sqrt{n}}$$

En la tabla se presenta la obtención de las diferencias y a continuación el cálculo del promedio y desviación estándar para dichas diferencias:

Patrón	Observación	Certificado	$d_i = \text{obs} - \text{Cert.}$
CRM322	2	3	-1
CRM323	6	5	1
CRM324	15	16	-1
CRM325	44	46	-2
CRM352	2	3	-1
CRM353	1	3	-2
CRM354	7	10	-3
CRM357	6	3	3
CRM358	12	7	5
CRM359	16	16	0
CRM360	29	30	-1

$$\bar{D} = \frac{\sum_{i=1}^n d_i}{n} = \frac{-1+1-1-2+\dots+0-1}{11} = -0.18$$

$$s_d^2 = \frac{\sum_{i=1}^n (d_i - \bar{D})^2}{n-1} = \frac{(-1+0.18)^2 + (1+0.18)^2 + \dots + (-1+0.18)^2}{10}$$

$$s_d^2 = 5.57$$

$$s_d = 2.36$$

El intervalo de confianza al 95%, para el promedio de las diferencias es por tanto:

$$I.C._{(95\%)} = -0.18 \pm \frac{(2.201) * (2.36)}{\sqrt{11}}$$

$$I.C._{(95\%)} = -0.18 \pm 1.57 \text{ ó } [-1.75, 1.39]$$

Dado que el intervalo de confianza contempla el valor de cero puede concluirse que no existe evidencia significativa al 5% como para dudar de la confiabilidad del método desarrollado en el laboratorio.

El intervalo de confianza creado mediante el uso de la hoja Excel puede calcularse mediante la opción herramientas con las sub-opciones análisis de datos y luego estadística descriptiva o bien mediante la función INTERVALO.CONFIANZA (alfa, desv. Est., n). Para este análisis se presenta en el archivo Zn.xls, que puede accersarse mediante el siguiente icono:



3.5.3 PRUEBA DE RECHAZO DE DATOS

1. Los resultados siguientes fueron obtenidos de un estudio de calculo del límite de detección (MDL):

10.2, 9.5, 10.1, 10.3, 9.8, 9.9, 11.9, 10

El analista sospecha que 11.9 es un dato anómalo. ¿Tiene razón?

Solución:

Prueba de Dixon

1. El rango de los datos existentes: $RA = 11.9 - 9.5 = 2.4$
2. La distancia entre el valor anómalo (11.9) y el valor más cercano a él existente (10.2), $DI = \text{abs}(11.9 - 10.3) = 1.6$
3. El cociente QU viene dado por $QU = 1.6/2.4 = 0.67$
4. El valor tabulado de Q para $n = 8$ según la tabla del anexo A3 es: 0.526

Dado que $Q_{\text{calculado}} (= 0.67) > Q_{\text{crítico}} (= 0.526)$ puede concluirse que el dato 11.9 es un dato anómalo con un nivel de confianza del 5%.

Prueba de Huber:

1. La mediana es: 10.05
2. Las desviaciones absolutas son por tanto: 0.15, 0.55, 0.05, 0.25, 0.25, 0.15, 1.85, 0.05
3. La mediana de las desviaciones absolutas es: MAD = 0.20
4. El intervalo para un 95% de confianza es: MEDIANA \pm kMAD= 10.05 \pm (3.5*0.20) = 10.05 \pm 0.20

Dado que 11.9 cae fuera del intervalo [9.85,10.25] este puede considerarse un dato anómalo con un nivel de significación del 0.05.

2. Analice mediante la prueba de GRUBBS los siguientes resultados de concentración en ppm de nitrito de un agua de río:

0.403, 0.410, 0.401, 0.380, 0.400, 0.413, 0.411

Solución:

Analizando la serie de datos se observa que el punto de 0.38 ppm, parece no corresponder adecuadamente a la serie de datos. Esta aparente no correspondencia se analizará mediante la prueba de GRUBB.

1. La desviación estándar (s) de los datos existentes es: 0.011ppm
2. La distancia DM es: $DM = |0.4026 - 0.38| = 0.0226$
3. El cociente es por tanto: $G = DM/s = 0.0226/0.011 = 2.05$
4. El estadístico de comparación viene dado por

$$\text{Estadístico} = \frac{(n-1)}{\sqrt{n}} \sqrt{\frac{t_{(\alpha/(2n), n-2)}^2}{n-2 + t_{(\alpha/(2n), n-2)}^2}}$$

Para la muestra: $t_{(\alpha/(2n), n-2)}^2 = t_{(0.05/(2*7), 7-2)}^2 = t_{(0.00357), 5}^2 = (4.299)^2$, que se ha calculado con la función DISTR.T.INV (probabilidad, grados de libertad) de la hoja Excel. Por tanto, el estadístico tiene el valor de:

$$\text{Estadístico} = \frac{(7-1)}{\sqrt{7}} \sqrt{\frac{18.48}{7-2 + 18.48}} = 2.011$$

Dado que $G(=2.05) > Estadístico (=2.011)$ puede concluirse con un nivel de significación del 5% que el dato de 0.38 es un dato anómalo.

3. Los datos de una muestra control de un experimento de HPLC son los siguientes:

2.79, 2.63, 2.64, 2.71, 2.85, 2.70

Se ha obtenido una nueva observación con un valor de 2.87. Determine, con un nivel de significación de 0.05, si el nuevo resultado pertenece a la distribución prescrita por las observaciones anteriores.

Solución:

La determinación de la consistencia del nuevo resultado se hará empleando una prueba t . El primer paso consiste en la determinación de los estadísticos de la muestra y el segundo en la aplicación del criterio de decisión establecido por la prueba al nuevo resultado.

1. El promedio (\bar{x}) y la desviación estándar (s) de los datos son (que pueden calcularse usando la función estadística de cualquier calculadora científica):

$$\bar{X} = 2.72 \quad ; \quad s = 0.086$$

2. La distancia entre el nuevo valor ($X_a = 2.87$) y el promedio ($\bar{x} = 2.72$) es:

$$D_p = \text{abs}(X_a - \bar{x}) = \text{abs}(2.87 - 2.72) = 0.15$$

3. Por lo tanto el cociente $t = D_p/s$, es:

$$t_{\text{calculado}} = 0.15/0.086 = 1.74$$

4. Por otra parte el valor de t tabulado (ver tabla del anexo C3) es $t_{(0.05/2, 5)} = 1.82$. Si se compara el valor del $t_{\text{calculado}}$ con el valor de t_{tabulado} se observa que el valor calculado de t pertenece al intervalo $-1.82 < t < 1.82$, por lo que puede concluirse que la nueva medición está bajo control estadístico.

REFERENCIAS BIBLIOGRÁFICAS

- **American Public Health Association (1996)**. Standard methods for the examination of water and wastewater. 19th Edition. United States of America: American Public Health Association.
- **Bode, P. (1996)**. Curso sobre validación e incertidumbre de métodos analíticos. El Salvador: Universidad de El Salvador.
- **Freund, J.E. y Walpole, R.E. (1990)**. Estadística matemática con aplicaciones. Cuarta edición. México: Prentice Hall Hispanoamericana.
- **Grubbs, F.E. (1979)**. Procedures for detecting outlying observations. In Army Statistics Manual DARCOM-P706-103, Chapter 3. U:S: Army Research and Development Center, Aberdeen Proving Ground, MD 21005.
- **Juran, J.M. y Gryma, F.M. (1993)**. Manual de control de calidad. Tomo II. México: McGRAW-HILL.
- **Kume, H.S. (1997)**. Herramientas estadísticas básicas para el mejoramiento de la calidad. Colombia: Grupo editorial NORMA.
- **Mendenhall, W. (1990)**. Estadística para administradores. Segunda edición. México: Grupo Editorial Iberoamérica.
- **Miller, J. C. y Miller, J. N. (1993)**. Estadística para química analítica. Segunda edición. Wilmington, Delaware, E.U.A: Addison-Wesley Iberoamericana, S.A.
- **Miller, I. y Freud, J. E. (1996)**. Probabilidad y estadística para ingenieros. Tercera edición. México: Prentice Hall hispanoamericana S.A.
- **Montgomery, D.C. (1991)**. Diseño y análisis de experimentos. México: Grupo Editorial Iberoamérica, S.A. de C.V.
- **Ostle, B. (1977)**. Estadística Aplicada. México: Editora López.
- **Pérez, C. (1999)**. Control estadístico de calidad. Mexico: Alfaomega grupo editor S.A de C.V.
- **Pérez, C. (2002)**. Estadística aplicada a través de Excel. España: Pearson Education.

- **Skoog, D.A., West, D.M., Holler, F.J. y Crouch, S.R. (2001).** Química analítica. Séptima edición. México: McGRAW-HILL.
- **Vitoriano, B. (1997).** Cursos de cálculo de probabilidades I y estadística. Maestría en Estadística. El Salvador: Universidad complutense de Madrid.

ANEXO A3: CURVAS DE POTENCIA PARA CONTRASTES DE HI POTESIS

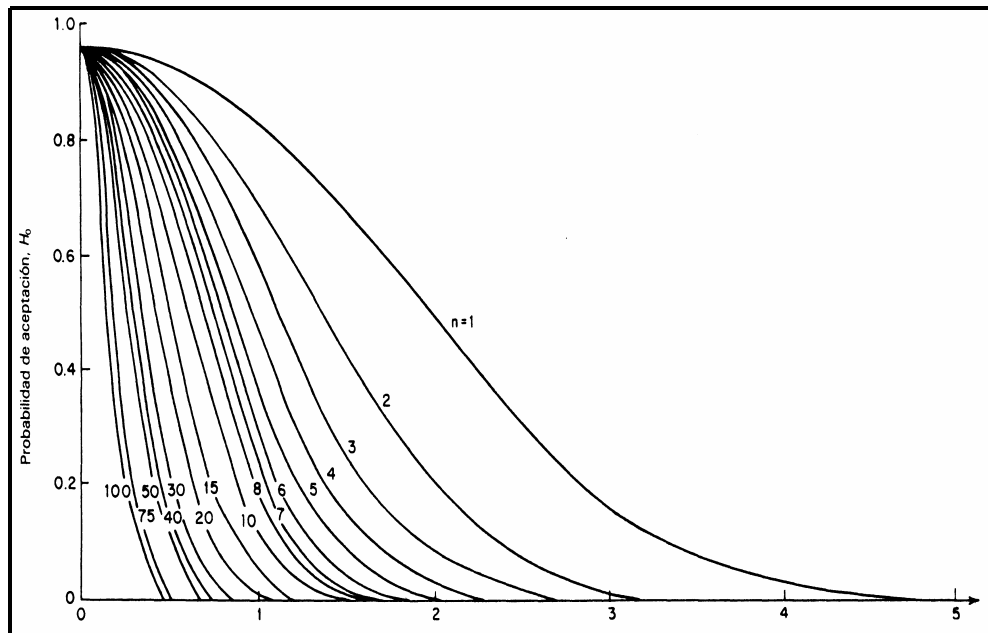


Figura a. Curva característica para contraste bilateral de comparación con un valor especificado ($\alpha = 0.05$) (JURAN y GRZYMA, 1993)

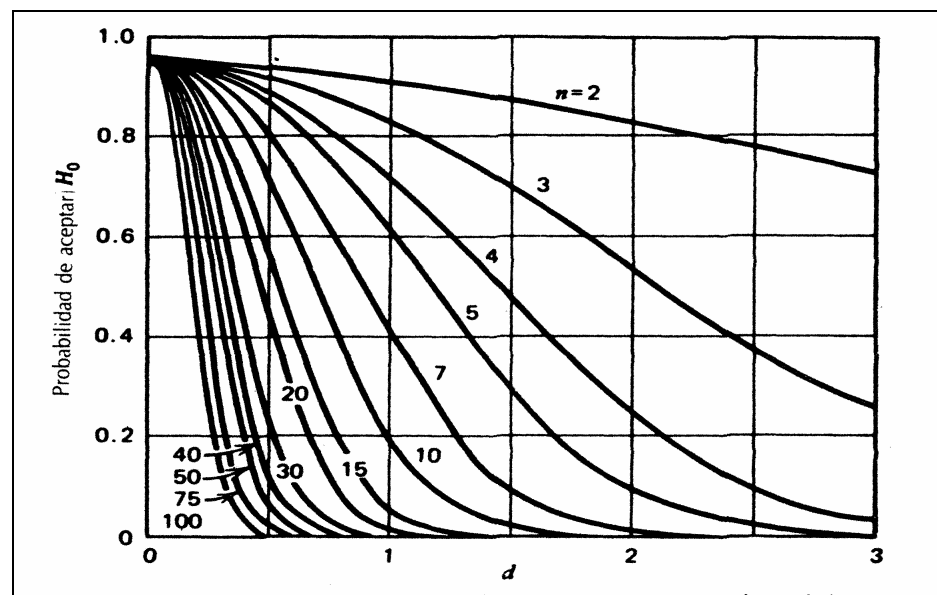


Figura b. Curva característica para contraste bilateral de comparación de dos promedios ($\alpha = 0.05$) (MONTGOMERY, 1990)

ANEXO B3: VALORES CRITICOS DE LA PRUEBA Q (SKOOG et al., 2001, pag.165)

n	90% de confianza	95% de confianza	99% de confianza
3	0.941	0.970	0.994
4	0.765	0.829	0.926
5	0.642	0.710	0.821
6	0.560	0.625	0.740
7	0.507	0.568	0.680
8	0.468	0.526	0.634
9	0.437	0.493	0.598
10	0.412	0.466	0.568

ANEXO C3: VALORES CRITICOS DE LA PRUEBA t
 (AMERICAN PUBLIC HEALTH ASOCIATION, 1996, cap. 1)

n	t crítico al 95% de confianza	t crítico al 99% de confianza
3	1.15	1.15
4	1.46	1.49
5	1.67	1.75
6	1.82	1.94
7	1.94	2.10
8	2.03	2.22
9	2.11	2.32
10	2.18	2.41
12	2.29	2.55
14	2.37	2.66
15	2.41	2.71
16	2.44	2.75
18	2.50	2.82
20	2.56	2.88
30	2.74	3.10
50	2.96	3.34
60	3.03	3.41
100	3.21	3.60
120	3.27	3.66

ANEXO D3: VALORES CRITICOS DE LA PRUEBA DE GRUBBS

(GRUBBS, 1979, cap. 3)

<i>n</i>	<u>Significance Level, •</u>		<i>n</i>	<u>Significance Level, •</u>	
	0.01	0.05		0.01	0.05
			31	3.119	2.759
			32	3.135	2.773
3	1.155	1.153	33	3.150	2.786
4	1.492	1.463	34	3.164	2.799
5	1.749	1.672	35	3.178	2.811
6	1.944	1.822	36	3.191	2.823
7	2.097	1.938	37	3.204	2.835
8	2.221	2.032	38	3.216	2.846
9	2.323	2.110	39	3.228	2.857
10	2.410	2.176	40	3.240	2.866
			41	3.251	2.877
11	2.485	2.234	42	3.261	2.887
12	2.550	2.285	43	3.271	2.896
13	2.607	2.331	44	3.282	2.905
14	2.659	2.371	45	3.292	2.914
15	2.705	2.409	46	3.302	2.923
16	2.747	2.443	47	3.310	2.931
17	2.785	2.475	48	3.319	2.940
18	2.821	2.504	49	3.329	2.948
19	2.854	2.532	50	3.336	2.956
20	2.884	2.557			
21	2.912	2.580			
22	2.939	2.603			
23	2.963	2.624			
24	2.987	2.644			
25	3.009	2.663			
26	3.029	2.681			
27	3.049	2.698			
28	3.068	2.714			
29	3.085	2.730			
30	3.103	2.745			

CAPITULO IV

DISEÑO DE EXPERIMENTOS APLICADO AL ANALISIS QUIMICO

CONTENIDO

	Pág.
4.1 Introducción al diseño de experimentos	4-2
4.2 Experimentos unifactoriales completamente al azar	4-8
4.3 Diseños experimentales en bloques completos al azar	4-34
4.4 Diseños cuadrados latinos	4-48
4.5 Diseños factoriales	4-56
4.6 Diseños factoriales 2^k	4-74
4.7 Problemas de diseño de experimentos	4-90
REFERENCIAS BIBLIOGRAFICAS	4-114
ANEXO A4. Tabla de números aleatorios	4-115
ANEXO B4. Tabla para la prueba de Dunett	4-116
ANEXO C4. Tabla para la prueba de Duncan	4-117
ANEXO D4. Curvas características de operación para el análisis de varianza	4-118
ANEXO E4. Tabla para la prueba de homogeneidad de varianzas F-máx	4-126

4.1 INTRODUCCIÓN AL DISEÑO DE EXPERIMENTOS

Es indudable que en el mundo competitivo en el que se mueve actualmente cualquier organización, uno de los factores claves del éxito es el ritmo en el que la organización es capaz de aprender. Los químicos analistas que se desarrollan en el ámbito de la metrología química, no están exentos de esto, y es muy importante que hagan esfuerzos para aprender y entender mejor los procesos de cuantificación química que ellos practican. Uno de los recursos que más puede contribuir a cumplir con ese objetivo es la *experimentación*. La *experimentación* no consiste más que en reproducir artificialmente la aparición de fuerzas "extrañas" en presencia de personas especialmente preparadas para interpretar y extraer conclusiones de lo que ocurra.

Sin embargo, para que un experimento se realice en la forma más eficiente, es necesario planificarlo, sistemática y objetivamente, puesto que éste se concibe como un medio para verificar o rechazar las hipótesis planteadas. La aplicación de la metodología estadística constituye la única forma objetiva de analizar un problema que involucre datos sujetos a errores experimentales. Es decir, se requiere de un enfoque estadístico de la planificación o del diseño de los experimentos para obtener conclusiones significativas a partir de los datos. El *diseño estadístico de experimentos*, es por tanto, el proceso de planear un experimento para obtener datos apropiados, que pueden ser analizados mediante métodos estadísticos, con el objeto de producir conclusiones válidas y objetivas (MONTGOMERY, 1991, cap. 1).

Los métodos de diseño experimental, tienen una aplicación importante en las actividades de diseño y desarrollo del laboratorio mediante la creación de nuevos procesos de medición o la mejora de los ya existentes. Algunas de las aplicaciones del diseño estadístico de experimentos en el diseño y desarrollo en metrología química son:

- a. Evaluar y comparar las metodologías de medición posibles.
- b. Evaluar materiales y procedimientos alternativos.
- c. Desarrollar procesos de medición que sean robustos.

La aplicación del *diseño estadístico de experimentos* en estas áreas puede dar como resultados una mejora del rendimiento de los procesos de medición, menor variabilidad y

mayor apego a los requerimientos de los procesos establecidos, y menor tiempo de desarrollo y en general menores costos globales.

El uso inteligente de las técnicas estadísticas en la experimentación requiere que el investigador tenga en mente los siguientes puntos (MONTGOMERY, 1991, cap.1):

1. Uso del conocimiento no estadístico del problema.
2. Mantener el diseño y análisis tan simple como sea posible.
3. Reconocer la diferencia entre la significación práctica y la estadística.
4. Reconocer que usualmente los experimentos son iterativos.

Para usar el enfoque estadístico al diseñar y analizar un experimento se requiere que todos los participantes tengan una idea clara de que es exactamente lo que se va a estudiar, como se van a recopilar los datos y, al menos, una idea cualitativa de cómo se van a analizar. A continuación se presenta una guía del procedimiento recomendado:

1. PLANTEAMIENTO DEL PROBLEMA

En la práctica no es sencillo diseñar un planteamiento claro y aceptable de un problema. Es necesario desarrollar todas las ideas sobre el objetivo del experimento. A continuación se dan a conocer las directrices para su planificación (KENETT y ZACKS, 2000, cap. 12):

- i) Investigar todas las fuentes disponibles de información.
- ii) Clasificar los problemas dados dentro del contexto de un problema conocido (marco teórico).
- iii) Contar con los medios económicos y humanos para alcanzar los objetivos de la investigación.
- iv) Identificar las respuestas que hay que medir, los factores que pueden ser variados, los que han de permanecer constantes y los que no pueden ser controlados.
- v) Identificar los intervalos o las limitaciones de las mediciones y los factores experimentales.

Las tres herramientas más importantes del aspecto estadístico de la planificación experimental son: *la creación de bloques, la aleatorización y la replicación* (MONTGOMERY, 1991, cap.1).

2. SELECCIÓN DE FACTORES Y NIVELES

En el momento de planear un experimento es indispensable definir claramente los diferentes elementos a incluir: Unidades experimentales, factores, niveles, etc.

Los objetos sobre los cuales se hacen las mediciones se denominan unidades experimentales.

La mayoría de los experimentos implican el estudio del efecto de una o más variables independientes en una respuesta. Las variables independientes se denominan factores. Los factores pueden ser cualitativos o cuantitativos.

Al grado de intensidad de un factor se le llama nivel. Cada tratamiento implica una combinación de uno o más niveles de un factor (o factores). En la selección de los tratamientos es importante definir claramente cada uno de ellos y entender el papel que jugarán para alcanzar los objetivos del experimento. La elección del número de niveles así como su espaciado es importante para lograr respuestas apropiadas a las preguntas planteadas.

Algunos experimentos incluyen un " Testigo " (mezcla inerte) que es utilizado en la mayoría de los casos como punto de referencia para medir el grado de intensidad de los demás tratamientos. El testigo debe ser una parte integral del experimento, de tal manera que los resultados para el testigo sean directamente comparables con los de otros tratamientos (COCHRAN, 1965, cap. 1).

3. ELECCIÓN DE VARIABLES RESPUESTA

Al seleccionar la respuesta o variable dependiente el experimentador debe estar seguro que la respuesta que se va a medir realmente provea información útil acerca del proceso de estudio. Con mayor frecuencia, el promedio o la desviación estándar (o ambos) de la característica u observación será la variable respuesta. La capacidad de medición también es un factor importante.

4. ELECCIÓN DEL DISEÑO EXPERIMENTAL Y MODELO ESTADÍSTICO

Los planes de experimentación deben guardar cierta relación con los objetivos, necesidades y limitaciones físicas de los experimentos. Asimismo, deben ofrecer ciertas ventajas en la economía de la experimentación y proporcionar estimaciones directas de los efectos experimentales y estimaciones válidas de la varianza.

Algunos de los aspectos que deben considerarse para la elección de un diseño experimental y modelo estadístico son los siguientes (LAWSON et al., 1992, cap. 2):

Tener metas bien definidas: Es responsabilidad del investigador definir claramente sus metas antes de proceder con el experimento y escoger el diseño que va a utilizar. Así mismo, debe indicar que información se obtendrá, así como la magnitud en la diferencia de medias que debe considerarse (con respecto al error experimental σ).

Evitar la confusión de los efectos: Los efectos de cada factor involucrado en el experimento no deben estar confundidos con otras variables. Este problema se puede evitar usando el diseño apropiado.

Procurar resultados insesgados: Los resultados no deben tener ningún sesgo (provocado de forma consciente o inconsciente). Esto se puede regular revisando cuidadosamente el lugar en que se planea ejecutar el experimento, así como el procedimiento que se intenta emplear. Adicionalmente, existen algunas técnicas estadísticas útiles para enfrentar este tipo de situaciones: a) el uso de bloques para considerar variables latentes, b) la aleatorización dentro de cada bloque que minimiza la confusión de los factores con las variables latentes y c) el uso de repeticiones para volver más efectiva la aleatorización.

Estimar la precisión de los resultados: El diseño debe permitir estimar la precisión de los resultados, de tal forma que pueda determinarse si los efectos de los factores son reales o producidos por el error aleatorio en los datos. El uso de repeticiones permite estimar la precisión y su validez es apoyada por el empleo de la aleatorización.

Establecer la precisión del diseño: La precisión del diseño experimental debe ser la suficiente como para satisfacer los objetivos del mismo.

En general, existen varias formas de clasificar a los diseños de experimentación, por ejemplo las siguientes (JURAN y GRAYNA, 1993, cap. 26):

1. Según el número de factores experimentales a investigar (por ejemplo, planes unifactoriales y multifactoriales).
2. Según la estructura del plan de experimentación (por ejemplo, planificado por bloques, factorial, anidado, o por el método de superficie respuesta).

3. Según la clase de información que se pretende que el experimento proporcione (por ejemplo, estimación de efectos, estimación de componentes de varianza, planificación empírica).

En el cuadro 4.1, se presenta una clasificación para los distintos planes de experimentación.

CUADRO 4.1 RESUMEN DE LA CLASIFICACION DE LOS PLANES DE EXPERIMENTACION

OBJETIVO BASICO	ESTRUCTURADO SOLO CON FACTORES CONTROLABLES	ESTRUCTURADO CON FACTORES CONTROLABLES Y NO CONTROLABLES
ANALISIS DE UNIFORMIDAD DEL EFECTO DE UN FACTOR	Diseño unifactorial de efectos fijos	Diseño por bloques al azar completos e incompletos Diseño cuadrado latino Diseño cuadrado greco-latino Diseño cuadrado de Youden
ANALISIS DEL EFECTO DE DOS O MAS FACTORES SOBRE UNA RESPUESTA	Diseños factoriales p^k Diseños factoriales 2^k Diseños mixtos	Diseños factoriales p^k con restricciones en la aleatorización (con bloques) Diseños factoriales 2^k (con bloques) Diseño en parcelas divididas o subdivididas
ANALISIS DEL EFECTO DE MUCHOS FACTORES SOBRE UNA RESPUESTA (CRIBADO DE VARIABLES)	Diseños factoriales fraccionarios 2^{k-p} Diseño Plackett-Burman	Diseños factoriales fraccionarios 2^{k-p} (con bloques)
OBTENCION DE TRAZADOS EMPIRICOS (DIAGRAMAS DE CONTORNO) ILUSTRATIVOS DE QUE FACTORES BAJO CONTROL INFLUYEN SOBRE LOS RESULTADOS	Diseños de superficie respuesta	
ESTUDIO DE LA VARIABILIDAD RELATIVA SOBRE UNA MISMA MUESTRA Y ENTRE MUESTRAS	Diseños unifactoriales de efectos aleatorios Diseños anidados	

5. RECOLECCION DE INFORMACION

Al llevar a cabo el experimento se deben tomar en cuenta las siguientes recomendaciones (COCHRAN, 1965, cap.2):

- a) Asegurar uniformidad en la aplicación de los tratamientos.
- b) Ejercitar suficiente control sobre las influencias externas, de tal manera que cada tratamiento produzca su efecto, bajo condiciones deseables.
- c) Proyectar medidas convenientes y no sesgadas de los efectos de los tratamientos.
- d) Prevenir errores importantes, de los cuales parece que ningún tipo de experimentos está enteramente libre. La supervisión y comprobación adecuada del trabajo de los ayudantes y un examen de los datos de cada unidad experimental por parte del experimentador, permitirá tener un mayor alcance en el descubrimiento y rectificación de errores.

Es aconsejable que el experimentador realice una prueba piloto para estimar parámetros importantes para el cálculo del número adecuado de réplicas, así como también para controlar el error producido por imprevistos en el experimento a escala completa.

La prueba piloto es la ejecución del experimento tomando una pequeña muestra arbitraria con el fin de poner a prueba los instrumentos de medición, verificar el manejo de las operaciones de campo y conocer el comportamiento de las unidades experimentales ante un tratamiento determinado, etc. Los resultados de una prueba piloto usualmente sugieren algunas modificaciones antes de realizar el experimento en su totalidad.

6. ANALISIS DE DATOS Y VALIDACION DE LA IDONEIDAD DEL MODELO

Hay dos aspectos en cualquier problema experimental: el diseño del experimento y el análisis estadístico de los datos. Estos dos temas están estrechamente relacionados, ya que el método de análisis depende directamente del diseño empleado (MONTGOMERY, 1991, cap.1).

El primer paso en el análisis del experimento consiste en determinar la ecuación para cada observación, con los componentes siguientes (COCHRAN, 1965, cap. 3):

- i) Un promedio general alrededor del cual fluctúan las observaciones.
- ii) Los componentes que representan el efecto de los tratamientos, bloques y sus interacciones.

- iii) Un componente residual que representa a las fuentes que influyeron en la observación y que se conoce como error experimental.

El segundo paso consiste, en el análisis propiamente dicho de los datos. El método más aplicable para el análisis de datos es el ANALISIS DE VARIANZA, que consiste en analizar la variación en una respuesta y asignar proporciones (componentes) de esa variación a cada una de las variables del conjunto de variables independientes controlables o no (COCHRAN, 1965, cap. 3).

Teóricamente es posible dividir la variabilidad del resultado de un experimento en dos partes:

1. La variabilidad originada por los factores que influyen directamente en el resultado del experimento, estudiados en sus distintos niveles o tratamientos.
2. La variabilidad producida por el resto de los factores con influencia en el resultado del experimento desconocida o no controlable, que se conoce como error experimental.

Existen otros métodos estadísticos que permiten el análisis de datos y validación de la idoneidad del modelo que pueden ser de tipo gráfico, como el análisis de residuos, elaboración de contornos, análisis de Pareto de efectos, etc. o bien de tipo analítico, como el estimación de efectos. Sin embargo, hay que recordar que los métodos estadísticos no pueden probar experimentalmente que un factor o varios factores tienen un efecto particular. Solo proporcionan directrices para la veracidad y validez de los resultados, mediante el establecimiento de la probabilidad de error de una conclusión o la asignación de un nivel de confiabilidad a los resultados. La ventaja principal de la aplicación de métodos estadísticos es que agregan objetividad al proceso de toma de decisiones. Las técnicas estadísticas, aunadas a un buen conocimiento técnico o del proceso y al sentido común, suelen llevar a conclusiones razonables.

4.2 EXPERIMENTOS UNIFACTORIALES COMPLETAMENTE AL AZAR

El diseño unifactorial completamente al azar o randomizado, es el más simple de todos los diseños. En éste, los tratamientos se distribuyen al azar sobre todas las unidades

experimentales y el número de réplicas o unidades por tratamiento pueden ser iguales o diferentes (COCHRAN, 1965, cap. 4).

Este diseño se utilizará cuando se estudian dos o más tratamientos, bajo las siguientes condiciones (COCHRAN, 1965, cap. 4):

- i) Cuando el material experimental es homogéneo, es decir tiene las mismas características deseadas para la investigación.
- ii) Cuando es probable que una parte apreciable de las unidades experimentales se destruya o no responda. En estos casos el análisis de los datos en otros diseños se vuelve difícil o imposible.
- iii) En experimentos pequeños en donde la mayor precisión de otros diseños no compensa la pérdida de grados de libertad del error (a través de otros diseños las unidades experimentales podrían darnos resultados más precisos; pero si esta precisión no sobrepase lo perdido con la disminución de los grados de libertad, no valdría la pena utilizar otro diseño).
- iv) Cuando el experimentador no tiene los criterios o no tiene los elementos necesarios para determinar la heterogeneidad en el material experimental.

Dentro de las ventajas del empleo de diseños unifactoriales se encuentran las siguientes (COCHRAN, 1965, cap. 4):

- i) Permite flexibilidad completa. Puede usarse cualquier número de tratamientos y repeticiones. Puede variarse a voluntad el número de repeticiones de un tratamiento a otro (pero no es recomendable sin una buena razón).
- ii) El análisis estadístico es fácil, aún si el número de repeticiones no es el mismo para todos los tratamientos, pues los cálculos para su análisis no presentan ninguna variante significativa.
- iii) Aún cuando los datos de alguna de las unidades se hayan perdido o se rechacen por alguna causa, el método de análisis sigue siendo sencillo. Por otra parte la pérdida relativa de información debida a los datos faltantes es de menos importancia que en cualquier otro diseño.
- iv) Permite el máximo número de grados de libertad para la suma de cuadrados del error.

Dentro de las limitaciones del uso de un diseño unifactorial se encuentran:

- i) Las unidades experimentales se suponen homogéneas. En dicho caso los errores debido a la heterogeneidad son asignados al error aleatorio.
- ii) No tiene mayor grado de precisión, hay otros diseños más eficientes en la estimación del error estándar por unidad experimental.

En general, el uso de diseños unifactoriales es recomendable en aquellos experimentos en donde el material de ensayo es suficientemente homogéneo. Es decir, cuando todas las unidades experimentales reúnen prácticamente las mismas características, de modo que el efecto de un tratamiento sobre la variable bajo estudio, es el mismo, independientemente de la unidad experimental donde se mide, salvo por variaciones aleatorias debido a las fuentes del error en la investigación. Tal condición de homogeneidad en el material experimental, permite que cada tratamiento pueda ensayarse con el número de repeticiones que se desee.

4.2.1 DISEÑOS UNIFACTORIALES Y DISTRIBUCION ALEATORIA DE LOS TRATAMIENTOS

Una vez que se ha definido claramente el problema y se han determinado los diferentes elementos presentes en la investigación: población, marco muestral, unidades experimentales, factores, variables a medir, tratamientos, etc., entonces se debe calcular el tamaño de la muestra y generarla a través de un procedimiento que asegure la aleatorización. Los tratamientos se alojan completamente al azar sobre las unidades experimentales, bajo la condición de que cada unidad deberá tener una probabilidad conocida de recibir un tratamiento en particular. Para la distribución, primeramente se enumeran las unidades experimentales y luego se obtiene una permutación aleatoria. Los procesos más comúnmente empleados en la práctica para la obtención de permutaciones aleatorias, son el uso de una tabla de números aleatorios como la que se presenta en la tabla anexo A4, o bien, el uso de programas de computadora. En este caso el investigador debe establecer un criterio para elegir las unidades experimentales y tratamientos; por ejemplo, tomar el primer dígito para el tratamiento y los dos últimos para la unidad experimental.

4.2.2 ANALISIS DEL DISEÑO UNIFACTORIAL

Supóngase que se desea comparar k tratamientos o niveles de un factor único. La respuesta que se desea observar en cada uno de los k tratamientos es una variable aleatoria. Los datos

aparecerían como en la tabla 4.2. En general, habrá n_i observaciones del tratamiento i (MONTGOMERY, 1991, cap. 3).

La generalización del modelo estadístico lineal para cada observación viene dada por (MONTGOMERY, 1991, cap. 3):

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad (\text{ec. 4.1})$$

(observación = media global + efecto del factor + error aleatorio)

donde: $i = 1, 2, \dots, k$; $j = 1, 2, \dots, n_i$; y_{ij} es la (ij) -ésima observación, μ es un parámetro común a todos los tratamientos denominado media global; τ_i es un parámetro único que manifiesta el efecto del i -ésimo tratamiento, y ε_{ij} es la componente aleatoria del error para cada observación.

TABLA 4.2 PLANTILLA DE RECOLECCION DE DATOS PARA UN EXPERIMENTO UNIFACTORIAL

Tratamiento	Observaciones	Totales (Y_i)	Media \bar{y}_i
(1)	$y_{11} \dots \dots \dots y_{1n_1}$	Y_1	\bar{y}_1
(2)	$y_{21} \dots \dots \dots y_{2n_2}$	Y_2	\bar{y}_2
.	.	.	.
.	.	.	.
(k)	$y_{k1} \dots \dots \dots y_{kn_k}$	Y_k	\bar{y}_k
		Y	\bar{Y}

El objetivo será probar hipótesis apropiadas con respecto a los efectos del tratamiento y hacer una estimación de ellos. Para probar las hipótesis, se supone que los errores del modelo son variables aleatorias independientes con distribución normal, con media cero y varianza σ^2 . Se supone que esta última es constante para todos los niveles del factor (MONTGOMERY, 1991, cap. 3).

Además se requiere que el experimento se realice en orden aleatorio, de manera que el medio ambiente en el que se usan los tratamientos sea lo mas uniforme posible.

El modelo estadístico, describe dos situaciones con respecto al efecto de los tratamientos. Primero, los tratamientos podrían haber sido seleccionados específicamente por el experimentador. En esta situación se desea probar hipótesis sobre las medias de los tratamientos y las conclusiones se aplican sólo a los niveles del factor considerados en el análisis. Las conclusiones no pueden hacerse extensivas a tratamientos similares que no

hayan sido consideradas específicamente. También será deseable estimar los parámetros del modelo (μ, τ_i, σ^2) . Este modelo se denomina *modelo de efectos fijos* (MONTGOMERY, 1991, cap. 3).

Alternativamente, los k tratamientos pueden ser una muestra aleatoria de una población mayor de tratamientos. En esta situación sería deseable generalizar las conclusiones (basadas en la muestra de tratamientos) a todos los tratamientos de la población, ya sea que hayan sido explícitamente considerados en el análisis o no. En este caso, los τ_i son variables aleatorias y resulta inútil conocer sus valores particulares para los tratamientos investigados. En su lugar se prueban hipótesis con referencia a la variabilidad de los τ_i y se interpreta dicha variabilidad. Este modelo se conoce como *modelo de efectos aleatorios o de componentes de varianza* (MONTGOMERY, 1991, cap. 3).

4.2.2.1 ANÁLISIS DE VARIANZA PARA EL MODELO DE EFECTOS FIJOS

Como se había mencionado en la sección anterior, en el modelo de efectos fijos los tratamientos son seleccionados específicamente por el experimentador (considerándose un número contable de ellos), por lo que interesa probar hipótesis relativas a la igualdad de los efectos de los tratamientos. Algunos ejemplos de aplicación de este tipo de diseño en análisis químico son los siguientes:

- Determinar si el tipo de envase en el que es envasada una solución influye sobre la concentración media de un analito contenido en la misma, después de un tiempo fijado de almacenaje.
- Comparar el trabajo de los analistas en un laboratorio.
- Comparar los resultados obtenidos al aplicar diferentes procesos de medición.
- Etc.

En el modelo unifactorial de efectos fijos, la media del i -ésimo tratamiento viene dada como $\mu_i = \mu + \tau_i$. Como puede observarse los efectos se definen usualmente como desviaciones con respecto a la media global μ , por esta razón $\sum \tau_i = 0$ (MONTGOMERY, 1991, cap. 3).

La hipótesis nula que se deseará probar es que las medias de todos los tratamientos sean iguales, es decir que (MONTGOMERY, 1991, cap. 3):

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

O que

$$H_0: \tau_i = 0 \text{ para } i = 1, 2, \dots, k$$

En forma correspondiente, la hipótesis alternativa es que no todas las medias de los tratamientos son iguales, o sea que (MONTGOMERY, 1991, cap. 3):

$$H_1: \mu_i \neq \mu_j \text{ para al menos un par } (i, j)$$

O que

$$H_1: \tau_i \neq 0 \text{ cuando menos para un valor de } i$$

Por tanto, es posible hablar de probar la igualdad de medias de los tratamientos, o bien de probar que los efectos de los tratamientos son cero. El procedimiento apropiado para probar la igualdad en el nivel medio de k tratamientos es el *análisis de varianza en un sentido*.

La prueba de hipótesis consiste en evaluar el estadístico F_o que se define como la razón (MONTGOMERY, 1991, cap. 3):

$$F_o = \frac{SC_{\text{Tratamientos}} / (k - 1)}{SC_{\text{Error}} / (N - k)} = \frac{CM_{\text{Tratamientos}}}{CM_{\text{Error}}}$$

Si la hipótesis nula es verdadera, F_o tiene una distribución F con $k-1$ y $N - k$ grados de libertad. Debe rechazarse la hipótesis, si el valor del estadístico es demasiado grande. Esto implica una región crítica unilateral superior. En otras palabras, se rechaza H_0 si:

$$F_o > F_{\alpha, k-1, N-k}$$

La denominación de análisis de varianza resulta de descomponer la variabilidad total de los datos en sus partes componentes, que en general, puede medirse por *la suma total de cuadrados corregida*, que viene dada por (MONTGOMERY, 1991, cap. 3):

$$SC_{\text{Total}} = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{Y})^2 \quad (\text{ec. 4.2})$$

en donde $\bar{Y} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}}{N}$ con $N = \sum_{i=1}^k n_i$ para tamaños muestrales diferentes en cada tratamiento, o bien, $N = kn$ cuando el número de observaciones es el mismo para todos los tratamientos.

Se puede demostrar que (MONTGOMERY, 1991, cap. 3):

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{Y})^2 = \sum_{i=1}^k n_i (\bar{y}_i - \bar{Y})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \quad (\text{ec. 4.3})$$

(Donde \bar{y}_i es el promedio obtenido para cada tratamiento)

Es decir, que la suma total de los cuadrados de las desviaciones de cada observación con respecto a la media global ($\mathbf{SC}_{\text{Total}}$), se descompone en la suma de los cuadrados de las desviaciones de las medias de los tratamientos en relación con la media global ($\mathbf{SC}_{\text{Tratamientos}}$), que representa a la variabilidad total debida a la diferencia entre los tratamientos, y la suma de los cuadrados de las desviaciones de las observaciones con respecto a sus propias medias de tratamiento ($\mathbf{SC}_{\text{Error}}$), que representa la variabilidad debida al error aleatorio. Por tanto puede escribirse (MONTGOMERY, 1991, cap. 3):

$$\mathbf{SC}_{\text{Total}} = \mathbf{SC}_{\text{Tratamientos}} + \mathbf{SC}_{\text{Error}} \quad (\text{ec. 4.4})$$

Entre más grande sea $\mathbf{SC}_{\text{Error}}$, mayor será la variación en las observaciones que pueden atribuirse al error aleatorio. Entre más grande es el valor de $\mathbf{SC}_{\text{Tratamientos}}$, mayor es la diferencia que existe entre las medias de los tratamientos y la media global.

Bajo la hipótesis nula y la suposición de que $\varepsilon_{ij} \sim \mathbf{N}(0, \sigma^2)$, $\mathbf{SC}_{\text{Tratamientos}} / \sigma^2$ y $\mathbf{SC}_{\text{Error}} / \sigma^2$ son dos variables aleatorias independientes con una distribución chi-cuadrado. Los grados de libertad se obtienen al separar la suma total de cuadrados. La $\mathbf{SC}_{\text{Total}}$ tiene $N - 1$ grados de libertad, debido a que se pierde un grado de libertad al ser necesario que (MONTGOMERY, 1991, cap. 3):

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{Y}) = 0$$

La suma de los cuadrados de los tratamientos tiene $k-1$ grados de libertad, debido a que se impone la restricción (MONTGOMERY, 1991, cap. 3):

$$\sum_{i=1}^k (\bar{y}_i - \bar{Y}) = 0$$

Entonces, como $SC_{Total} = SC_{Error} + SC_{Tratamientos}$, el número de grados de libertad para SC_{Error} será igual a la diferencia entre el número de grados de libertad SC_{Total} y $SC_{Tratamientos}$, es decir $N - k$.

De esta forma pueden definirse las cantidades (MONTGOMERY, 1991, cap. 3):

- Media de cuadrados debida a los tratamientos:

$$MC_{Tratamientos} = \frac{SC_{Tratamientos}}{k-1} \quad (\text{ec. 4.5})$$

El cálculo de la suma de cuadrados de los tratamientos se facilita si se aplica la siguiente ecuación:

$$SC_{Tratamientos} = \sum_{i=1}^k \frac{y_i^2}{n_i} - \frac{Y^2}{N} \quad (\text{ec. 4.6})$$

Donde y_i es la suma total para cada tratamiento e Y es la suma total de todos los datos.

- Media de cuadrados debida al error, que viene dada como:

$$MC_{Error} = \frac{SC_{Error}}{N-k} \quad (\text{ec. 4.7})$$

La suma de cuadrados del error puede calcularse a partir de:

$$SC_{Error} = SC_{Total} - SC_{Tratamientos}$$

Donde la suma de cuadrados total puede calcularse como:

$$SC_{Total} = \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}^2 - \frac{Y^2}{N} \quad (\text{ec. 4.8})$$

El procedimiento de prueba de hipótesis se resume en la tabla 4.3. Las fuentes de variación, grados de libertad, suma de cuadrados, cuadrados medios y el cociente F_{α} juntos, constituyen lo que se conoce como tabla de análisis de varianza (ANDEVA o ANOVA).

TABLA 4.3 ANÁLISIS DE VARIANZA PARA EL DISEÑO COMPLETAMENTE ALEATORIO

Fuente de Variación	g.l.	Suma de Cuadrados (SC)	Cuadrados medios (CM)	Estadístico (F_p)
Tratamiento	k-1	$SC_{Tratamientos}$	$SC_{Tratamientos} / k-1$	
Error	N-k	SC_{Error}	$SC_{Error} / N-k$	$F_o = \frac{MC_{Tratamientos}}{MC_{Error}}$
Total	N-1	SC_{Total}		

El MC_{Error} es un estimador insesgado de σ^2 sin importar si la hipótesis nula es cierta. Por otro lado, si la hipótesis nula es cierta entonces $E(MC_{Tratamientos}) = \sigma^2$; es decir, si H_o es cierta, tanto MC_{Error} como $MC_{Tratamientos}$ son estimadores insesgados de la varianza del error. Pero si la hipótesis nula no es cierta, $MC_{Tratamientos}$ tiende a ser generalmente mayor que MC_{Error} .

4.2.2.2 ESTIMACION DE PARAMETROS DEL MODELO

Los estimadores mínimos cuadrados de las medias poblacionales son las medias observadas de los grupos en tratamiento \bar{y}_i . Las medias observadas son los promedios de n_i observaciones independientes, de forma que la varianza de la media de un grupo de tratamiento es $\sigma_{\bar{y}_i}^2 = \sigma^2 / n_i$, siendo su estimador (KUEHL, 2001, cap. 3):

$$s_{\bar{y}_i}^2 = \frac{s^2}{n_i}$$

En donde a partir del análisis de varianza, $s^2 = MC_{Error}$.

La construcción de un intervalo al $100(1 - \alpha)\%$ de confianza, para cada media del grupo en tratamiento i , viene dado por (MONTGOMERY, 1991, cap. 3):

$$IC_{(1-\alpha)\%} = \bar{y}_i \pm t_{\alpha/2, N-k} \sqrt{\frac{MC_{Error}}{n_i}} \quad (\text{ec. 4.9})$$

En gran parte, el interés no se centra solamente en la estimación de la media del i -ésimo tratamiento, sino también en la diferencia entre las medias de los tratamientos. Un intervalo de confianza al $100(1 - \alpha)\%$ para la diferencia entre las medias de dos tratamientos cualesquiera, será (MONTGOMERY, 1991, cap. 3):

$$IC_{(1-\alpha)\%} = (\bar{y}_i - \bar{y}_j) \pm t_{\alpha/2, N-k} \sqrt{\frac{2MC_{Error}}{(n_i + n_j)/2}} \quad (\text{ec. 4.10})$$

4.2.2.3 COMPARACION DE TRATAMIENTOS CON UN CONTROL

En muchos estudios, uno de los tratamientos actúa como un control para algunos o todos los restantes. En este caso, el interés se centra en determinar si las respuestas medias de los tratamientos difieren de la del tratamiento control (por ejemplo, comparar el trabajo de una serie de laboratorios con un laboratorio de referencia). Dunnett (1955) introdujo un procedimiento basado en una tasa de error con respecto al experimento con este propósito (MONTGOMERY, 1991, cap 3).

Las hipótesis a contrastar son: $H_0: \mu_i = \mu_{control}$ contra $H_1: \mu_i \neq \mu_{control}$. El estadístico de contraste para la prueba, viene dado por:

$$|\bar{y}_i - \bar{y}_{control}|$$

El criterio de Dunnett para comparar k tratamientos con el control es (MONTGOMERY, 1991, cap 3):

$$D(k, \alpha_E) = d_{\alpha, k, v} \sqrt{s^2 \left(\frac{1}{n_i} + \frac{1}{n_{control}} \right)} = d_{\alpha, k, v} \sqrt{MC_{Error} \left(\frac{1}{n_i} + \frac{1}{n_{control}} \right)}$$

La hipótesis nula se rechaza, a un nivel de significación α si: $|\bar{y}_i - \bar{y}_{control}| > D(k, \alpha_E)$

Los valores de $d_{\alpha, k, v}$ para el método de Dunnett de uno o dos lados, se encuentran en la tabla del anexo B4, para k tratamientos, un error tipo I de α con respecto al experimento y v grados de libertad para la estimación de la varianza del error experimental.

Las estimaciones de los intervalos de confianza simultáneos bilaterales para la diferencia entre las medias de los tratamientos individuales y la media del control ($\mu_i - \mu_{control}$) son (KUEHL, 2001, cap.3):

$$\bar{y}_i - \bar{y}_{control} \pm D(k, \alpha)$$

4.2.2.4 METODO DE COMPARACION DE MEDIAS INDIVIDUALES

Al rechazar la hipótesis nula puede concluirse que las medias o el efecto de los tratamientos difieren significativamente; pero esto no es información suficiente para rechazar una alternativa en particular, por ejemplo, sea la hipótesis nula $H_0: \mu_1 = \mu_2 = \mu_3$ rechazada, en este caso se pudieran haber dado alguna de las siguientes situaciones $\mu_1 = \mu_2$, $\mu_1 \neq \mu_3$, $\mu_1 = \mu_3$, $\mu_1 \neq \mu_2$, $\mu_1 \neq \mu_2 \neq \mu_3$, por lo tanto, esta es una buena razón para que el investigador

necesite un análisis más completo para encontrar las diferencias estadísticamente significativas entre cierto número de medias de tratamientos en la población.

Hay una diversidad de métodos para realizar análisis pareados de las medias o para realizar contrastes entre dichas medias, según sea la necesidad del investigador, entre ellas se encuentran *la prueba de Duncan, método de la diferencia mínima significativa (ó LSD), la prueba de Newman-Keuls, la prueba de Tuckey.*

A continuación se describe *la prueba de intervalos múltiples de Duncan* para contrastes pareados de medias.

La prueba de intervalos múltiples de Duncan se basa en la noción general del rango estudentizado, que establece que la diferencia entre las medias consideradas debe exceder un cierto rango para ser consideradas significativamente diferentes. En este caso dicho rango recibe el nombre de rango menos significativo y se representa por R_p (en donde $p = 2, 3, \dots, k$). En un primer momento para la prueba de Duncan es necesario colocar en forma ascendente los k promedios de tratamientos, y se determina el error estándar de cada promedio, a partir de la expresión (MONTGOMERY, 1991, cap. 3):

$$S_{\bar{y}_i} = \sqrt{\frac{CM_{Error}}{n}} \quad (\text{ec. 4.11})$$

Para el caso de muestras del mismo tamaño. Cuando el tamaño es diferente hay que reemplazar n por la media armónica n_h de $\{n_i\}$ (que es utilizada en el caso de promediar razones) es decir (MONTGOMERY, 1991, cap. 3):

$$n_h = \frac{k}{\sum_{i=1}^k (1/n_i)}$$

Como puede deducirse el promedio armónico de una serie de muestras de igual tamaño es igual a n ($n_h = n$).

En segundo lugar, deben calcularse los rangos menos significativos a partir de los cuales se rechaza o no la diferencia entre dos medias determinadas, su ecuación es (MONTGOMERY, 1991, cap. 3):

$$R_p = r_{\alpha}(p, f) S_{\bar{y}_i} \quad \text{para } p = 2, 3, \dots, k \quad (\text{ec. 4.12})$$

Como se mencionó anteriormente, los valores de las cantidades R_p son llamados rangos estudentizados menos significativos o mínimos intervalos significativos y dependen del nivel de significación deseado (α) y del número de grados de libertad del cuadrado del error. Esta dependencia es expresada en el término $r_{\alpha}(p, f)$. Los valores de $r_{\alpha}(p, f)$ pueden obtenerse de la tabla del anexo C4, para $p = 2, 3, \dots, 10$ medias (MONTGOMERY, 1991, cap. 3).

A continuación, se prueban las diferencias observadas entre las medias, comenzando por el valor más alto contra el más pequeño, comparando esta diferencia con el rango menos significativo R_k . Después se calcula la diferencia entre el valor más alto y el segundo más pequeño y se compara con R_{k-1} . Este procedimiento continúa hasta que todas las medias han sido comparadas con la media más grande. Posteriormente, se obtiene la diferencia entre la segunda media más grande y la más pequeña y se compara contra el rango menos significativo R_{k-1} . Este proceso continúa hasta que han sido consideradas las diferencias entre todos los $k(k-1)/2$ posibles pares. Si una diferencia observada es mayor que el rango menos significativo correspondiente, se concluye que la pareja de medias en cuestión es significativamente diferente (MONTGOMERY, 1991, cap. 3).

Para evitar contradicciones, ninguna diferencia entre una pareja de medias se considera significativa si las dos medias se encuentran entre otras dos que no difieren significativamente (MONTGOMERY, 1991, cap. 3); es decir si sucede el siguiente caso:

Sean $\bar{y}_1 = \bar{y}_2$ y \bar{y}_3 , parte del conjunto de k medias y $\bar{y}_1 - \bar{y}_2$, $\bar{y}_1 - \bar{y}_3$, $\bar{y}_2 - \bar{y}_3$ diferencias de las $k(k-1)/2$ parejas de medias.

Entonces si:

$$\bar{y}_1 - \bar{y}_2 < R_4$$

$$\bar{y}_1 - \bar{y}_3 < R_3$$

$$\bar{y}_3 - \bar{y}_2 < R_2$$

Se concluirá que $\bar{y}_3 - \bar{y}_2$ tienen una diferencia no significativa a pesar de ser mayor que R_2 , ya que, en caso contrario, se estaría generando una contradicción, puesto que se está considerando no significativa la diferencia de $\bar{y}_1 - \bar{y}_2$ y $\bar{y}_1 - \bar{y}_3$ siendo $R_4 > R_3 > R_2$.

Es importante hacer notar, que a medida que el número de medias incluidas en el grupo aumenta, la prueba requiere una diferencia observada más grande para detectar parejas de medias significativamente diferentes. Lo anterior puede explicarse a partir del siguiente razonamiento:

Los valores de $r_{\alpha}(p, f)$ deben elegirse para obtener un nivel de protección o seguridad específico. En otras palabras cuando se comparan dos medias separadas p pasos, el nivel de protección es $(1-\alpha)^{p-1} = (1-\alpha)(1-\alpha) \dots(1-\alpha)$; hasta $p - 1$ veces, en donde α es el nivel de significación especificado para dos medias adyacentes. Así, el nivel del error de informar al menos una diferencia significativa incorrecta entre dos medias cuando el tamaño del grupo es p , será igual a: $1 - (1-\alpha)^{p-1}$ (MONTGOMERY, 1991, cap. 3).

Por lo general, si el nivel de protección es α , las pruebas sobre las medias tienen un nivel de significación mayor o igual que α . En consecuencia, el procedimiento de Duncan es muy eficiente para determinar diferencias entre medias cuando estas diferencias realmente existen (MONTGOMERY, 1991, cap. 3).

4.2.2.5 ANÁLISIS DE VARIANZA PARA EL MODELO DE EFECTOS ALEATORIOS

Muchas veces al experimentador le interesa analizar un factor que tiene un número infinito de posibles niveles, o bien un número de niveles lo suficientemente grande como para considerarse infinito. El objetivo de la investigación en este caso, debe ser la obtención de conclusiones relativas a toda la población de niveles del factor y para lograrlo el investigador debe seleccionar al azar k de estos niveles y estructurar un diseño unifactorial conocido como *diseño unifactorial de componentes de varianza o de efectos aleatorios*. En este diseño las hipótesis a probar hacen referencia a la variabilidad de los tratamientos y a su interpretación (MONTGOMERY, 1991, cap. 3). Un ejemplo de aplicación del *diseño unifactorial de componentes de varianza*, en análisis químico, es concluir respecto a la homogeneidad de la concentración de un analito en un producto en polvo contenido en un recipiente, ya que se puede concebir que el número de puntos a seleccionar en el interior del contenedor es de carácter infinito.

El modelo estadístico lineal para este diseño, es el mismo que el presentado en la ecuación 4.1, en donde τ_i e ε_{ij} son variables aleatorias. Si τ_i tienen una varianza a σ_{τ}^2 y es independiente de ε_{ij} , la varianza de cualquier observación es: $V(y_j) = \sigma_{\tau}^2 + \sigma^2$. Las varianzas σ_{τ}^2 y σ^2 se conocen como componentes de varianza y para probar las hipótesis del modelo se requiere que los ε_j sean independientes y normalmente distribuidos (NID(0, σ^2)), que los t_i

sean independientes y normalmente distribuidos (NID $(0, \sigma_{\tau}^2)$) y que τ_i e ε_{ij} sean independientes (MONTGOMERY, 1991, cap. 3).

La variabilidad total en las observaciones se descompone en una componente que mide la variación entre los tratamientos. ($SC_{\text{Tratamientos}}$) y otra que mide la variación dentro de los tratamientos (SC_{Error}). La suma de identidad de cuadrados: $SC_{\text{Total}} = SC_{\text{Tratamientos}} + SC_{\text{Error}}$, aun es válida (MONTGOMERY, 1991, cap. 3).

En este caso, carece de sentido probar hipótesis que se refieren a los efectos de tratamientos individuales y en su lugar deben probarse las hipótesis (MONTGOMERY, 1991, cap. 3):

$$H_0: \sigma_{\tau}^2 = \mathbf{0} \text{ (todos los tratamientos son idénticos)}$$

Contra

$$H_1: \sigma_{\tau}^2 > \mathbf{0} \text{ (existe variabilidad entre los tratamientos)}$$

Si la hipótesis nula es verdadera:

$$F_o = \frac{SC_{\text{Tratamientos}} / (k - 1)}{SC_{\text{Error}} / (N - k)} = \frac{CM_{\text{Tratamientos}}}{CM_{\text{Error}}}$$

Tiene una distribución F con $k - 1$ y $N - k$ grados de libertad. Por tanto, H_0 se rechaza si:

$$F_o > F_{\alpha, k-1, N-k}$$

El procedimiento computacional y tabla de análisis de varianza para el modelo de efectos aleatorios es idéntico al de efectos fijos. Sin embargo, las conclusiones son un tanto diferentes, ya que se aplican a toda la población de tratamientos (MONTGOMERY, 1991, cap. 3).

Usualmente interesa calcular las componentes de varianza en el modelo (σ_{τ}^2 y σ^2). El procedimiento usado para estimar σ_{τ}^2 y σ^2 se conoce como método de análisis de varianza (MONTGOMERY, 1991, cap. 3).

El procedimiento consiste en igualar los valores esperados de las medias de cuadrados con su correspondiente valor observando en la tabla de análisis de varianza, para después despejar

las componentes de varianza. Al igualar los valores observados y los esperados de las medias de cuadrados, se obtiene (MONTGOMERY, 1991, cap. 3):

$$CM_{\text{Tratamientos}} = \sigma^2 + n\sigma_{\tau}^2 \quad (\text{ec. 4.13})$$

y

$$CM_{\text{Error}} = \sigma^2$$

Los estimadores de las componentes de varianza son (MONTGOMERY, 1991, cap. 3):

$$\hat{\sigma}^2 = CM_{\text{Error}}$$

y

$$\hat{\sigma}_{\tau}^2 = \frac{CM_{\text{Tratamientos}} - CM_{\text{Error}}}{n}$$

En donde n es el número de observaciones por tratamiento.

Para tamaños de muestras desiguales, n debe reemplazarse por (MONTGOMERY, 1991, cap. 3):

$$n_0 = \frac{1}{k-1} \left[\sum_{i=1}^k n_i - \frac{\sum_{i=1}^k n_i^2}{\sum_{i=1}^k n_i} \right]$$

El método de análisis de varianza para la estimación de las componentes de varianza no requiere la suposición de normalidad.

Se pueden calcular las estimaciones de los intervalos de confianza al 100(1 - α)%, para ambos componentes de varianza. Para σ^2 el intervalo viene dado como (KUEHL, 2001, cap. 5):

$$\frac{SC_{\text{Error}}}{A} < \sigma^2 < \frac{SC_{\text{Error}}}{B} \quad (\text{ec.4.14})$$

donde $A = \chi_{\alpha/2, N-k}^2$ y $B = \chi_{(1-\alpha/2), N-k}^2$

Un intervalo con al menos 100(1 - 2 α)% de confianza para σ_{τ}^2 es (KUEHL, 2001, cap. 5):

$$\frac{SC_{Tratamientos}(1-(F_u/F_o))}{nC} < \sigma_\tau^2 < \frac{SC_{Tratamientos}(1-(F_l/F_o))}{nD} \quad (\text{ec.4.15})$$

En donde $C = \chi_{\alpha/2, k-1}^2$, $D = \chi_{(1-\alpha/2), k-1}^2$, F_o es el estadístico observado. Las cantidades $F_u = F_{\alpha/2, (k-1), (N-k)}$ y $F_l = F_{(1-\alpha/2), (k-1), (N-k)}$ que son valores de la variable F .

Por definición, una componente de varianza es positiva; sin embargo, las estimaciones de σ_τ^2 mediante la ecuación (4.15) pueden ser negativas. Existen varios cursos de acción sugeridos en caso de estimaciones negativas (SEARLE, 1971; SEARLE, CASELLA Y MCCULLOCH, 1992):

1. Aceptar la estimación como evidencia de un valor verdadero de cero y usar el cero como la estimación, reconociendo que el estimador será sesgado.
2. Conservar la estimación negativa, tomando en cuenta que es posible que los cálculos subsecuentes con los resultados no tengan mucho sentido.
3. Interpretar la estimación de la componente negativa como una indicación de un modelo estadístico incorrecto.
4. Utilizar un método diferente de análisis de varianza para estimar sus componentes.
5. Reunir más datos y analizarlos por separado o junto con los existentes y esperar que el aumento de información conduzca a estimaciones positivas.

4.2.3 CALCULO DEL TAMAÑO DE LA MUESTRA¹

Uno de los problemas principales que se enfrentan al llevar a cabo una investigación experimental, es la elección del número de réplicas necesario para detectar alguna diferencia predeterminada entre las medias de los tratamientos. En análisis químico, esta elección muchas veces está condicionada a la disponibilidad de recursos humanos, materiales, y económicos.

No obstante, se considera de gran importancia poder definir el número de réplicas que sería conveniente tomar para un determinado estudio y de esta forma establecer comparaciones y estimaciones objetivas sobre la potencia de los contrastes a la hora de validar los modelos empleados.

¹Para el cálculo del tamaño de la muestra se estará bajo el supuesto que $n = n_1 = n_2 = \dots = n_k$.

Por otra parte, siempre que sea aplicable, toda investigación experimental que involucre un diseño unifactorial debe partir con el cálculo del número de réplicas pues ello conduce a un control más eficiente del error tipo II (β).

El cálculo del número de réplicas en un diseño unifactorial depende en primer lugar, de la aleatoriedad o no de los niveles del factor. Y en segundo lugar, de la sensibilidad de la prueba, que describe la capacidad del procedimiento para detectar diferencias en las medias poblacionales y se mide por la potencia de la prueba $1 - \beta^2$.

A continuación, se describe el cálculo del número de réplicas para los modelos unifactoriales de efectos fijos y efectos aleatorios.

a. Calculo del número de réplicas para el modelo unifactorial de efectos fijos:
(MONTGOMERY, 1991, cap. 3)

Al plantear un diseño completamente aleatorizado, el objetivo principal es probar la hipótesis nula (H_0), es decir: $H_0: \tau_i = 0$ para $i = 1, 2, \dots, k$, mediante el uso del estadístico ($F_o = CM_{\text{Tratamientos}}/CM_{\text{Error}}$). La potencia de la prueba se define en general, como:

$$1 - \beta = P(F_o > F_{\alpha, k-1, N-k} | H_0 \text{ falsa})$$

Cuando H_0 es falsa, F_o tiene una distribución F desplazada con $K - 1$ y $N - k$ grados de libertad y *parámetro de desplazamiento* $\lambda = n \sum_i^k \tau_i^2 / \sigma^2$ (*valor del efecto poblacional* = $\mu_i - \bar{\mu}$). Si la hipótesis nula es cierta, entonces el parámetro de desplazamiento tienen un valor de cero, puesto que todo $\tau_i = 0$ y F_o tiene una distribución F centrada.

Se cuenta con tablas de valores y gráficas de la potencia de la prueba F desplazada en función del nivel de significación α ; la potencia $1 - \beta$; los grados de libertad $K - 1$ y $N - k$, y Φ (una función del parámetro de desplazamiento que es: $\Phi = \sqrt{\lambda/k}$), y que se usan para estimar el número de réplicas necesario para realizar un diseño

² Como se presentó en la sección 3.3.2 del capítulo 3, la potencia de una prueba es la probabilidad de rechazar la hipótesis nula (H_0) siendo la hipótesis alternativa verdadera (decisión correcta). La potencia de una prueba puede calcularse como $1 - \beta$, donde β es la probabilidad de cometer un error tipo II, el cual se comete al aceptar H_0 cuando en realidad es falsa (decisión incorrecta).

unifactorial de efectos fijos. Como puede observarse un valor de Φ requiere de un valor de σ^2 y valores específicos de las medias de los tratamientos que conducen a rechazar la hipótesis nula ($\mu_1, \mu_2, \dots, \mu_k$).

Dado que es difícil especificar los efectos deseados para un conjunto de niveles del factor, es más conveniente especificar la diferencia entre cualquier pareja de medias de los tratamientos ($D = \mu_i - \mu_j$), que puede establecerse como una diferencia significativa dentro del fenómeno de estudio. En este caso se trabaja con el cálculo de un valor mínimo de Φ^2 , que viene dado como:

$$\Phi^2 = \frac{nD^2}{2k\sigma^2}$$

En el anexo D4 se presentan tablas y curvas de potencia para la distribución F desplazada.

b. Calculo del número de réplicas para el modelo unifactorial de efectos aleatorios:
(KUEHL, 2001, cap.5)

En el caso de efectos fijos, el cálculo de potencia requiere de la distribución no central de F. Este no es el caso para el modelo de efectos aleatorios.

La hipótesis nula de interés en el modelo de efectos aleatorios, $H_0: \sigma_\tau^2 = 0$ se prueba con $F_o = CM_{Tratamientos}/CM_{Error}$, y la potencia de la prueba es:

$$1 - \beta = P(F > F_{\alpha, k-1, k(n-1)} | H_0 \text{ falso}) = P(F > F_{\alpha, k-1, k(n-1)} | \sigma_\tau^2 > 0)$$

Cuando $\sigma_\tau^2 > 0$, la distribución F es la distribución central $F_{\alpha, k-1, k(n-1)}$ multiplicada por una constante $1/\lambda^2$, donde:

$$\lambda = \sqrt{1 + n \frac{\sigma_\tau^2}{\sigma^2}}$$

La potencia de la prueba se puede determinar a partir de la distribución F central como sigue:

$$1 - \beta = P[F > 1/\lambda^2 (F_{\alpha, k-1, k(n-1)})]$$

El número de réplicas necesario, se determina a partir de valores tabulados o gráficos de potencia para la prueba F central (similares a los de los efectos fijos), tal como los que se presentan en el anexo D4, dados el número de tratamientos k , un nivel de significación (α), una potencia deseada ($1 - \beta$) y λ .

Para determinar el valor de λ es necesario establecer una razón deseada de los componentes de varianza, que puede determinarse mediante juicio o experiencia anterior, o según la desviación estándar de una observación individual, σ_y .

El valor necesario para la razón σ_τ^2/σ^2 , en términos del porcentaje fijo de aumento en σ_y , aceptable (P), y que con uno mayor se rechaza la hipótesis nula es:

$$\frac{\sigma_y}{\sigma} = \frac{\sqrt{\sigma_\tau^2 + \sigma^2}}{\sigma} = 1 + 0.01P$$

Por lo que el valor necesario para la razón σ_τ^2/σ^2 es:

$$\frac{\sigma_\tau^2}{\sigma^2} = (1 + 0.01P)^2 - 1$$

4.2.4 COMPROBACIÓN DE LA IDONEIDAD DEL MODELO

La validez de las estimaciones y pruebas de hipótesis con respecto al análisis de los datos, se apoyan en el establecimiento de varias suposiciones clave. Las suposiciones que fundamentan el análisis de varianza son:

- a. Que los datos estén descritos de manera adecuada por el modelo lineal:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \dots, k \\ j = 1, 2, \dots, n_i \end{cases}$$

- b. Que los errores (ε_{ij}) sean variables aleatorias independientes y normalmente distribuidas con media cero y varianza constante ($\varepsilon_{ij} \sim N(0, \sigma^2)$).

- c. En el caso de los modelos de efectos aleatorios se hace la suposición adicional de que los τ_i son independientes y están distribuidos normalmente con media cero y varianza σ_τ^2 , y además que σ_τ^2 y ε_{ij} son independientes.

Las condiciones ideales raramente se cumplen en los estudios reales. En el análisis de varianza, la magnitud de las discrepancias en los datos, con respecto a las suposiciones de independencia, la distribución normal supuesta y las varianzas homogéneas, pueden conducir a inferencias erróneas. Las discrepancias menores generalmente, no ocasionan modificaciones sustanciales en la eficiencia de las estimaciones y en los niveles de significación de las pruebas. Sin embargo, es posible que discrepancias mayores, en especial una heterogeneidad excesiva de la varianza o alguna heterogeneidad de la varianza con números de réplicas desiguales, afecten en forma importante las inferencias estadísticas.

Lo anterior implica, la necesidad de implementar el proceso de validación del modelo como un paso indispensable en cada proyecto que se realice.

La validación del modelo consiste en el uso de herramientas de tipo gráfico o analítico que permiten "diagnosticar" en que medida las suposiciones tomadas para el análisis son adecuadas.

Las herramientas principales para el diagnóstico están basadas en el análisis de los *residuos observados*. En general, los residuos se definen como la diferencia entre la respuesta observada y valor de respuesta predicho con el modelo supuesto, es decir:

$$\varepsilon_{ij} = y_{ij} - \hat{y}_{ij}$$

Para el caso del modelo unifactorial los residuos quedan definidos como:

$$\varepsilon_{ij} = y_{ij} - \bar{y}_i$$

En otras palabras, los residuos del *i*-ésimo tratamiento se determinan restando el promedio del tratamiento a cada observación dentro del tratamiento.

La comprobación de la idoneidad del modelo consiste en observar la magnitud de los residuos y su relación con otras variables, proporcionando evaluaciones visuales de las suposiciones del análisis de varianza. La suposición de varianza homogénea se evalúa con una gráfica de los residuos contra las medias estimadas de los tratamientos. La suposición de independencia se evalúa con una gráfica de los residuos contra los valores predichos, y la

suposición de distribución normal mediante una gráfica de probabilidad normal de los residuos.

A continuación se describe el análisis de residuos para la prueba de normalidad, independencia y homogeneidad de varianzas.

a. PRUEBA DE SUPOSICIÓN DE NORMALIDAD:

Un procedimiento útil para comprobar la suposición de normalidad de los residuos es construir una gráfica de probabilidad normal con los residuos, tal como la que se presenta en la figura 4.1.

Una gráfica de probabilidad normal organiza los residuos en orden creciente y los grafica contra los cuantiles correspondientes de la distribución normal estándar. El residuo de orden i -ésimo tiene una frecuencia acumulada de i/N en una muestra de tamaño N , valor al que se le aplica una corrección por continuidad $f_i = (i - 0.5)/N$.

El cuantil de una variable normal se conoce por una probabilidad acumulada dada, y el valor del cuantil de un residuo se conoce por su frecuencia acumulada en el conjunto de datos residuales. Se forman pares con los cuantiles correspondientes de los residuos y la distribución normal estándar y se grafican como valores X e Y en una gráfica bivariada llamada gráfica cuantil-cuantil. Si los cuantiles de los residuos son iguales a los de la variable normal para la misma frecuencia acumulada, la gráfica será una línea recta (ver capítulo 2).

En general, se dice que un análisis de varianza (y otros procedimientos relacionados como los de comparaciones múltiples) es consistente o robusto ante la suposición de normalidad, porque la prueba F se ve muy poco afectada por la falta de normalidad, sobre todo para el caso de efectos fijos. Usualmente las desviaciones de la normalidad hacen que tanto el nivel de significación verdadero como la potencia difieran ligeramente de los valores que se especifican, siendo generalmente la potencia menor. El modelo de efectos aleatorios se ve más afectado por la no normalidad. En particular, los niveles reales de confianza en las estimaciones por intervalo de los componentes de varianza pueden diferir mucho de los valores especificados (MONTGOMERY, 1991, cap. 4).

Un defecto que se presenta a menudo en las gráficas de probabilidad normal, es un residuo que es mucho menor o mayor que el resto (residuo inusitado). La presencia de uno o más residuos inusitados puede distorsionar seriamente el análisis de varianza, por

lo que cuando se detecta uno en potencia debe realizarse una cuidadosa investigación (MONTGOMERY, 1991, cap. 4).

Un procedimiento informal para detectar residuos distanciados es analizar residuos estandarizados (MONTGOMERY, 1991, cap. 4):

$$d_{ij} = \frac{\varepsilon_{ij}}{\sqrt{CM_{Error}}}$$

Si los ε_{ij} son $N(0, \sigma^2)$, los residuos estandarizados deben ser aproximadamente normales con media cero y varianza igual a uno. Por lo tanto, es aplicable la regla empírica general a los residuos estandarizados (ver capítulo 1). Un residuo a una distancia mayor que 3 ó 4 desviaciones estándar del origen es potencialmente un residuo distanciado ó anómalo (MONTGOMERY, 1991, cap. 4).

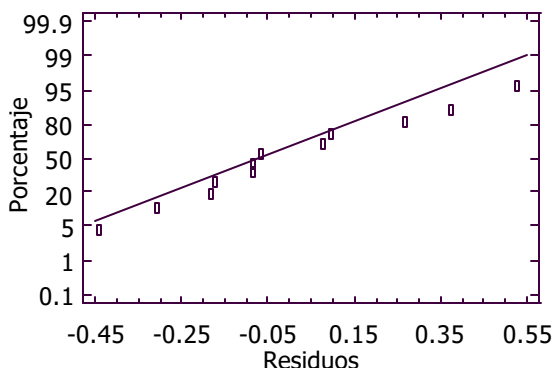


FIGURA 4.1 GRAFICA DE PROBABILIDAD NORMAL PARA RESIDUOS

b. SUPOSICION DE HOMOGENEIDAD DE LAS VARIANZAS:

Si el modelo es correcto y las suposiciones se satisfacen los residuos no deben tener algún patrón, ni deben estar relacionados con alguna otra variable incluyendo la respuesta y_j . Una comprobación sencilla consiste en graficar los residuos contra los valores ajustados \hat{y}_j . En esta gráfica no debe revelarse ningún patrón obvio. En la figura 4.2 se presentan algunos posibles patrones de comportamiento de los residuos cuando se grafican en función de los valores predichos \hat{y}_j . El último caso (gráfica 4.2, c) es el que se presenta cuando la suposición de varianza constante en los grupos de tratamiento se cumple.

Algunas veces la varianza de las observaciones aumenta a medida que la magnitud de los residuos lo hace. Esto resulta cuando el residuo es proporcional a la magnitud de la observación, en este caso la gráfica de comportamiento de los residuos contra los valores predichos (\hat{y}_{ij}) parecerá un embudo que se ensancha. La varianza variable de los residuos también ocurre en casos cuyos datos no tienen una distribución normal y están sesgados (MONTGOMERY, 1991, cap. 4).

Si la suposición de homogeneidad no se cumple, la prueba F es solo ligeramente afectada en los modelos balanceados de efectos fijos. Sin embargo, el problema es serio en el caso desbalanceado y en el de efectos aleatorios (MONTGOMERY, 1991, cap. 4).

El enfoque usual para tratar varianzas variables de los residuos consiste en aplicar una transformación para igualar varianzas y volver aplicar el análisis de varianza a los datos transformados (MONTGOMERY, 1991, cap. 4).

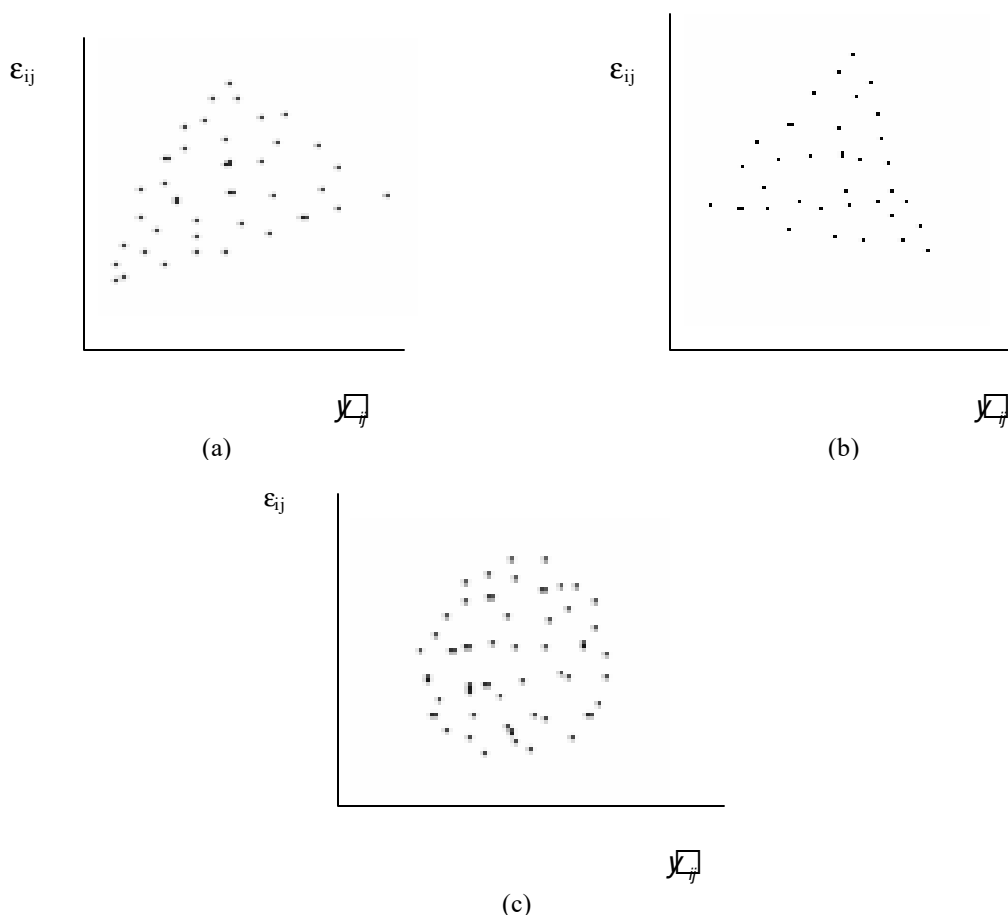


FIGURA 4.2 PATRONES DE COMPORTAMIENTO DE LOS RESIDUOS CONTRA LOS VALORES PREDICHOS \hat{y}_{ij} :
CASO (a), el error no está normalmente distribuido ni la varianza es constante. CASO (b), la distribución de los residuos no presenta varianza constante. CASO (c), representa el comportamiento cuando existe homogeneidad en la varianza de las observaciones (de varianza constante de los residuos).

La gráfica de dispersión (spread-location) s-l, es otra gráfica residual que revela más sobre las varianzas heterogéneas. Las tendencias en la gráfica s-l se pueden usar para revelar las posibles relaciones entre las medias y las varianzas de los grupos en tratamiento (KUEHL, 2001, cap. 4).

Las raíces cuadradas de los valores absolutos de los residuales, $\sqrt{|\epsilon_{ij}|}$ se usan para medir la dispersión de los residuos, ya que su magnitud reflejará la dispersión o variación dentro de un grupo de tratamiento. La raíz cuadrada elimina parte de la asimetría en los residuales, mientras las estimaciones de las medias de grupos de tratamientos miden la localización. En la figura 4.3 se muestra una gráfica s-l (KUEHL, 2001, cap. 4).

Uniendo, mediante líneas rectas, las medianas de las $\sqrt{|\epsilon_{ij}|}$ para cada sitio en la gráfica muestran el incremento en su magnitud con un incremento en las medias de los sitios. Esta tendencia del incremento en la magnitud de los valores absolutos de los residuos, representa el aumento en la varianza del sitio cuando crece la media (KUEHL, 2001, cap. 4).

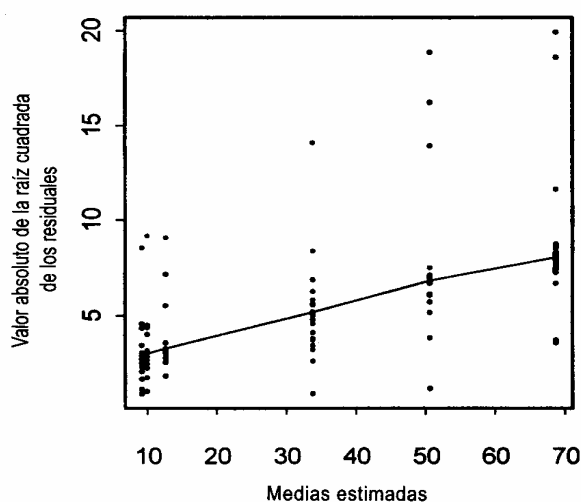


FIGURA 4.3 GRAFICO DE DISPERSION (SPREAD-LOCATION) s-l

Aparte de las gráficas de residuos que frecuentemente se usan para diagnosticar la igualdad en las varianzas, se han propuesto algunas pruebas estadísticas formales de homogeneidad de las varianzas para diseños completamente aleatorizados, entre las cuales se encuentran: la prueba de Bartlett, la prueba de Levene y la prueba F máx.

En general, las hipótesis nula y alternativa para todas las pruebas son:

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_k^2$$

$$H_1: \text{lo anterior no es cierto al menos para una } \sigma_j^2$$

Una técnica ampliamente usada es la prueba de Bartlett. Este procedimiento consiste en calcular un estadístico cuya distribución muestral es, aproximadamente, chi-cuadrado con $k-1$ grados de libertad cuando las k muestras aleatorias provienen de poblaciones normales independientes. La estadística de prueba es (MONTGOMERY, 1991, cap. 4):

$$\chi_o^2 = 2.3026 \frac{q}{c}$$

Donde:

$$q = (N - k) \log_{10} S_p^2 - \sum_{i=1}^k (n_i - 1) \log_{10} S_i^2$$

$$c = 1 + \frac{1}{3(k-1)} \left(\sum_{i=1}^k (n_i - 1)^{-1} - (N - k)^{-1} \right)$$

$$S_p^2 = \frac{\sum_{i=1}^k (n_i - 1) S_i^2}{N - k}$$

y S_i^2 es la varianza muestral de la i -ésima población.

El valor de q es grande cuando hay una gran diferencia entre las varianzas muestrales S_i^2 y es igual a cero si todas las S_i^2 son iguales. Por tanto, debe rechazarse H_0 para valores grandes de χ_o^2 . En otras palabras, se rechaza H_0 , sólo si, $\chi_o^2 > \chi_{\alpha, k-1}^2$, en donde $\chi_{\alpha, k-1}^2$ es el punto α porcentual de la distribución chi-cuadrado con $k-1$ grados de libertad. Varios estudios indican que la prueba de Bartlett es muy sensible a la suposición de normalidad y no debe ser aplicada cuando exista alguna duda en cuanto a esta suposición (MONTGOMERY, 1991, cap. 4).

Una prueba menos sensible a las suposiciones de normalidad, es la prueba de Levene (1960). Sea y_j la j -ésima observación del i -ésimo grupo de tratamiento y \hat{y}_i su mediana. Sea $z_j = |y_j - \hat{y}_i|$ el valor absoluto de la diferencia entre una observación y la mediana del grupo de tratamiento i . Para probar la homogeneidad de varianzas, se calcula el análisis de varianza en un sentido para z_j y a partir del estadístico (KUEHL, 2001, cap. 4):

$$F_o = \frac{CM_{Tratamientos}}{CM_{Error}} = \frac{\sum_{i=1}^k n_i (\bar{z}_i - \bar{z})^2 / (k-1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2 / (N-k)}$$

Se rechaza H_0 si $F_o > F_{\alpha, (k-1), (N-k)}$.

Otra prueba sencilla válida cuando los tamaños de las muestras son iguales y las observaciones tienen una distribución normal, es la prueba F máx (Hartley, 1950). El estadístico de la prueba es el siguiente (KUEHL, 2001, cap. 4):

$$F_o \text{ Máx} = \frac{\text{máx}(s_i^2)}{\text{min}(s_i^2)}$$

Donde $\text{máx}(s_i^2)$ y $\text{min}(s_i^2)$ son las varianzas mayor y menor respectivas dentro de los grupos de tratamiento. La hipótesis nula se rechaza al nivel de significación α si $F_o \text{ Máx}$, es el valor de la variable $F \text{ Máx}$ excedida, con una probabilidad α para k grupos de tratamientos y $(k-1)$ grados de libertad para cada s_i^2 . Los valores críticos del estadístico $F \text{ Máx}$ se encuentran en la tabla del anexo E4.

c. **SUPOSICION DE INDEPENDENCIA DE LOS RESIDUOS:**

Además, también es importante graficar los residuos contra el orden del tiempo en que se recopilaron los datos ya que ello permitirá detectar cuando la suposición de independencia de los residuos se ha violado; pudiendo haber sido originado por las siguientes razones: Los tratamientos no han sido correctamente distribuidos (en forma completamente aleatoria), cambios en los instrumentos de medición conforme el tiempo transcurre, etc. Una tendencia a tener rachas con residuos positivos y negativos indica una correlación positiva de los residuos. Potencialmente este es un problema serio y difícil de corregir, por ello es importante prevenir el problema durante la recolección de los datos. Un paso importante para asegurar la independencia es un procedimiento apropiado de aleatorización (MONTGOMERY, 1991, cap. 4).

La figura 4.4, presenta un ejemplo de gráfico de residuos contra la secuencia en el tiempo.

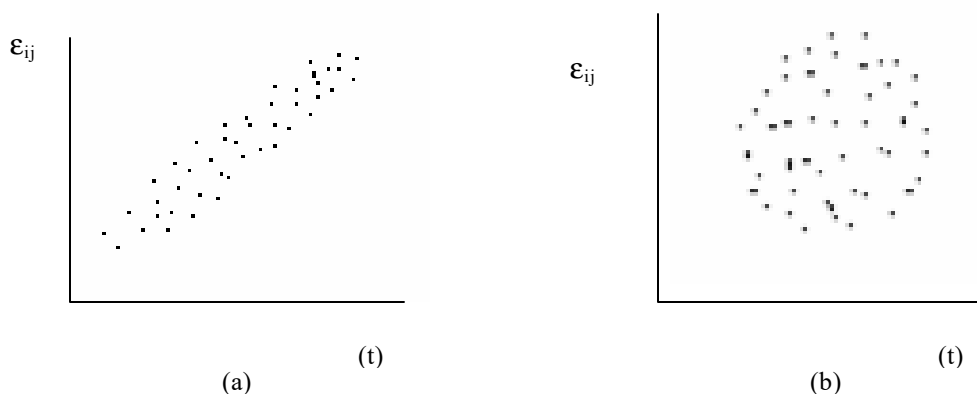


FIGURA 4.4 GRAFICO DE RESIDUOS CONTRA SECUENCIA EN TIEMPO
CASO (a) no se cumple la suposición de independencia de los residuos; CASO (b) si se cumple el Principio de independencia

4.3 DISEÑOS EXPERIMENTALES EN BLOQUES COMPLETOS AL AZAR

Cuando las unidades experimentales no son homogéneas, la variación entre las unidades puede encubrir los verdaderos efectos de los tratamientos sobre una respuesta, evitando tener comparaciones precisas entre los mismos. Factores como soluciones patrón de distinta concentración, muestras de orígenes diferentes, lotes distintos de reactivos o material fabricado, etc., son ejemplos de variables externas al tratamiento que marcan la diferencia entre las unidades experimentales, aumentando por tanto, la variación entre las observaciones de la variable respuesta.

El bloquear estratifica a las unidades experimentales en grupos homogéneos, o unidades parecidas. Una buena elección de los criterios de bloqueo es un medio para reducir y controlar la varianza del error experimental con el fin de lograr una mayor precisión. Algunos criterios de bloqueo son: 1) proximidad (reactivos de un mismo lote); 2) características físicas (soluciones de distintas concentraciones); 3) tiempo; 4) administración de la tarea del experimento (KUEHL, 2001, cap. 8).

El *diseño de bloques al azar* es uno de los primeros válidos para estimar el error experimental y para probar si los efectos de los tratamientos son significativos a pesar de la heterogeneidad de las unidades experimentales sobre las que se hacen las observaciones.

Un bloque es simplemente una porción de material experimental que se espera que sea más homogéneo que el total o por lo menos más homogéneo que el material primitivo no clasificado (KENETT y ZACKS, 2000, cap. 12).

Los bloques completos al azar comprenden el mismo número de unidades experimentales. Cada tratamiento se ensaya el mismo número de veces sobre otras tantas unidades experimentales. El objetivo consiste en mantener la variabilidad entre unidades experimentales dentro de un bloque tan pequeña como sea posible y maximizar la diferencia entre bloques.

El diseño se recomienda en experimentos donde la variación debida a la heterogeneidad en las unidades experimentales es tan grande que se puede reducir formando bloques. En general, las condiciones bajo las cuales se hará uso del diseño de bloques completos al azar son las siguientes (COCHRAN y COX, 1965, cap. 4):

- i) Cuando las unidades experimentales no sean homogéneas; es decir que no puedan reaccionar o responder a los tratamientos de la misma manera o con la misma capacidad debido a sus diferencias y ello no significa necesariamente que no tengan el mismo aspecto.
- ii) Cuando el investigador tenga suficientes elementos que le garanticen la heterogeneidad en las unidades experimentales.
- iii) Cuando se tenga un solo criterio de formación de bloques.
- iv) Donde es poco probable que las unidades experimentales se destruyan o no respondan a un tratamiento determinado.

Entre las ventajas del empleo de este tipo de diseño se encuentran (COCHRAN y COX, 1965, cap. 4):

- 1) En el caso de que se presente heterogeneidad en las unidades experimentales, por medio de la agrupación se reduce el error aleatorio.

- 2) Puede incluirse cualquier número de tratamientos y repeticiones. En el diseño de bloques completos al azar, cada tratamiento tendrá el mismo número de repeticiones.
- 3) Si por un accidente se hace necesario la omisión de un bloque completo o de todos los datos de uno o más tratamientos, el análisis no presenta ninguna dificultad.
- 4) Si la varianza del error experimental es mayor para algunos tratamientos que para otros aún puede obtenerse un error insesgado para probar cualquier combinación específica de las medias de los tratamientos.
- 5) Las medias de las repeticiones dan comparaciones insesgadas de las diferencias entre las repeticiones.

Las limitaciones principales en la aplicación de este tipo de diseño son las siguientes (COCHRAN y COX, 1965, cap. 4):

- 1) Se pierden grados de libertad para la suma de los cuadrados del error.
- 2) Cuando faltan datos de algunas unidades separadas, el análisis requiere de un cálculo adicional.

4.3.1 ALEATORIZACION EN DISEÑOS POR BLOQUES AL AZAR

El muestreo y la distribución de las unidades experimentales en el diseño de bloques está constituido por dos etapas. La primera consiste en reunir entre sí las unidades parecidas para formar un grupo homogéneo; el grupo así formado recibe el nombre de *bloque*.

La segunda etapa consiste en asignar los diferentes tratamientos a las unidades experimentales, sorteando los tratamientos independientemente en cada bloque de la misma manera en que se realizó para el diseño completamente aleatorio. Esta clase de diseño se caracteriza porque cada tratamiento aparece representado una vez en cada uno de los bloques.

Debido a que los tratamientos se asignan aleatoriamente a las unidades experimentales dentro de cada bloque, pero que los bloques se forman decididamente en forma no aleatoria,

en el diseño de bloques completos al azar no se deberá efectuar una prueba estadística del efecto de bloques (ANDERSON y MC LEAN, 1974).

4.3.2 ANÁLISIS DE DATOS EN EL DISEÑO DE BLOQUES AL AZAR

Supóngase que se tienen k tratamientos asignados a b bloques, se realiza una observación por tratamiento en cada bloque, y los tratamientos se asignan aleatoriamente a los bloques. El modelo lineal para un experimento en diseño por bloques al azar requiere un término que represente la variación identificable en las observaciones como consecuencia de los bloques. La respuesta de la unidad con el i ésimo tratamiento en el j ésimo bloque se escribe como (MONTGOMERY, 1991, cap. 5):

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

$$i = 1, 2, \dots, k \quad j = 1, 2, \dots, b \quad (\text{ec. 4.16})$$

Donde μ es una media general, τ_i es el efecto del i ésimo tratamiento, β_j es el efecto del j ésimo bloque y ε_{ij} es el error aleatorio. Según el modelo los efectos del tratamiento y el bloque son aditivos, sin existir interacción; también se supone que los errores experimentales son independientes, con media cero y varianza común σ^2 . La suposición de independencia se justifica a través de la asignación aleatoria de los tratamientos a las unidades experimentales. En la tabla 4.4 se muestra la plantilla de recolección de datos para un diseño completo de bloques al azar.

TABLA 4.4 PLANTILLA DE RECOLECCIÓN DE DATOS PARA EL DISEÑO DE BLOQUES AL AZAR

TRATAMIENTO \ BLOQUE	BLOQUE			Suma	Medias de tratamiento
	1	2 b		
1	Y_{11}	Y_{12}	... Y_{1b}	Y_{k1}	\bar{y}_{k1}
2	Y_{21}	Y_{22}	... Y_{2b}	Y_{k2}	\bar{y}_{k2}
.
.
k	Y_{k1}	Y_{k2}	... Y_{kb}	Y_{kk}	\bar{y}_{kk}
Sumas de bloque	Y_{b1}	Y_{b2}	... Y_{bb}	Y	
Medias de bloque	\bar{y}_{b1}	\bar{y}_{b2}	... \bar{y}_{bb}		\bar{y}

En donde:

Y_{ki} : Suma de las observaciones para el i -ésimo tratamiento.

Y_{bj} : Suma de las observaciones en el j -ésimo bloque.

Y : Suma de todas las bk observaciones.

\bar{y}_{ki} : Media de las observaciones para el i -ésimo tratamiento.

\bar{y}_{bj} : Media de las observaciones en el j -ésimo bloque.

\bar{y} : Media de todas las bk observaciones.

N : Número total de observaciones ($N = bk$)

La desviación de cualquier observación con respecto a la gran media en la tabla 4.4, se puede escribir como la identidad algebraica (KUEHL, 2001, cap. 8):

$$y_{ij} - \bar{y} = (\bar{y}_{i\cdot} - \bar{y}) + (\bar{y}_{\cdot j} - \bar{y}) + (y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}) \quad (\text{ec. 4.17})$$

Los términos del lado derecho de la ecuación 4.17 son:

- Una desviación de tratamiento: $(\bar{y}_{i\cdot} - \bar{y})$
- Una desviación de bloque: $(\bar{y}_{\cdot j} - \bar{y})$
- Error experimental: $(y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y})$

El contraste de hipótesis en el diseño por bloques al azar es el mismo que para el caso del diseño unifactorial, es decir (Montgomery, 1991, cap.5):

$$H_0: \tau_i = 0 \text{ para } i = 1, 2, \dots, k$$

y

$$H_1: \tau_i \neq 0 \text{ cuando menos para un valor de } i$$

El reordenamiento de los términos de la ecuación 4.17, forma una identidad algebraica para la desviación de las observaciones de la media del tratamiento (KUEHL, 2001, cap. 8):

$$y_{ij} - \bar{y}_{i\cdot} = (\bar{y}_{\cdot j} - \bar{y}) + (y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}) \quad (\text{ec. 4.18})$$

Como puede observarse de la ecuación 4.18, se hace una partición de la desviación estándar del error experimental del diseño totalmente aleatorizado, $y_{ij} - \bar{y}_{i\bullet}$, en dos componentes, el primer término se identifica con el criterio de bloques como $(\bar{y}_{\bullet j} - \bar{y})$, el segundo término se identifica sólo como un residual o error experimental, $(y_{ij} - \bar{y}_{i\bullet} - \bar{y}_{\bullet j} + \bar{y})$.

Elevando al cuadrado y sumando ambos términos de la ecuación 4.17, se tiene (KUEHL, 2001, cap. 8):

$$\sum_{i=1}^k \sum_{j=1}^b (y_{ij} - \bar{y})^2 = n \sum_{i=1}^k (\bar{y}_{i\bullet} - \bar{y})^2 + k \sum_{j=1}^b (\bar{y}_{\bullet j} - \bar{y})^2 + \sum_{i=1}^k \sum_{j=1}^b (y_{ij} - \bar{y}_{i\bullet} - \bar{y}_{\bullet j} + \bar{y})^2 \quad (\text{ec. 4.19})$$

La que se puede expresar simbólicamente como:

$$\mathbf{SC}_{\text{Total}} = \mathbf{SC}_{\text{Tratamientos}} + \mathbf{SC}_{\text{Bloques}} + \mathbf{SC}_{\text{Error}} \quad (\text{ec. 4.20})$$

La suma de cualquier producto cruzado en el lado derecho es cero (lo que evidencia que se asume que no existe efecto de interacción entre el factor y el factor de bloque).

Como se tienen N observaciones, la suma de cuadrados total ($\mathbf{SC}_{\text{Total}}$) tiene N - 1 grados de libertad. Las sumas de cuadrados de los tratamientos ($\mathbf{SC}_{\text{Tratamientos}}$) y de los bloques ($\mathbf{SC}_{\text{Bloques}}$), tienen (k - 1) y (b - 1) grados de libertad respectivamente.

Los grados de libertad asociados con la suma de cuadrados del error ($\mathbf{SC}_{\text{Error}}$) se obtienen por sustracción, esto es:

$$\begin{aligned} \text{g.l.}(\mathbf{SC}_{\text{Total}}) &= \text{g.l.}(\mathbf{SC}_{\text{Tratamientos}}) + \text{g.l.}(\mathbf{SC}_{\text{Bloques}}) + \text{g.l.}(\mathbf{SC}_{\text{Error}}) \\ kb - 1 &= (k - 1) + (b - 1) + \text{g.l.}(\mathbf{SC}_{\text{Error}}) \\ \text{g.l.}(\mathbf{SC}_{\text{Error}}) &= kb - 1 - (k - 1) - (b - 1) \\ \text{g.l.}(\mathbf{SC}_{\text{Error}}) &= kb - k - b + 1 \\ \text{g.l.}(\mathbf{SC}_{\text{Error}}) &= k(b - 1) - (b - 1) \\ \text{g.l.}(\mathbf{SC}_{\text{Error}}) &= (k - 1)(b - 1) \end{aligned}$$

Por lo que, los grados de libertad de la suma de cuadrados del error son $(k - 1)(b - 1)$.

Los valores esperados de las sumas de cuadrados para los tratamientos, bloques y el error aleatorio, están dados por (MONTGOMERY, 1991, cap. 5):

$$E(SC_{Tratamientos}) = \sigma^2 + \frac{b \sum_{i=1}^k \tau_i^2}{k-1}$$

$$E(SC_{Bloques}) = \sigma^2 + \frac{k \sum_{j=1}^b \beta_j^2}{b-1}$$

$$E(SC_{Error}) = \sigma^2$$

Suponiendo que los errores se distribuyen normalmente, entonces las cantidades $SC_{Tratamientos}/\sigma^2$, $SC_{Bloques}/\sigma^2$ y SC_{Error}/σ^2 , son variables aleatorias independientes con distribución chi-cuadrado. Por otra parte, cada suma de cuadrados dividida por sus respectivos grados de libertad es igual a una media de cuadrados. Si no hay efecto de los tratamientos o de los bloques los cuadrados medios asociados constituyen ambos estimadores insesgados de σ^2 . Por otra parte, la media de cuadrados del error, constituye un estimador insesgado de σ^2 sin importar que la hipótesis nula sea verdadera o falsa (MONTGOMERY, 1991, cap. 5).

Para probar la hipótesis nula de que los efectos de los tratamientos son todos iguales a cero, se calcula la relación (MONTGOMERY, 1991, cap. 5):

$$F_o = \frac{CM_{Tratamientos}}{CM_{Error}}$$

La cual es un valor de una variable aleatoria que tiene una distribución F con $k-1$ y $(k-1)(b-1)$ grados de libertad cuando la hipótesis nula es verdadera. La hipótesis nula se rechaza con un nivel de significación α cuando (MONTGOMERY, 1991, cap. 5):

$$F_o > F_{\alpha, (k-1), (k-1)(b-1)}$$

La tabla 4.5, resume la partición de la suma de cuadrados. Este procedimiento constituye el análisis de varianza en el diseño de bloques al azar.

TABLA 4.5 ANÁLISIS DE VARIANZA PARA EL DISEÑO DE BLOQUES AL AZAR

Fuente de Variación	g.l.	Suma de Cuadrados (SC)	Cuadrados medios (CM)	Estadístico (F_p)
Tratamiento	k - 1	$SC_{\text{Tratamientos}}$	$SC_{\text{Tratamientos}} / k - 1$	$F_o = \frac{CM_{\text{Tratamientos}}}{CM_{\text{Error}}}$
Bloque	b - 1	SC_{Bloques}	$SC_{\text{Bloques}} / b - 1$	
Error	(k-1)(b-1)	SC_{Error}	$SC_{\text{Error}} / (k-1)(b-1)$	
Total	N-1	SC_{Total}		

Se debe recordar que la aleatorización fue aplicada solo a los tratamientos dentro de los bloques, y que la prueba F del análisis de varianza solamente se puede justificar con base en la aleatorización sin necesidad de usar la suposición de normalidad; es por esa razón que dicha prueba se excluye del análisis de varianza.

La razón entre $CM_{\text{Bloques}}/CM_{\text{Error}}$ puede ser un procedimiento aproximado para investigar el efecto de la variable bloque. Un valor grande de esa razón, implica que el factor bloque tiene un efecto grande y que la reducción del error obtenido al analizar por bloques probablemente fue útil al mejorar la precisión de las comparaciones de las medias de los tratamientos.

Existen ecuaciones más sencillas para el cálculo de la suma de cuadrados que son:

$$SC_{\text{Total}} = \sum_{i=1}^k \sum_{j=1}^b y_{ij}^2 - \frac{Y^2}{bk}$$

$$SC_{\text{Tratamientos}} = \sum_{i=1}^k \frac{y_{i\cdot}^2}{b} - \frac{Y^2}{bk}$$

$$SC_{\text{Bloques}} = \sum_{j=1}^b \frac{y_{\cdot j}^2}{k} - \frac{Y^2}{bk}$$

$$SC_{\text{Error}} = SC_{\text{Total}} - SC_{\text{Tratamientos}} - SC_{\text{Bloques}}$$

Hasta el momento se han descrito procedimientos de prueba, considerando que los tratamientos y los bloques son factores fijos. Sin embargo, el mismo procedimiento de análisis puede usarse si tanto los tratamientos como los bloques (o ambos) son aleatorios. Por supuesto que hay cambios correspondientes en la interpretación de los resultados. Los lotes de material fabricado (como tela, asfalto o producto químico) usados como bloques de tratamientos experimentales son lotes aleatorios, el lote se divide en unidades experimentales más pequeñas a las que se administra el tratamiento (MONTGOMERY, 1991, cap. 5).

Por ejemplo, si los bloques son aleatorios, se espera que las comparaciones entre los tratamientos sean las mismas en toda la población de bloques de la que se eligieron aleatoriamente los que se usaron en el experimento. También existen cambios que corresponden a la esperanza de la media de cuadrados. Por ejemplo, si los bloques son aleatorios, $E(CM_{Bloques}) = \sigma^2 + k\sigma_{Bloques}^2$ en donde $\sigma_{Bloques}^2$ es la componente de varianza del efecto de bloques. La varianza de una media de tratamiento con bloques aleatorios incluye la componente de la varianza para los bloques, $\sigma_{Bloques}^2$, y será mayor que la varianza con efectos de bloques fijos (MONTGOMERY, 1991, cap. 5).

En cualquier caso, $E(CM_{Tratamientos})$ siempre es independiente del efecto de los bloques y el estadístico de prueba para la variabilidad entre los tratamientos siempre es: $F_0 = CM_{Tratamientos}/CM_{Error}$ (MONTGOMERY, 1991, cap. 5).

Si existe la interacción entre los tratamientos y los bloques, la prueba para las medias de los tratamientos no se ve afectada en el caso de que los bloques sean aleatorios. La razón de ello es que tanto el valor esperado de la media de cuadrados de tratamiento como la del error contienen el efecto de interacción. En consecuencia, la prueba para la diferencia en el nivel medio de los tratamientos se realiza en la forma usual comparando la media de cuadrados de tratamiento con la media de cuadrados del error. Esta prueba no produce información sobre la interacción (MONTGOMERY, 1991, cap. 5).

4.3.3 CALCULO DEL TAMAÑO DE LA MUESTRA

La elección adecuada del número de bloques, es una decisión importante cuando se trabaja con un diseño de bloques completos al azar. Al aumentar el número de bloques, aumenta el número de réplicas por tratamiento, ya que cada tratamiento aparece representado en cada uno de los bloques. En otras palabras a mayor número de bloques, pruebas más confiables³. Por otro lado, al aumentar el número de bloques los costos aumentan, el investigador corre el riesgo de perder control sobre el experimento, etc. Por lo que hay necesidad de tomar el número de bloques óptimo que garanticen una probabilidad mínima ($1 - \beta$) de no equivocarse al rechazar la hipótesis nula siendo ésta falsa. Dicha prueba, como se vio en la sección

³bajo el supuesto que el investigador tiene los recursos necesarios para controlar grandes cantidades de datos.

anterior se basa en los cuadrados medios de los tratamientos y el error (MONTGOMERY, 1991, cap. 5).

Matemáticamente la potencia de la prueba $1 - \beta$ viene dada por:

$$1 - \beta = P \left[\frac{CM_{Tratamientos}}{CM_{Error}} > F_{\alpha, (k-1), (k-1)(b-1)} \text{ cuando } H_1 \text{ es verdadera} \right]$$

$$= P \left[\frac{CM_{Tratamientos}}{CM_{Error}} > F_{\alpha, (k-1), (k-1)(b-1)} \text{ cuando } \sum_{i=1}^k \tau_i \neq 0 \right]$$

Cualquiera de las técnicas presentadas en la sección 4.2.3, para la selección del número de réplicas pueden aplicarse a un diseño aleatorizado por bloques (MONTGOMERY, 1991, cap. 5).

Las expresiones para λ^2 y Φ^2 así como los grados libertad para el diseño de bloques aleatorizados se presentan en la tabla 4.6.

TABLA 4.6 VALORES DE λ^2 Y Φ^2 PARA EL DISEÑO DE BLOQUES AL AZAR

Diseño de Bloques			
v_1	v_2	λ^2	Φ^2
$k-1$	$(k-1)(b-1)$	$\frac{b \sum_{i=1}^k \tau_i^2}{\sigma^2}$	$\frac{b \sum_{i=1}^k \tau_i^2}{k\sigma^2}$

En caso que sea difícil encontrar τ_i para calcular Φ^2 , puede obtenerse en función de la diferencia mínima entre dos medias de bloque que se desea determinar (D). De esta forma (MONTGOMERY, 1991, cap. 5):

$$\Phi^2 = \frac{b D^2}{2k \sigma^2}$$

La potencia de la prueba puede incrementarse utilizando un tamaño de bloque b más grande. Para el diseño en bloques se desea determinar un valor de b con el cual se satisfagan los requerimientos de potencia.

El objetivo de encontrar el número de bloques es para asegurar que el número de éstos den una prueba satisfactoria, con una probabilidad $1 - \beta$ de rechazar H_0 siendo H_0 falsa.

Cuando el factor es aleatorio, λ viene dado por (MONTGOMERY, 1991, cap. 5):

$$\lambda = \sqrt{1 + \frac{b\sigma_{\tau}^2}{\sigma^2}}$$

en donde $k-1$ son los grados de libertad del numerador y $(k-1)(b-1)$, los grados de libertad del denominador.

4.3.4 ESTIMACION Y COMPARACION DE PARAMETROS EN EL MODELO DE BLOQUES AL AZAR DE EFECTOS FIJOS

Si tanto los tratamientos como los bloques son fijos, los parámetros del modelo de diseño aleatorio por bloques completos pueden estimarse mediante mínimos cuadrados.

Dado que los efectos de tratamientos y de bloques se consideran como desviaciones de la media general, se imponen las restricciones (MONTGOMERY, 1991, cap. 5)::

$$\sum_{i=1}^k \tau_i = 0 \quad \text{y} \quad \sum_{j=1}^b \beta_j = 0$$

Por lo tanto las ecuaciones normales se reducen a (MONTGOMERY, 1991, cap. 5):

$$kb\mu = Y$$

$$b\mu + b\tau_i = y_{i\bullet} \quad i = 1, 2, \dots, k$$

$$k\mu + k\beta_j = y_{\bullet j} \quad j = 1, 2, \dots, b$$

Cuyas soluciones son (MONTGOMERY, 1991, cap. 5):

$$\mu = \bar{y}$$

$$\tau_i = \bar{y}_{i\bullet} - \bar{y}$$

$$\beta_j = \bar{y}_{\bullet j} - \bar{y}$$

Al sustituir la solución de las ecuaciones normales en las ecuaciones anteriores se encuentra que los valores estimados o ajustados de y_{ij} , son:

$$\hat{y}_{ij} = \mu + \tau_i + \beta_j$$

$$\hat{y}_{ij} = \bar{y}_{i\bullet} + \bar{y}_{\bullet j} - \bar{y}$$

Si los tratamientos en un diseño aleatorizado por bloques son fijos, y el análisis indica una diferencia significativa entre las medias de tratamiento, sería interesante llevar a cabo comparaciones múltiples, para determinar qué medias de tratamiento son diferentes. Cualquiera de los métodos de comparación planteados en la sección 4.2.2.4, puede usarse

con este propósito. Simplemente, hay que reemplazar el número de réplicas del diseño unifactorial completamente aleatorizado (n), por el número de bloques (b) en las fórmulas, por otra parte, también deben usarse como grados de libertad del error, los correspondientes al diseño por bloques al azar: $(k-1)(b-1)$ (MONTGOMERY, 1991, cap. 5).

4.3.5 ANÁLISIS DE RESIDUOS PARA EL DISEÑO POR BLOQUES AL AZAR

Las suposiciones relativas a los errores experimentales del modelo lineal para el diseño de bloques completos al azar se pueden evaluar con un análisis de residuos, como el presentado en la sección 4.2.4. Se calculan los residuos a partir de la componente de la desviación del error experimental mostrada en la ecuación 4.17, como:

$$\varepsilon_{ij} = y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}$$

4.3.6 DISEÑO POR BLOQUES INCOMPLETOS

Es posible que en algunos experimentos en los que se usan diseños aleatorizados por bloques no puedan realizarse los ensayos de todas las combinaciones de tratamiento dentro de cada bloque. Situaciones como esta ocurren debido a la escasez en los recursos del experimento, por el tamaño físico de los bloques. En estos casos es posible usar diseños aleatorizados por bloques en los que cada tratamiento no está presente en cada bloque. Estos diseños se conocen como **diseños aleatorizados por bloques incompletos** (MONTGOMERY, 1991, cap. 6).

Cuando las comparaciones entre todos los tratamientos tienen la misma importancia, estas deben elegirse de manera que ocurran en forma balanceada en cada bloque. Esto significa que cada par de tratamientos ocurren juntos el mismo número de veces que cualquier otro par. Por lo tanto, un diseño balanceado (o equilibrado) por bloques incompletos es un diseño en el que cualquier par de tratamientos ocurren juntos el mismo número de veces (MONTGOMERY, 1991, cap. 6).

Los diseños de bloques incompletos balanceados son los que asignan k tratamientos a b bloques de tamaño r , $r \geq k$, como sigue (KENETT y ZACKS, 2000, cap. 12):

- I. Cada tratamiento sólo se asigna una vez a cualquier bloque.

- II. Cada tratamiento aparece en t bloques; t es la cantidad de réplicas.
- III. Cada par de dos tratamientos distintos aparece en a bloques.
- IV. El orden de tratamientos dentro de cada bloque está aleatorizado.
- V. El orden de los bloques está aleatorizado.

De acuerdo con estos requisitos, hay en total $N = kt = br$ ensayos. Además se debe cumplir la siguiente igualdad: $a(k-1) = t(r-1)$. La cuestión es como diseñar bloques balanceados para determinados k y r . Se puede obtener un diseño de estos haciendo una lista combinatoria completa de las $\binom{k}{r}$ selecciones sin reposición, de r letras tomadas de k letras. En este caso la cantidad de bloques es (KENETT y ZACKS, 2000, cap. 12):

$$b = \binom{k}{r}$$

La cantidad de réplicas es (KENETT y ZACKS, 2000, cap. 12):

$$t = \binom{k-1}{r-1} \quad \text{y} \quad a = \binom{k-2}{r-2}$$

La cantidad total de ensayos es (KENETT y ZACKS, 2000, cap. 12):

$$N = kt = k \binom{k-1}{r-1} = \frac{k!}{(r-1)!(k-r)!} = r \binom{k}{r} = rb$$

Estos diseños de bloques incompletos balanceados se llaman *diseños combinatorios*. Sin embargo, pueden ser demasiado grandes; con frecuencia es posible obtener el balance con menos de $\binom{k}{r}$ bloques. No existe un método único para construir todas las clases de diseños de bloques incompletos pero existen métodos para construir algunos. La construcción de estos diseños ha sido tema de diversas investigaciones matemáticas que han arrojado un vasto arreglo de bloques balanceados y parcialmente balanceados. Al final

del tema se proporcionan algunos planes para pocos tratamientos (KENETT y ZACKS, 2000, cap. 12).

Una vez construido el diseño básico con los números de código de los tratamientos, los pasos para aleatorizar son los siguientes (KENETT y ZACKS, 2000, cap. 12):

Paso 1. Aleatorizar el arreglo de los bloques conformados por grupos de los números código asignados a los tratamientos.

Paso 2. Aleatorizar el arreglo de los números de tratamiento dentro de cada bloque.

Paso 3. Aleatorizar la asignación de los tratamientos a los números de código en el plan.

El modelo estadístico lineal para un diseño de bloques incompleto balanceado es (MONTGOMERY, 1991, cap. 6):

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

$$i = 1, 2, \dots, k \quad j = 1, 2, \dots, b \quad (\text{ec. 4.16})$$

Donde μ es una media general, τ_i es el efecto del i -ésimo tratamiento, β_j es el efecto del j -ésimo bloque y ε_{ij} son los errores aleatorios independientes, con media cero y varianza común σ^2 .

Los efectos de tratamientos y bloques no son ortogonales en los diseños de bloques incompletos porque no aparecen todos los tratamientos en cada bloque; por lo tanto no sería correcto para el diseño incompleto calcular la partición de la suma de cuadrados para los tratamientos igual que para el diseño de bloques completos, tampoco las medias de tratamiento proporcionan estimaciones insesgadas de μ_i . Los parámetros estimados y las sumas de cuadrados de los tratamientos para los diseños de bloque incompleto balanceados se calculan con fórmulas relativamente directas.

Supóngase que T_j es el conjunto de los índices de bloques que contienen el i -ésimo tratamiento. Los estimados de los efectos del tratamiento, por mínimos cuadrados, se obtienen como sigue (MONTGOMERY, 1991, cap. 6):

Supóngase que $W_i = \sum_{i \in T_j} y_{ij}$ es la suma de todos los valores de y pertenecientes al i ésimo

tratamiento. Y que $W_i^* = \sum_{j=1}^b n_{ij} y_{bj}$ es la suma de valores en todos los t bloques que contienen al i ésimo tratamiento, con $n_{ij} = 1$ si el tratamiento i ocurre en el bloque j , y $n_{ij} = 0$ en otro caso. Se calcula (MONTGOMERY, 1991, cap. 6)

$$Q_i = W_i - (1/r)W_i^* \quad i = 1, 2, \dots, k$$

El estimador de τ_i por mínimos cuadrados es

$$\tau_i = \frac{rQ_i}{ka} \quad i = 1, 2, \dots, k$$

Observe que

$$\sum_{i=1}^k Q_i = 0$$

por consiguiente, $\sum_{i=1}^k \tau_i = 0$, sea $\bar{y} = 1/N \sum_{j=1}^b \sum_{i \in B_j} y_{ij}$ (B_j conjunto de tratamientos en el j ésimo bloque).

El promedio ajustado del tratamiento se define como: $\bar{y}_i^* = \bar{y} + \tau_i$, $i = 1, \dots, k$. El análisis de varianza para un diseño de bloques incompleto balanceados se presenta en la tabla 4.7.

TABLA 4.7 ANÁLISIS DE VARIANZA PARA EL DISEÑO DE BLOQUES INCOMPLETOS

Fuente de Variación	g.l.	Suma de Cuadrados (SC)	Cuadrados medios (CM)	Estadístico (F_p)
Tratamiento	k - 1	$\frac{r \sum Q_i^2}{ak}$	SC _{Tratamientos} / k-1	$F_o = \frac{CM_{Tratamientos}}{CM_{Error}}$
Bloque	b - 1	$\sum \frac{y_{bj}^2}{r} - \frac{Y^2}{N}$	SC _{Bloques} / b-1	
Error	N - k - b + 1	SC _{Error} = SC _{Total} - SC _{Bloques (no ajustada)} - SC _{Tratamientos (ajustada)}	SC _{Error} / N-k-b+1	
Total	N - 1	$\sum \sum y_{ij}^2 - \frac{Y^2}{N}$		

4.4 DISEÑO EN CUADRADOS LATINOS

Cuando las unidades experimentales se agrupan en bloques homogéneos en dos direcciones, formando un arreglo de filas y columnas, con la particularidad de que cada fila o columna constituye una repetición completa de los tratamientos, se genera un diseño experimental que se conoce con el nombre de *cuadrado latino*. Si las unidades difieren con respecto a una característica los bloques se basan únicamente en ella; cuando las unidades difieren en cuanto a dos características, se supone básicamente que tanto una como la otra actúan independientemente y aditivamente entre sí, con esta suposición se pueden formar los bloques de acuerdo con una característica y después de acuerdo a la otra. Es decir la columna son los bloques que se basan en la primera característica y las filas son bloques que se basan en la segunda. Tanto las filas como las columnas satisfacen la condición de un bloque con relación a una característica cada vez.

Como se puede observar el *diseño de cuadrados latinos* es una ampliación lógica del diseño de bloques completamente aleatorizados, en él se impone la misma restricción para los bloques, más otra restricción igual para las columnas. En el cuadrado latino, cada tratamiento debe aparecer una vez en cada fila y columna; aquí se requiere que el número de filas y columnas sea igual al número de tratamientos (diseño restrictivo) (KUEHL, 2001, cap.8).

El diseño de *cuadrados latinos* se usa con más frecuencia en experimentos agrícolas e industriales; y en general, cuando se dispone de unidades experimentales cuya variabilidad ocurre en dos sentidos. La particularidad del diseño, de construir bloques completos en sentido de filas y columnas permite absorber en ambos sentidos la variabilidad de las unidades experimentales. La utilidad de este diseño se limita a proyectos de experimentos que comprenden unos pocos tratamientos, tal vez hasta un número no mayor que diez (JURAN y GRZYNA, 1997, cap. 26).

Este diseño permite controlar la variabilidad de las unidades experimentales en dos direcciones (a través del doble bloqueo).

El diseño de *cuadrados latinos* se aplica bajo las siguientes condiciones (COCHRAN, 1965, cap. 4):

i) Cuando las unidades experimentales no sean homogéneas; es decir que no puedan reaccionar o responder a los tratamientos de la misma manera o con la misma capacidad debido a sus diferencias.

ii) Cuando el investigador tenga suficientes elementos que le garanticen la heterogeneidad en las unidades experimentales en dos direcciones.

iii) Cuando sea posible formar bloques según dos características.

iv) Donde es poco probable que las unidades experimentales se destruyan o no respondan a un tratamiento determinado.

Entre las ventajas del empleo del diseño cuadrado latino se encuentran (COCHRAN, 1965, cap. 4):

- Disminuyen los efectos de dos fuentes de variación en las unidades experimentales; ya que el efecto de filas y columnas no está incluido en el error.
- Reduce el número de combinaciones de tratamiento.
- El análisis de varianza no es complicado.

Entre las limitaciones de la aplicación del diseño cuadrado latino (COCHRAN, 1965, cap. 4):

- Como el número de tratamientos depende del número de filas y columnas esto le quita flexibilidad al diseño para su uso.
- Se cuenta con menos grados de libertad para el error experimental que con los diseños anteriores.
- El número de unidades experimentales requeridas para establecer un experimento, se incrementa notablemente a medida que aumenta el número de tratamientos en ensayo.

4.4.1 DISTRIBUCION ALEATORIA DE TRATAMIENTOS EN DISEÑOS CUADRADOS LATINOS

Para una mejor comprensión, la distribución en el *diseño de cuadrados latinos* se explicará mediante un ejemplo. Supóngase que se desean probar cuatro tratamientos (A, B, C, D). Se deberá hacer cuatro réplicas de cada tratamiento. De acuerdo con el principio básico de los bloques al azar, los tratamientos deberán asignarse al azar a las cuatro unidades de un bloque.

Puesto que cada fila es un bloque, cada fila de cuatro unidades deberá tener A, B, C, D en un orden cualquiera.

Asimismo, puesto que cada columna es un bloque, deberá tener A, B, C, D en un orden cualquiera. Se debe comprobar detenidamente que en cada fila y en cada columna aparecen A, B, C, D sin duplicación.

Un ejemplo de la distribución de unidades para el diseño cuadrado latino seleccionado podría ser:

		COLUMNAS			
FILAS		1	2	3	4
1		A	B	C	D
2		D	A	B	C
3		C	D	A	B
4		B	C	D	A

En general, la aleatorización en un cuadrado latino de tamaño $t \times t$ se logra con los siguientes pasos (KUEHL, 2001, cap. 8):

- Paso 1.* Seleccionar al azar uno de los cuadrados estándar
- Paso 2.* Ordenar al azar todos menos el primer renglón
- Paso 3.* Ordenar al azar todas las columnas
- Paso 4.* Asignar al azar los tratamientos con letras

En la tabla 4.8, se presentan distribuciones sugeridas de cuadrados latinos de diversos tamaños.

TABLA 4.8 CUADRADOS LATINOS SELECCIONADOS
(JURAN y GRZYNA, 1997, cap. 26)

		4 × 4			
3 × 3		1	2	3	4
<i>ABC</i>	<i>ABCD</i>	<i>ABCD</i>	<i>ABCD</i>	<i>ABCD</i>	<i>ABCD</i>
<i>BCA</i>	<i>BADC</i>	<i>BCDA</i>	<i>BDAC</i>	<i>BADC</i>	<i>BADC</i>
<i>CAB</i>	<i>CDBA</i>	<i>CDAB</i>	<i>CADB</i>	<i>CDAB</i>	<i>CDAB</i>
	<i>DCAB</i>	<i>DABC</i>	<i>DCBA</i>	<i>DCBA</i>	<i>DCBA</i>
5 × 5		6 × 6		7 × 7	
<i>ABCDE</i>		<i>ABCDEF</i>		<i>ABCDEFG</i>	
<i>BAECD</i>		<i>BFDCAE</i>		<i>BCDEFGA</i>	
<i>CDAEB</i>		<i>CDEFBA</i>		<i>CDEFGAB</i>	
<i>DEBAC</i>		<i>DAFE CB</i>		<i>DEFGABC</i>	
<i>ECDBA</i>		<i>ECABFD</i>		<i>EFGABCD</i>	
		<i>FEBAD C</i>		<i>FGABCDE</i>	
				<i>GABCDEF</i>	

4.4.2 ANÁLISIS DE LOS DISEÑOS CUADRADOS LATINOS

Como ya se mencionó el diseño de cuadrado latino se usa para eliminar dos fuentes de variación no controlada, es decir, permite analizar sistemáticamente por bloques en dos direcciones. En este diseño, las filas y las columnas representan las restricciones de la aleatorización. Un cuadrado latino $r \times r$, es un cuadrado que contiene r filas y r columnas. El modelo estadístico lineal para este diseño es el siguiente (MONTGOMERY, 1991, cap. 5):

$$y_{mij} = \mu + \alpha_m + \tau_i + \beta_j + \varepsilon_{mij}$$

$$i = 1, 2, \dots, k \quad j = 1, 2, \dots, b \quad (\text{ec. 4.21})$$

$$m = 1, 2, \dots, l$$

Donde:

y_{mij} : Representa una observación en la m -ésima fila y en la j -ésima columna correspondiente al i -ésimo tratamiento.

α_m y β_j : Son los efectos de la m -ésima fila y de la j -ésima columna respectivamente.

τ_i : Es el efecto del i -ésimo tratamiento.

μ : Es la media principal

ε_{mij} : El error aleatorio, independiente con media 0 y varianza σ^2

Las hipótesis a probar en este tipo de diseño son las mismas que para el caso del diseño unifactorial, es decir:

$$H_0: \tau_i = 0 \text{ para todo } i = 1, 2, \dots, k$$

En forma correspondiente, la hipótesis alternativa es:

$$H_1: \tau_i \neq 0 \text{ cuando menos para un valor de } i$$

Aquí se imponen las restricciones (KENETT y ZACKS, 2000, cap. 12):

$$\sum_{m=1}^l \alpha_m = \sum_{i=1}^k \tau_i = \sum_{j=1}^b \beta_j = 0$$

Se supone que no hay interacción entre los tratamientos y las columnas y los renglones. La notación para totales y las medias de las observaciones para renglones y columnas siguen la convención usual con $Y_m = \sum_j^l y_{ij}$ para el total de un renglón y $Y_j = \sum_i^k y_{ij}$ para el total de columna. El total de tratamiento estará representado como Y_i que implica una suma de observaciones sobre las unidades experimentales que reciben el tratamiento k . De igual forma, \bar{y}_{ki} representa la media de las observaciones en el k -ésimo tratamiento. Se asume que y_{mij} son valores de variables aleatorias independientes que tienen distribuciones normales con media μ_{mij} donde $\mu_{mij} = \mu + \alpha_m + \tau_i + \beta_j$. Siendo la varianza común σ^2 .

Las particiones de sumas de cuadrados se pueden derivar de la identidad algebraica (KUEHL, 2001, cap. 8):

$$y_{imj} - \bar{y} = (\bar{y}_{ki} - \bar{y}) + (\bar{y}_{bj} - \bar{y}) + (\bar{y}_{lm} - \bar{y}) + (y_{imj} - \bar{y}_{ki} - \bar{y}_{bj} - \bar{y}_{lm} + 2\bar{y}) \quad (\text{ec. 4.22})$$

Los términos del lado derecho de la ecuación 4.22 son:

- Una desviación de tratamiento: $(\bar{y}_{ki} - \bar{y})$
- Una desviación de la columna: $(\bar{y}_{bj} - \bar{y})$
- Una desviación de la fila: $(\bar{y}_{lm} - \bar{y})$
- Error experimental: $(y_{imj} - \bar{y}_{ki} - \bar{y}_{bj} - \bar{y}_{lm} + 2\bar{y})$

Al elevar al cuadrado ambos lados de la ecuación 4.22 y sumar los términos se llega a la partición aditiva de:

$$\sum_{i=1}^k \sum_{m=1}^l \sum_{j=1}^b (y_{imj} - \bar{y})^2 = l \sum_{l=1}^m (\bar{y}_{ml} - \bar{y})^2 + l \sum_{j=1}^b (\bar{y}_{bj} - \bar{y})^2 + l \sum_{i=1}^k (\bar{y}_{ki} - \bar{y})^2 +$$

$$\sum_{i=1}^k \sum_{j=1}^b \sum_{m=1}^l (y_{imj} - \bar{y}_{ki} - \bar{y}_{bj} - \bar{y}_{ml} + \bar{y})^2$$

Es decir,

$$SC_{Total} = SC_{Tratamientos} + SC_{Columnas} + SC_{Renglones} + SC_{Error} \quad (\text{ec. 4.23})$$

Al dividir cada una de las sumas de cuadrados del lado derecho de la ecuación 4.23 por sus correspondientes números de grados de libertad, considerando que $b = k = l = r$, se obtienen cuatro estimaciones independientes de σ^2 que son: el cuadrado medio de los renglones, columnas, tratamientos y error, respectivamente, y que se denotan por:

$$\begin{aligned} CM_{Renglones} &= SC_{Renglones} / r - 1 \\ CM_{Columnas} &= SC_{Columnas} / r - 1 \\ CM_{Tratamientos} &= SC_{Tratamientos} / r - 1 \\ CM_{Error} &= SC_{Error} / (r-2)(r-1). \end{aligned}$$

Para probar la hipótesis nula de que los efectos de los tratamientos son todos iguales a cero, se calcula la relación (MONTGOMERY, 1991, cap. 5):

$$F_o = \frac{CM_{Tratamientos}}{CM_{Error}}$$

La cual es un valor de una variable aleatoria que tiene una distribución F con $k-1$ y $(r-2)(r-1)$ grados de libertad cuando la hipótesis nula es verdadera. La hipótesis nula se rechaza con un nivel de significación α cuando (MONTGOMERY, 1991, cap. 5):

$$F_o > F_{\alpha, (k-1), (r-2)(r-1)}$$

El cuadrado medio del error es un estimador insesgado de σ^2 , sin importar si la hipótesis nula es cierta. Si H_o es cierta, $CM_{Renglones}$, $CM_{Columnas}$, $CM_{Tratamientos}$ son estimadores insesgados de σ^2 también; en caso contrario se dan las siguientes relaciones (BOX et al., 1989, cap. 8):

$$\begin{aligned} E(CM_{Renglones}) &= \sigma^2 + \frac{r \sum_{i=1}^r \alpha_i^2}{r-1} & E(CM_{Columnas}) &= \sigma^2 + \frac{r \sum_{j=1}^r \beta_j^2}{r-1} \\ E(CM_{Tratamientos}) &= \sigma^2 + \frac{r \sum_{i=1}^r \tau_i^2}{r-1} & E(CM_{Error}) &= \sigma^2 \end{aligned}$$

En la tabla 4.9, se presenta el análisis de varianza para un diseño cuadrado latino $r \times r$.

TABLA 4.9 ANÁLISIS DE VARIANZA PARA EL DISEÑO CUADRADO LATINO $r \times r$

Fuentes de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	Estadístico F_o
Renglones	$SC_{\text{Renglones}}$	$r-1$	$CM_{\text{Renglones}}$	
Columnas	SC_{Columnas}	$r-1$	CM_{Columnas}	
Tratamientos	$SC_{\text{Tratamientos}}$	$r-1$	$CM_{\text{Tratamientos}}$	$F_o = \frac{CM_{\text{Tratamientos}}}{CM_{\text{Error}}}$
Error	SC_{Error}	$(r-1)(r-2)$	CM_{Error}	
Total	SC_{Total}	r^2-1		

Para facilitar el cálculo manual de las expresiones de la tabla se presentan las alternativas siguientes (MONTGOMERY, 1991, cap.5):

$$SC_{\text{Total}} = \sum_{i=1}^r \sum_{j=1}^r \sum_{l=1}^r y_{ijl}^2 - \left(\frac{Y}{r}\right)^2$$

$$SC_{\text{Renglones}} = \frac{\sum_{l=1}^r y_{ml}^2}{r} - \left(\frac{Y}{r}\right)^2$$

$$SC_{\text{Columnas}} = \frac{\sum_{j=1}^r y_{bj}^2}{r} - \left(\frac{Y}{r}\right)^2$$

$$SC_{\text{Tratamientos}} = \frac{\sum_{i=1}^r y_{ki}^2}{r} - \left(\frac{Y}{r}\right)^2$$

$$SC_{\text{Error}} = SC_{\text{Total}} - SC_{\text{Renglones}} - SC_{\text{Columnas}} - SC_{\text{Tratamientos}}$$

4.4.3 CÁLCULO DEL TAMAÑO DE LA MUESTRA EN DISEÑOS CUADRADOS LATINOS

Como en todo diseño experimental, en el cuadrado latino es necesario también calcular un tamaño de muestra que garantice una probabilidad alta de rechazar la hipótesis nula si ésta es falsa. La potencia de la prueba $(1 - \beta)$ para el diseño cuadrado latino viene dada como:

$$1 - \beta = P \left[\frac{CM_{Tratamientos}}{CM_{Error}} > F_{\alpha, (r-1), (r-1)(r-2)} \text{ cuando } H_1 \text{ es verdadera} \right]$$

$$= P \left[\frac{CM_{Tratamientos}}{CM_{Error}} > F_{\alpha, (r-1), (r-1)(r-2)} \text{ cuando } \sum_{i=1}^k \tau_i \neq 0 \right]$$

Siguiendo el mismo análisis del tamaño de muestra para diseños anteriores se define Φ^2 por la relación:

$$\Phi^2 = \frac{b \sum_{i=1}^r \tau_i^2}{r\sigma^2} = \frac{\lambda^2}{r}$$

El procedimiento para determinar el número de bloques fila o columna para garantizar una prueba potente es similar a los estudiados anteriormente.

4.4.4 ANÁLISIS DE RESIDUOS PARA EVALUAR SUPOSICIONES

Los residuos se pueden usar para evaluar las suposiciones del modelo (como se describe en la sección 4.2.4, de la ecuación 4.22, el residuo de las observaciones del i -ésimo tratamiento en el m -ésimo renglón y la j -ésima columna es:

$$\varepsilon_{imj} = y_{imj} - \bar{y}_{ki} - \bar{y}_{bj} - \bar{y}_{lm} + 2\bar{y} \quad (\text{ec. 4.24})$$

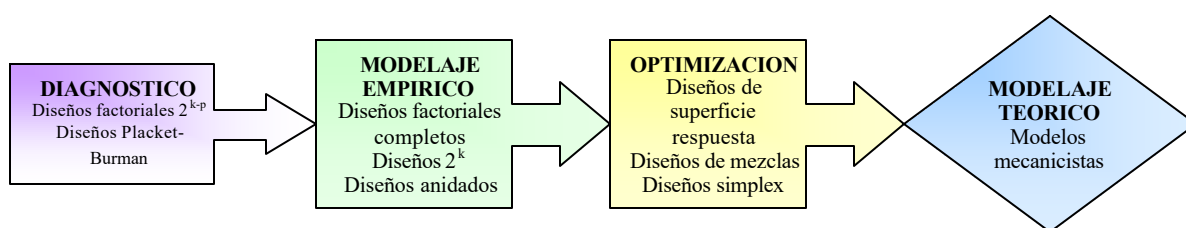
4.5 DISEÑOS FACTORIALES

El experimento con un factor se ha convertido en el modelo típico de la investigación científica, ya que ha resultado ser muy adecuado. Sin embargo, en muchas situaciones se requiere del estudio del efecto de más de un factor. Por ejemplo, en análisis químico, muchas veces es de interés el estudio de cómo dos o más factores controlables en un proceso de medición pueden influir en la cuantificación de un analito contenido en una determinada muestra, o bien cuales son los niveles de los factores que permiten cuantificar óptimamente la cantidad de dicho analito.

En éstos casos los diseños unifactoriales además de ser inadecuados por lo restringido, son imprecisos en cuanto a los resultados obtenidos. En ésta sección se estudian diseños experimentales que permiten evaluar simultáneamente el efecto de dos o más factores de

tipo cualitativo o cuantitativo sobre una respuesta, y que en general, son conocidos como **diseños factoriales**.

Los diseños factoriales pueden clasificarse como: *diseños factoriales completos*, *diseños factoriales 2^k* , *diseños factoriales fraccionarios 2^{k-p}* , *diseños jerárquicos o anidados*, *diseños de mezclas*, *experimentos factoriales con restricciones en la aleatorización*, y *diseños de superficie respuesta*. En el cuadro 4.1 (sección 4.1), se presentan los objetivos de la aplicación de cada uno de estos tipos de diseños. En el esquema 4.1 se presenta la secuencia de utilización de los diseños factoriales.



ESQUEMA 4.1 SECUENCIA DEL DESARROLLO DE EXPERIMENTOS FACTORIALES

Un diseño factorial, en general, es aquel proceso en el que además de investigar independientemente el efecto producido por cada factor, sobre la respuesta, también se estudia el efecto producido por la interacción entre dichos factores, ya que toma en cuenta todas las posibles combinaciones de los niveles de los factores, o combinación de tratamientos, en cada ensayo completo o réplica del experimento.

En un experimento factorial, se eligen los niveles de cada factor y se hacen las mediciones en cada una de las posibles combinaciones de niveles de los factores, manteniendo el principio de la aleatorización. Los factores se representan por letras mayúsculas mientras que sus niveles con las letras minúsculas.

En el análisis de los experimentos factoriales, se habla de estimación de *efectos principales* y de *efectos de interacción* (o simplemente *interacciones*). Los efectos principales de un factor dado son siempre función de la respuesta o resultado medio de los distintos niveles de dicho factor sobre la respuesta. Las interacciones resultan cuando en dos o más factores el efecto de uno de ellos sobre la respuesta depende de los niveles del otro factor o factores. Las interacciones pueden ser de primer orden, cuando se presenta el efecto combinado de dos factores entre sí, de segundo orden cuando intervienen tres factores; de tercer orden cuando intervienen cuatro factores y así sucesivamente.

Entre las ventajas de la aplicación de este tipo de diseño experimental se encuentran:

- i. Facilitan el estudio de varios factores simultáneamente.
- ii. Requieren menor número de observaciones que las necesarias, para estudiar cada factor independientemente.
- iii. Proporcionan una información más completa de cada factor en estudio, pues no sólo se comprueban los efectos aislados de cada uno, sino también las interacciones entre sí.
- iv. Las conclusiones pueden tener mayor valor desde un punto de vista práctico, pues los distintos niveles de cada factor se experimentan en presencia del otro factor o factores.
- v. La precisión en el análisis de varianza es mayor, ya que el cálculo estadístico es más preciso.

Entre las limitaciones de la aplicación de este tipo de diseños se encuentran:

- i. El planeamiento requiere una cuidadosa administración así como de la interpretación de resultados.
- ii. El análisis de los datos se complica cuando se incluyen muchos factores.

En las secciones 4.5 y 4.6 se presenta la descripción para la realización de los diseños factoriales completos y 2^k respectivamente. Un estudio detallado de los demás tipos de diseños factoriales se considera fuera del alcance de este trabajo, pudiéndose estudiar con más detalle en MONTGOMERY (1991), PRAT et al. (1999) y LAWSON et al. (1992) (ver referencias bibliográficas).

4.5.1 EXPERIMENTOS FACTORIALES COMPLETOS

Los experimentos factoriales completos son aquellos en los que se efectúan pruebas completas de todas las combinaciones de los diversos factores de carácter fijo o aleatorio.

El número de combinaciones de tratamiento a obtener en un diseño factorial viene dado por el producto: $(a_A) \times (b_B) \times \dots \times (q_Q)$, en donde i_i indica el número de niveles o tratamientos para el factor i . Todas ellas se prueban. También el experimento factorial completo se puede replicar n_i veces para cada combinación de tratamientos. Si se tiene el

mismo número de niveles para cada factor y el mismo número de replicas para cada combinación, la representación del diseño puede reducirse a $P^k \times n$, en donde P indica el número de niveles común de todos los factores, el superíndice k indica el número de factores y n el número de replicas para cada combinación de tratamientos. El orden en el que se efectúan las pruebas es aleatorio.

Los niveles de un factor *cuantitativo* toman valores métricos, mientras que los niveles de un factor *cualitativo* son las categorías del factor. Por ejemplo el factor temperatura de secado esta expresado en escala métrica (25°C, 30°C, etc.) y el factor material del envase es de carácter cualitativo (plástico, vidrio, etc.).

El *efecto* de un factor es un cambio en la respuesta media ocasionado por un cambio en el nivel de ese factor. Los dos efectos de interés en un experimento factorial son los *simples o principales* y los *de interacción*. Los efectos simples o principales de un factor son las comparaciones entre los niveles del mismo. Los efectos de interacción miden las diferencias entre los efectos simples de un factor a diferentes niveles del otro. Por ejemplo, si hay l factores, A, B, C,... a dos niveles cada uno hay l tipos de efectos principales, hay $\binom{l}{2}$ tipos de interacciones apareadas AB, AC, BC,....., $\binom{l}{3}$ interacciones entre tres factores, ABC, ABD,, etcétera. En total junto con la media general μ , hay 2^l tipos de parámetros.

4.5.1.1 ANÁLISIS DE VARIANZA PARA DISEÑOS FACTORIALES COMPLETOS

El análisis de varianza para *diseños factoriales completos* se hace para probar las hipótesis de que los efectos principales o los de interacción son igual a cero. A continuación se presentará el análisis de varianza para un caso con dos factores: el factor A con a niveles y el factor B con b niveles. El método se puede generalizar a cualquier otra cantidad de factores.

La estructura del experimento es tal que se prueban todas las $a \times b$ combinaciones. Cada combinación de tratamientos se repite n veces. El modelo estadístico lineal es:

$$y_{ijk} = \mu + \tau_i^A + \tau_j^B + \tau_{ij}^{AB} + \varepsilon_{ijk} \quad (\text{ec. 4.25})$$

$$i = 1, 2, \dots, a \quad j = 1, 2, \dots, b \quad k = 1, 2, \dots, n$$

En donde μ es la media general; τ_i^A es el efecto principal del factor A ; τ_j^B es el efecto principal del factor B ; τ_{ij}^{AB} es el efecto de interacción de primer orden entre los factores A y B ; y ε_{ijk} es el error aleatorio k -ésima réplica. Se supone que los ε_{ijk} son variables aleatorias independientes para las que $E[\varepsilon_{ijk}] = 0$ y $V[\varepsilon_{ijk}] = \sigma^2$ para toda i, j y k .

Sean a y b el número de niveles del factor A y B respectivamente, n el número de réplicas en cada tratamiento (combinación de niveles) y y_{ijk} la respuesta observada cuando el factor A se encuentra en el i -ésimo nivel, el factor B en el j -ésimo nivel y en la k -ésima réplica. En general, los datos para un diseño bifactorial se organizan como en la tabla 4.10.

TABLA 4.10 PLANTILLA DE RECOLECCION DE DATOS PARA UN DISEÑO BIFACTORIAL

FACTOR A	FACTOR B			
	1	2	...	b
1	$Y_{111}, Y_{112}, \dots, Y_{11n}$	$Y_{121}, Y_{122}, \dots, Y_{12n}$...	$Y_{1b1}, Y_{1b2}, \dots, Y_{1bn}$
2	$Y_{211}, Y_{212}, \dots, Y_{21n}$	$Y_{221}, Y_{222}, \dots, Y_{22n}$...	$Y_{2b1}, Y_{2b2}, \dots, Y_{2bn}$
.
.
a	$Y_{a11}, Y_{a12}, \dots, Y_{a1n}$	$Y_{a21}, Y_{a22}, \dots, Y_{a2n}$...	$Y_{ab1}, Y_{ab2}, \dots, Y_{abn}$

El análisis de varianza cuando se considera más de un factor es similar al unifactorial; prácticamente la lógica y fundamentos teóricos que lo sustentan son los mismos.

La partición de la suma de cuadrados total se puede deducir de la ecuación:

$$y_{ij} = \sum_{k=1}^n y_{ijk} \rightarrow \bar{y}_{ij} = \frac{1}{n} \sum_{k=1}^n y_{ijk}$$

$$y_{i\cdot} = \sum_{j=1}^b \bar{y}_{ij} \rightarrow \bar{y}_{i\cdot} = \frac{1}{b} \sum_{j=1}^b \bar{y}_{ij} \quad i = 1, \dots, a$$

$$y_{\cdot j} = \sum_{i=1}^a \bar{y}_{ij} \rightarrow \bar{y}_{\cdot j} = \frac{1}{a} \sum_{i=1}^a \bar{y}_{ij} \quad j = 1, \dots, b$$

$$Y = \sum_{i=1}^a \sum_{j=1}^b \bar{y}_{ij} \rightarrow \bar{Y} = \frac{1}{ab} \sum_{i=1}^a \sum_{j=1}^b \bar{y}_{ij}$$

El análisis de varianza divide primero la suma total de cuadrados de desviaciones respecto a \bar{Y} :

$$SC_{Total} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{Y})^2$$

en dos componentes, la suma dentro de tratamientos (SC_{Error}) y la suma entre tratamientos ($SC_{Tratamientos}$), es decir:

$$SC_{Error} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{ij})^2 \quad \text{y} \quad SC_{Tratamientos} = n \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij} - \bar{Y})^2$$

Se puede demostrar que: $SC_{Total} = SC_{Error} + SC_{Tratamientos}$. En la segunda etapa, se divide la suma de cuadrados de los tratamientos ($SC_{Tratamientos}$), en tres componentes: de las interacciones ($SC_{Interacciones}$), del efecto principal de A (SC_A) y del efecto principal de B (SC_B), son:

$$SC_{Interacciones} = SC_{AB} = n \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{Y})^2$$

$$SC_A = nb \sum_{j=1}^b (\bar{y}_{.j} - \bar{Y})^2$$

$$SC_B = na \sum_{i=1}^a (\bar{y}_{i.} - \bar{Y})^2$$

Es decir,

$$SC_{Tratamientos} = SC_{Interacciones} + SC_A + SC_B$$

En la tabla 4.11, se presenta el análisis de varianza para el modelo bifactorial general.

TABLA 4.11 ANÁLISIS DE VARIANZA PARA EL MODELO BIFACTORIAL GENERAL

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	F
Factor A	SC _{F_A}	a-1	CM _{F_A}	F_A
Factor B	SC _{F_B}	b-1	CM _{F_B}	F_B
Interacción AB	SC _{F_{AB}}	(a-1)(b-1)	CM _{F_{AB}}	F_{AB}
Error	SC _{Error}	ab(n-1)	CM _{Error}	
Total	SC _{Total}	abn-1		

A continuación se presentan formulas que permitirán calcular las diferentes sumas de cuadrados de una forma más sencilla. Para la suma de cuadrados total se tiene:

$$SC_{Total} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 - \frac{nY^2}{ab}$$

Las sumas de cuadrados para los efectos principales son:

$$SC_A = \sum_{i=1}^a \frac{ny_{i\cdot}^2}{b} - \frac{nY^2}{ab} \quad SC_B = \sum_{j=1}^b \frac{ny_{\cdot j}^2}{a} - \frac{nY^2}{ab}$$

Es conveniente obtener la $SC_{Interacciones}$ en dos etapas. Primero se calcula la suma de cuadrados entre los totales de las ab celdas, conocida como la suma de cuadrados de los "subtotales":

$$SC_{Subtotales} = \sum_{i=1}^a \sum_{j=1}^b \frac{y_{ij}^2}{n} - \frac{nY^2}{ab}$$

Esta suma de cuadrados contiene a la suma de cuadrados de los factores A y B . Por lo tanto, la segunda etapa consiste en calcular $SC_{Interacciones}$ mediante (MONTGOMERY, 1991, cap. 7):

$$SC_{Interacciones} = SC_{Subtotales} - SC_A - SC_B$$

Mientras que la suma de cuadrados del error se calcula por diferencia:

$$SC_{Error} = SC_{Total} - SC_{Subtotales}$$

4.5.1.2 INFERENCIA PARA EL MODELO DE EFECTOS FIJOS

En un diseño factorial de dos factores fijos, tanto los factores (o tratamientos) de renglón como los de columna tienen la misma importancia. Específicamente el interés consiste en probar hipótesis acerca de la igualdad de los efectos de renglón, columna y de los de interacción, es decir:

Para el factor A:

$$H_0 : \tau_i^A = 0$$

$$H_1 : \tau_j^A \neq 0 \text{ al menos para un } i$$

Para el factor B:

$$H_o : \tau_j^B = 0$$

$$H_1 : \tau_j^B \neq 0 \text{ al menos para un } j$$

Para el factor de interacción AB:

$$H_o : \tau_{ij}^{AB} = 0$$

$$H_1 : \tau_{ij}^{AB} \neq 0 \text{ al menos para un par } ij$$

Cada suma de cuadrados divididos entre sus grados de libertad produce una media de cuadrados. Los valores esperados de las medias de cuadrados son (MONTGOMERY, 1991, cap. 7):

$$E(CM_A) = E\left(\frac{SC_A}{a-1}\right) = \sigma^2 + \frac{bn \sum_{i=1}^a (\tau_i^A)^2}{a-1}$$

$$E(CM_B) = E\left(\frac{SC_B}{b-1}\right) = \sigma^2 + \frac{an \sum_{j=1}^b (\tau_j^B)^2}{b-1}$$

$$E(CM_{AB}) = E\left(\frac{SC_{AB}}{(a-1)(b-1)}\right) = \sigma^2 + \frac{n \sum_{i=1}^a \sum_{j=1}^b (\tau_{ij}^{AB})^2}{(a-1)(b-1)}$$

$$E(CM_{Error}) = E\left(\frac{SC_{Error}}{ab(n-1)}\right) = \sigma^2$$

Si las hipótesis nulas son verdaderas, entonces los cuadrados medios para los factores de renglón, columna y de interacción (CM_A , CM_B y CM_{AB}) son estimadores insesgados de σ^2 . Sin embargo, al existir diferencias en los tratamientos de renglón entonces $CM_A > CM_{Error}$. En forma similar, si hay efecto de tratamiento columna o interacciones, las medias de los cuadrados correspondientes serán mayores que CM_{Error} . Por lo tanto, para probar el significado de ambos efectos principales, así como de su interacción simplemente deben dividirse las medias de cuadrados correspondientes entre las medias del cuadrado del error; valores grandes de estas razones implican que los datos no concuerdan con la hipótesis nula. En otras palabras, si se considera que el modelo es adecuado y que los términos del error ε_{ij} son independientes con distribución normal y varianza constante, entonces las razones de las medias de cuadrados de los factores principales y de la interacción con respecto a la media

de cuadrados del error (CM_A/CM_{Error} , CM_B/CM_{Error} y CM_{AB}/CM_{Error}) tienen una distribución F con $a - 1$, $b - 1$ y $(a - 1)(b - 1)$ grados de libertad en el numerador, y $ab(n - 1)$ grados de libertad en el denominador respectivamente (MONTGOMERY, 1991, cap. 7).

De esta forma la hipótesis nula para el factor A se rechaza si (MONTGOMERY, 1991, cap. 7):

$$F_A = CM_A/CM_{Error} > F_{\alpha, a-1, ab(n-1)}$$

Para el factor B la hipótesis nula se rechaza si (MONTGOMERY, 1991, cap. 7):

$$F_B = CM_B/CM_{Error} > F_{\alpha, b-1, ab(n-1)}$$

Y finalmente, la hipótesis nula para el factor de interacción se rechaza si (MONTGOMERY, 1991, cap. 7):

$$F_{AB} = CM_{AB}/CM_{Error} > F_{\alpha, (a-1)(b-1), ab(n-1)}$$

Los parámetros en el modelo del diseño bifactorial, pueden estimarse por el método de los mínimos cuadrados, imponiendo las restricciones (MONTGOMERY, 1991, cap. 7):

$$\sum_{i=1}^a \tau_i^A = 0 \quad \sum_{j=1}^b \tau_j^B = 0 \quad \sum_{i=1}^a \tau_{ij}^{AB} = 0 \quad j = 1, 2, \dots, b \quad \sum_{j=1}^b \tau_{ij}^{AB} = 0 \quad i = 1, 2, \dots, a$$

Al aplicar estas restricciones, las ecuaciones normales se simplifican considerablemente y se obtienen las soluciones (MONTGOMERY, 1991, cap. 7):

$$\begin{aligned} \mu &= \bar{Y} \\ \tau_i^A &= \bar{y}_i - \bar{Y} \\ \tau_j^B &= \bar{y}_j - \bar{Y} \\ \tau_{ij}^{AB} &= \bar{y}_{ij} - \bar{y}_i - \bar{y}_j + \bar{Y} \end{aligned}$$

Usando las ecuaciones anteriores puede determinarse el valor ajustado de y_{ijk} mediante:

$$\hat{y}_{ijk} = \bar{y}_{ij}$$

En otras palabras, la k -ésima observación de la ij -ésima celda se estima mediante el promedio de las n observaciones de dicha celda.

4.5.1.3 INFERENCIA PARA EL MODELO DE EFECTOS ALEATORIOS

Considérese ahora la situación en la que los niveles de los factores A y B se eligen al azar de poblaciones más grandes, este caso es relativo al modelo de efectos aleatorios o de componentes de varianza. Las inferencias pueden generalizarse a todos los niveles de las poblaciones bajo estudio, porque los a niveles de A y los b niveles de B se eligieron al azar. Las observaciones se pueden representar mediante el modelo lineal, de la ecuación 4.25, sin embargo, se hace la suposición adicional de que $\tau_i^A, \tau_j^B, \tau_{ij}^{AB}$ son variables aleatorias independientes normalmente distribuidas con media cero y varianzas respectivas σ_A^2, σ_B^2 y σ_{AB}^2 constantes. En general, la varianza de cualquier observación es (MONTGOMERY, 1991, cap. 7):

$$V(y_{ijk}) = \sigma_A^2 + \sigma_B^2 + \sigma_{AB}^2 + \sigma^2$$

en donde σ_A^2, σ_B^2 y σ_{AB}^2 y σ^2 se conocen como componentes de varianza.

Las hipótesis que deben probarse para el caso de dos factores aleatorios son:

$$H_0: \sigma_A^2 = 0, \quad H_0: \sigma_B^2 = 0, \quad H_0: \sigma_{AB}^2 = 0$$

El análisis de varianza básico permanece sin cambios, en otras palabras, SC_A, SC_B, SC_{AB} , y SC_{Error} se calculan como en el modelo de efectos fijos. Sin embargo, para determinar las estadísticas con objeto de probar las hipótesis, deben examinarse antes los valores esperados de las medias de cuadrados (MONTGOMERY, 1991, cap. 7):

$$E(CM_A) = \sigma^2 + n\sigma_{AB}^2 + bn\sigma_A^2$$

$$E(CM_B) = \sigma^2 + n\sigma_{AB}^2 + an\sigma_B^2$$

$$E(CM_{AB}) = \sigma^2 + n\sigma_{AB}^2$$

$$E(CM_{Error}) = \sigma^2$$

A partir de estos valores esperados, puede apreciarse que la estadística apropiada para

probar la hipótesis $H_0: \sigma_{AB}^2 = 0$ es $F_{AB} = \frac{CM_{AB}}{CM_{Error}}$, porque si la hipótesis nula es verdadera,

tanto el numerador como el denominador de F_{AB} tienen un valor esperado igual a σ^2 , y solo si H_0 es falsa $E(CM_{AB})$ será mayor que $E(CM_{Error})$. La razón F_{AB} tiene una distribución $F_{\alpha, (a-1)(b-1), ab(n-1)}$ (MONTGOMERY, 1991, cap. 7).

La estadística apropiada, para probar $H_0: \sigma_A^2 = 0$, es:

$$F_A = \frac{CM_A}{CM_{AB}}$$

que tiene una distribución $F_{\alpha, a-1, (a-1)(b-1)}$.

En forma similar para probar $H_0: \sigma_B^2 = 0$, la estadística debe ser:

$$F_B = \frac{CM_B}{CM_{AB}}$$

Y tiene una distribución $F_{\alpha, a-1, (a-1)(b-1)}$.

Con igual número de observaciones para cada combinación de tratamiento, se puede usar el análisis de varianza estudiado en la sección 4.5.1.2 para estimar los componentes de varianza. Esta estimación se hace igualando los cuadrados medios observados con los cuadrados medios esperados correspondientes y se despejan los valores desconocidos de los componentes. Las estimaciones de los cuatro componentes de varianza son (MONTGOMERY, 1991, cap. 7):

$$\text{Error: } \sigma^2 = CM_{Error}$$

$$\text{Interacción : } \sigma_{AB}^2 = \frac{CM_{AB} - CM_{Error}}{n}$$

$$\text{Factor A : } \sigma_A^2 = \frac{CM_A - CM_{AB}}{bn}$$

$$\text{Factor B : } \sigma_B^2 = \frac{CM_B - CM_{AB}}{an}$$

4.5.1.4 INFERENCIA PARA EL MODELO DE EFECTOS MIXTOS

Muchos experimentos se diseñan para estudiar los efectos de un factor sobre la media de la población y los efectos de otro sobre la varianza de la misma, estos experimentos tienen una mezcla de factores fijos y aleatorios. Los modelos para arreglos factoriales que incluyen

factores fijos y aleatorios se llaman *modelos mixtos*, pues contienen una mezcla de ambos tipos de efectos (KUEHL, 2001, cap. 7).

El modelo y análisis para los efectos mixtos se compone de dos partes porque hay dos tipos de inferencias. Las inferencias para los factores de efectos aleatorios se aplican a la variación en una población de efectos, mientras que las inferencias para los factores de efectos fijos están restringidas a los niveles específicos usados en el experimento (KUEHL, 2001, cap. 7).

El modelo lineal para este experimento es el mismo que el presentado en la ecuación 4.25, suponiendo en este caso que las variables aleatorias independientes y normalmente distribuidas con varianza constante, son además de ε_{ij} , los efectos de uno de los factores principales, por ejemplo τ_j^B , y el de la interacción τ_{ij}^{AB} (o bien en el caso más general, el de las interacciones que involucran al factor aleatorio).

Los cuadrados medios esperados para el análisis de varianza de efectos mixtos vienen dados como (KUEHL, 2001, cap. 7):

$$E(CM_A) = \sigma^2 + n\sigma_{AB}^2 + \frac{bn \sum_{i=1}^a (\tau_i^A)^2}{a-1}$$

$$E(CM_B) = \sigma^2 + n\sigma_{AB}^2 + an\sigma_B^2$$

$$E(CM_{AB}) = \sigma^2 + n\sigma_{AB}^2$$

$$E(CM_{Error}) = \sigma^2$$

Por lo tanto la estadística apropiada para probar $H_0 : \tau_i^A = 0$ es:

$$F_A = \frac{CM_A}{CM_{AB}}$$

que tiene una distribución $F_{\alpha, a-1, (a-1)(b-1)}$. Para probar $H_0 : \sigma_{AB}^2 = 0$, la estadística correspondiente es:

$$F_{AB} = \frac{CM_{AB}}{CM_{Error}}$$

cuya distribución es $F_{\alpha, (a-1)(b-1), ab(n-1)}$. Finalmente, para probar $H_0 : \sigma_B^2 = 0$, la estadística es:

$$F_B = \frac{CM_B}{CM_{AB}}$$

que tiene una distribución $F_{\alpha, b-1, (a-1)(b-1)}$

La estimación de los efectos del factor fijo (para el caso A), del modelo, viene dada por (KUEHL, 2001, cap. 7):

$$\begin{aligned}\mu &= \bar{y} \\ \tau_i^A &= \bar{y}_{i.} - \bar{y} \quad \text{para } i = 1, \dots, a\end{aligned}$$

Y las estimaciones para los componentes de varianza del factor aleatorio B y la interacción vienen dadas por (KUEHL, 2001, cap. 7):

$$\begin{aligned}\text{Interacción : } \sigma_{AB}^2 &= \frac{CM_{AB} - CM_{Error}}{n} \\ \text{Factor B : } \sigma_B^2 &= \frac{CM_B - CM_{AB}}{an}\end{aligned}$$

4.5.1.5 DETERMINACION DEL NUMERO DE REPLICAS

Para determinar el tamaño apropiado de la muestra (número de réplicas n) en los diseños factoriales; el procedimiento es similar a los explicados para el caso de los diseños unifactoriales (MONTGOMERY, 1991, cap. 7).

En la tabla 4.12, se muestra el valor apropiado del parámetro Φ^2 , para el caso del modelo bifactorial de efectos fijos.

σ^2 puede ser estimado a través de experiencias anteriores con diseños similares, una prueba piloto o estimaciones propuestas.

Φ^2 se puede estimar eligiendo valores de las medias de tratamientos para las cuales se debe rechazar H_0 , con una probabilidad alta. Otra forma es elegir un tamaño de muestra, de tal manera que la hipótesis nula se rechace si la diferencia entre cualquier par de medias de tratamientos excede un valor específico o si la diferencia entre dos medias es por lo menos D . Si este es el caso el valor mínimo de Φ^2 viene dado por (MONTGOMERY, 1991, cap. 7):

$$\Phi^2 = \frac{n' D^2}{2\sigma^2(v+1)}$$

Donde: n' es el número de unidades experimentales consideradas; D es la diferencia mínima para considerar que dos medias son diferentes; σ^2 es la varianza y v son los grados de libertad.

Así se tiene que si la diferencia entre dos medias de renglón es D , el valor mínimo de Φ^2 será (MONTGOMERY, 1991, cap. 7):

$$\Phi^2 = \frac{nbD^2}{2b\sigma^2}$$

Mientras que si la diferencia entre dos medias de columna es D el valor mínimo de Φ^2 será (MONTGOMERY, 1991, cap. 7):

$$\Phi^2 = \frac{naD^2}{2b\sigma^2}$$

Finalmente, el valor mínimo de Φ^2 que corresponde a una diferencia igual a D entre cualquier par de efectos de interacción será (MONTGOMERY, 1991, cap. 7):

$$\Phi^2 = \frac{nD^2}{2\sigma^2 [(a-1)(b-1) + 1]}$$

TABLA 4.12 PARAMETROS DE LA CURVA CARACTERÍSTICA DE OPERACIÓN MODELO BI FACTORIAL DE EFECTOS FIJOS (MONTGOMERY, 1991, cap. 7)

FACTOR	Φ^2	Grados de libertad del numerador	Grados de libertad del denominador
A	$\frac{bn \sum_{i=1}^a (\tau_i^A)^2}{a\sigma^2}$	a-1	ab(n-1)
B	$\frac{an \sum_{j=1}^b (\tau_j^B)^2}{b\sigma^2}$	b-1	ab(n-1)
AB	$\frac{n \sum_{i=1}^a \sum_{j=1}^b (\tau_{ij}^{AB})^2}{\sigma^2 [(a-1)(b-1) + 1]}$	(a-1)(b-1)	ab(n-1)

Para los modelos de efectos aleatorios y mixtos deben usarse las curvas de operación característica del anexo D4. El parámetro λ así como los grados de libertad del numerador y denominador para ambos modelos se muestran en las tablas 4.13 y 4.14 respectivamente.

TABLA 4.13 PARAMETROS DE LA CURVA CARACTERÍSTICA DE OPERACIÓN MODELO BI FACTORIAL DE EFECTOS ALEATORIOS (MONTGOMERY, 1991, cap. 7)

FACTOR	λ	Grados de libertad del numerador	Grados de libertad del denominador
A	$\sqrt{1 + \frac{bn \sigma_A^2}{\sigma^2 + n \sigma_{AB}^2}}$	a-1	(a-1)(b-1)
B	$\sqrt{1 + \frac{an \sigma_B^2}{\sigma^2 + n \sigma_{AB}^2}}$	b-1	(a-1)(b-1)
AB	$\sqrt{1 + \frac{n \sigma_{AB}^2}{\sigma^2}}$	(a-1)(b-1)	ab(n-1)

TABLA 4.14 PARAMETROS DE LA CURVA CARACTERÍSTICA DE OPERACIÓN MODELO BI FACTORIAL DE EFECTOS MIXTOS

FACTOR	Parámetro	Grados de libertad del numerador	Grados de libertad del denominador
A (fijo)	$\Phi^2 = \frac{bn \sum_{i=1}^a (\tau_i^A)^2}{a(\sigma^2 + \sigma_{AB}^2)}$	a-1	(a-1)(b-1)
B (aleatorio)	$\sqrt{1 + \frac{an \sigma_B^2}{\sigma^2}}$	b-1	(a-1)(b-1)
AB	$\sqrt{1 + \frac{n \sigma_{AB}^2}{\sigma^2}}$	(a-1)(b-1)	ab(n-1)

4.5.1.6 PRUEBA DE IDONEIDAD DEL MODELO

Las suposiciones para el modelo en cuanto a la homogeneidad de las varianzas y la distribución normal, se pueden evaluar con el *análisis de residuos* desarrollado en la sección 4.2.4. Los residuos para el diseño factorial con dos factores se calculan como la desviación de los valores observados de las medias estimadas para cada celda en el arreglo, el residual de cualquier celda es por tanto:

$$\varepsilon_{ij} = y_{ijk} - \hat{y}_{ijk} = y_{ijk} - \bar{y}_{ij}.$$

4.5.1.7 ANÁLISIS GRÁFICO DE LOS EFECTOS DE INTERACCIÓN

En las figuras 4.5 y 4.6, se ilustra de manera gráfica la ausencia o presencia de interacción para un arreglo con dos factores, A y B, cada uno con dos niveles; en la figura 4.5 se muestra la medición gráfica de un solo efecto de cada uno de los factores. La respuesta causada por A se grafica por separado para cada nivel de B, pero cuando no hay interacción las líneas son paralelas, mostrando que existe una respuesta idéntica al factor A en ambos niveles de B. En este caso, los factores son independientes y los efectos principales pueden usarse para interpretar por separado los efectos de ambos factores.

En la figura 4.6, se muestra la presencia de la interacción, entre los dos factores, aquí las líneas de respuesta no son paralelas. La diferencia en la magnitud (figura 4.6 a) o en la dirección (figura 4.6 b) de las respuestas representan la interacción entre los factores, estos no actúan de forma independiente y las interpretaciones deben basarse en los contrastes para cada combinación de los dos factores (KUEHL, 2001, cap. 6).

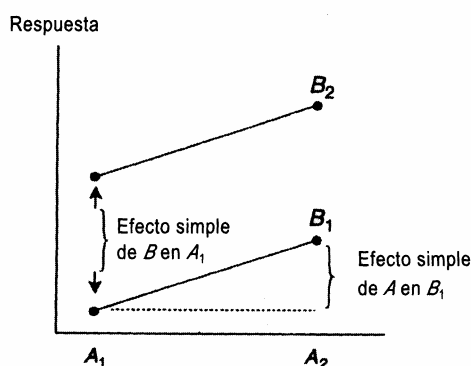


FIGURA 4.5 COMPORTAMIENTO DE NO INTERACCIÓN EN UN ARREGLO FACTORIAL

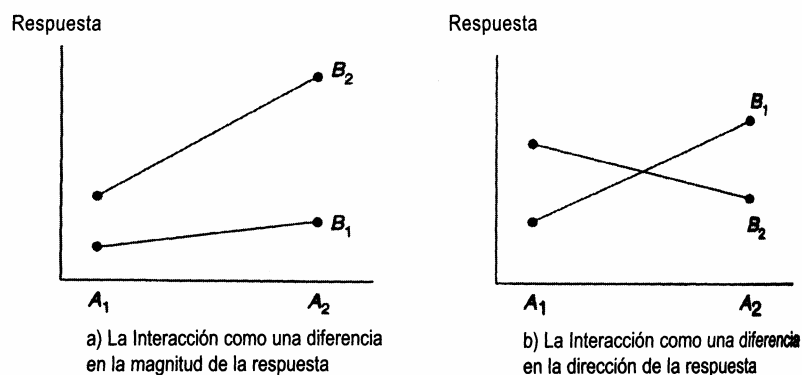


FIGURA 4.6 PRESENCIA DE EFECTOS DE INTERACCIÓN EN UN ARREGLO FACTORIAL

4.5.1.8 TEMAS GENERALES PARA LOS DISEÑOS FACTORIALES COMPLETOS

a) DISEÑO FACTORIAL GENERAL

Los resultados del diseño factorial de dos factores pueden extenderse al caso general en el que existen a niveles del factor A , b niveles del factor B , c niveles del factor C , y así sucesivamente, arreglados en un diseño factorial. En general habrá un total de $(abc...n)$ observaciones; si hay n réplicas por cada tratamiento. Hay que notar que deben recopilarse al menos dos réplicas para poder determinar la suma de cuadrados del error si todas las posibles interacciones en el modelo están incluidas (MONTGOMERY, 1991, cap. 7).

En el caso de que todos los factores del experimento sean de efectos fijos, fácilmente pueden formularse hipótesis acerca de los efectos principales y de las interacciones. Los estadísticos que prueban cada efecto principal y cada interacción se pueden construir dividiendo la media de cuadrados del efecto de las interacciones correspondientes entre las medias de cuadrados del error. Todas las pruebas F corresponden a pruebas unilaterales del extremo superior. Los grados de libertad de cualquier efecto principal son iguales al número de niveles del factor menos uno; y los grados de libertad de una interacción son el producto de los grados de libertad asociados con los componentes individuales de la interacción (MONTGOMERY, 1991, cap. 7).

b) UNA SOLA REPLICAS EN EL DISEÑO FACTORIAL COMPLETO:

En los estudios de investigación surgen situaciones en las que solo se dispone de una observación por celda del arreglo factorial. La varianza del error experimental no puede estimarse con una sola réplica de las combinaciones de los tratamientos, porque las particiones de las sumas de cuadrados para los efectos principales e interacción del factor son iguales a la suma de cuadrados total para las observaciones. Lo anterior implica que no se puedan efectuar contrastes de hipótesis sobre los efectos en estudio. Una forma de solventar esta situación es usando la asunción de la *aditividad* la cual describe que no hay interacción entre los factores. Bajo la aditividad de factores se puede usar la partición de los cuadrados medios de la interacción como una estimación del error experimental. Como la aditividad de los efectos principales o de la ausencia de interacción no esta garantizada, son necesarios algunos medios para evaluar la presencia de interacción (KUEHL, 2001, cap 6). Aun cuando existan algunos métodos para evaluar la presencia de interacciones que en general, serán función de que los factores que se combinan sean cualitativos y/o cuantitativos, lo más recomendable es prever desde el principio de la investigación los recursos disponibles y en función de ellos pensar en el uso de diseños experimentales alternativos, que aun con el uso de una sola réplica, permiten métodos de análisis más versátiles para el estudio de los efectos de los factores principales y de interacción, como por ejemplo pueden mencionarse los diseños 2^k (ver sección 4.6).

c) REPLICAS DESIGUALES EN LOS TRATAMIENTOS O DISEÑOS FACTORIALES DESBALANCEADOS:

Muchas veces puede darse el caso de datos faltantes en los estudios de investigación, situación que puede ocurrir por ejemplo, por la perdida imprevista de datos y la escasez de recursos para hacer repeticiones, o bien, porque el investigador crea a propósito diseños desbalanceados dado que algunas combinaciones por celda son más caras o difíciles de ensayar. Por otro lado, algunas combinaciones de tratamientos pueden ser de mayor interés para el experimentador porque representan condiciones nuevas o inesperadas (MONTGOMERY, 1991, cap. 7).

En general, sea cual sea la causa de datos faltantes, cuando se tienen diseños "desbalanceados" se pierde la propiedad de ortogonalidad de los efectos principales y de interacción, lo que implica que para estos diseños no puedan aplicarse las técnicas usuales de análisis de varianza.

Los métodos generales para el análisis de diseños factoriales desbalanceados pueden dividirse en (MONTGOMERY, 1991, cap. 7):

- i. *Métodos aproximados*: que son útiles cuando los datos no están muy desbalanceados para convertir este problema en uno balanceado. Para la aplicación de estos métodos aproximados se supone que cada celda tiene al menos una observación. Entre estos métodos están: *el método de la estimación de datos faltantes, el método de la eliminación de datos y el método de los promedios no ponderados* (una descripción detallada de estos métodos la presenta MONTGOMERY, 1991, sección 7.7.2).
- ii. *Métodos exactos*: En casos en los que los análisis aproximados se vuelven inapropiados, debido a que los datos están bien desbalanceados (como por ejemplo, presencia de celdas vacías o con números de réplicas bien diferentes), deben usarse *análisis exactos* que se basan en el desarrollo de métodos de partición del análisis de varianza, cuyo enfoque es representar el análisis de varianza mediante el ajuste de modelos de regresión. Como es lógico existen diferentes formas de hacer esto, y puede ser que cada método produzca resultados diferentes, sin embargo, las rutinas estadísticas generales, programadas para aceptar la teoría estadística conocida, han eliminado en parte, las diferencias asociadas con el análisis de conjuntos incompletos de datos. Una descripción detallada de la aplicación de un método exacto para el análisis de diseños desbalanceados se presenta en KUEHL, 2001, sección 6.9).

4.6 DISEÑOS FACTORIALES 2^k

El diseño 2^k es particularmente útil en las primeras etapas del trabajo experimental, cuando es probable que haya muchos factores por investigar. La difusión industrial de los diseños 2^k , se fundamenta en tres motivos (PRAT et al., 2000, cap. 7):

- i. Proporcionan una excelente relación entre el esfuerzo experimental y la información obtenida.
- ii. Son sencillos de construir, realizar, analizar e interpretar.
- iii. Son fáciles de combinar entre ellos para obtener otros diseños más complejos.

En estos diseños el plan experimental tiene por objeto el estudio del efecto sobre una respuesta, de k factores, cada uno a dos niveles. El modelo estadístico del diseño 2^k incluye el cálculo de $2^k - 1$ efectos que se distribuyen como: k efectos principales, $k(k-1)/2$ interacciones de dos factores, $k(k-1)(k-2)/3!$ interacciones de tres factores, ..., y una interacción de k factores (JURAN y GRZYNA, 1997, cap. 26). Los diseños 2^k , aun cuando constituyen una variante de los diseños factoriales, se estudian por separado debido a que pueden aplicarse técnicas de análisis diferentes al análisis de varianza.

Los diseños en los que cada factor varía a dos niveles sólo permiten estudiar relaciones lineales entre la respuesta y los diferentes factores, lo que puede considerarse como un grave inconveniente si este hecho no se compensa con la elección de niveles lo suficientemente cercanos, de tal manera que una línea recta sea una buena aproximación a la verdadera forma de la respuesta en la región de interés.

En general, las fases para un trabajo experimental que involucre un diseño factorial 2^k , son las mismas que para cualquier otro tipo de diseño, estas son: a) fase de diseño, b) fase de recolección, c) fase de estimación de efectos, d) fase de significación de efectos, y e) fase de evaluación o validación del modelo. En las secciones de la 4.6.1 a la 4.6.2, se presenta la descripción de la aplicación de cada una de estas fases.

4.6.1 MATRIZ DEL DISEÑO 2^k

En el diseño 2^k , los valores correspondientes a los dos niveles se codifican asignando al nivel bajo el valor de -1 (o simplemente -) y al alto +1 (o simplemente +). Si el factor es cualitativo, a un nivel se le asigna -1 y al otro +1 arbitrariamente. El arreglo de experimentos en un diseño con k factores, se puede generar fácilmente. En general, el número total de experimentos cuando se tienen k factores es igual a 2^k si no hay repeticiones. En problemas que tengan hasta cinco factores, la estructuración de todas las combinaciones posibles se puede elaborar empleando la tabla 4.15 como referencia. Una forma de recordar estos arreglos es observando que la primera columna alterna signos - y +, la segunda alterna pares de signos - y +, la tercera alterna cuartetos, etc. Obsérvese que el tamaño de los grupos de signos es una potencia de 2. Esto es, el tamaño en la primera columna equivale a 2^0 , en la segunda 2^1 , en la siguiente 2^2 , y así sucesivamente hasta 2^k grupos de signos. Cada fila de signos representa el conjunto de condiciones a que estará sujeto dicho experimento. Conviene señalar que en cada experimento se pueden medir diversas variables dependientes,

las cuales se analizan individualmente. La columna de signos de los factores de interacción se obtiene multiplicando los signos correspondientes en las columnas (factores) involucrados en la interacción, esta operación se realiza por hileras (LAWSON et al., 1992, cap. 2).

TABLA 4.15 ORDEN DE YATES DE EXPERIMENTOS 2^k (LAWSON et al., 1992, cap. 2)

	No.	X_1	X_2	X_3	X_4	X_5
k = 1	1	-	-	-	-	-
	2	+	-	-	-	-
k = 2	3	-	+	-	-	-
	4	+	+	-	-	-
k = 3	5	-	-	+	-	-
	6	+	-	+	-	-
	7	-	+	+	-	-
	8	+	+	+	-	-
k = 4	9	-	-	-	+	-
	10	+	-	-	+	-
	11	-	+	-	+	-
	12	+	+	-	+	-
	13	-	-	+	+	-
	14	+	-	+	+	-
	15	-	+	+	+	-
	16	+	+	+	+	-
	17	-	-	-	-	+
	18	+	-	-	-	+
	19	-	+	-	-	+
	20	+	+	-	-	+
	21	-	-	+	-	+
	22	+	-	+	-	+
	23	-	+	+	-	+
	24	+	+	+	-	+
25	-	-	-	+	+	
26	+	-	-	+	+	
27	-	+	-	+	+	
28	+	+	-	+	+	
29	-	-	+	+	+	
30	+	-	+	+	+	
k = 5	31	-	+	+	+	+
	32	+	+	+	+	+

4.6.2 RECOLECCION DE DATOS EN EL DISEÑO 2^k

El orden en que se presentaron los experimentos en la tabla 4.15 se conoce como orden estándar u orden de Yates. Este representa una forma conveniente de mostrar los resultados, pero a pesar de ello no es recomendable que los experimentos se ejecuten en ese orden, ya que se pueden producir sesgos. El orden de la experimentación debe ser totalmente

aleatorio, ya que de esta forma se evita o minimiza la influencia de los factores no controlables, al no permanecer constantes durante el experimento.

En ocasiones, el costo de cambiar un factor de nivel es muy elevado (esto ocurre con frecuencia cuando alguno de los factores y el coste energético y tiempo de calentamiento es muy elevado, así como el tiempo para el enfriamiento). Se puede recurrir a aleatorizaciones restringidas para minimizar el número de veces que se cambia el nivel de un factor. Sin embargo, siempre que sea posible y no se origine ningún problema serio, ni coste adicional importante en la realización del experimento, es conveniente aleatorizar totalmente el orden de la experimentación (LAWSON et al., 1992, cap. 2).

4.6.3 REPLICAS EN EL DISEÑO 2^k

Cuando se sabe que el sistema objeto de experimentación es muy variable, se hacen varios experimentos bajo cada condición experimental. A estas repeticiones se les denomina réplicas. Como se verá en la sección 4.6.4, al analizar los resultados se considera la media de estas réplicas como la única respuesta, y se procede a calcular los efectos como si sólo se hubiese experimentado una vez bajo cada condición y el resultado hubiese sido precisamente esa media. Conviene destacar que una réplica implica la realización de todo el experimento (bajo la condición señalada) y no recoger dos medidas de la misma respuesta, ambas consecuencias del mismo experimento. Las réplicas así consideradas tienen la ventaja de que permiten medir la variabilidad del sistema con el que se está experimentando y esta variabilidad puede ser utilizada para determinar la significación de los efectos de los factores. Tal como se verá en la sección 4.6.5 (PRAT et al., 2000, cap. 7).

4.6.4 CALCULO DE EFECTOS EN EL DISEÑO 2^k

Existen dos formas rápidas y sencillas de calcular los efectos principales y de interacción en un diseño 2^k :

- a. El algoritmo de los signos también llamado de BOX, HUNTER y HUNTER (1989).
- b. El algoritmo de Yates

A continuación se describe cada uno de estos métodos.

- a. Algoritmo de los signos: Este algoritmo calcula los efectos a partir de los signos utilizados para definir los niveles en la matriz de diseño, añadiendo las columnas de signos generadas para las interacciones (ver sección 4.6.1). Se añade además una columna con +1, que servirá para calcular la media. Cada efecto estimado es un estadístico de la forma $\bar{y}_+ - \bar{y}_-$, es decir esta expresado por la diferencia entre dos medias, cada una conteniendo 2^{k-1} observaciones. Otra forma de conceptualizar, la cuantificación de cualquiera de los efectos por este procedimiento, es mediante la suma algebraica de la respuesta, de acuerdo con los signos definidos en la columna del efecto correspondiente, y se divide por la mitad del número de condiciones experimentales. Para hallar la media, la operación es exactamente la misma que antes, excepto que se divide por el número de condiciones experimentales especificadas en la matriz de diseño, este valor es el promedio de todas las respuestas obtenidas durante la experimentación, pero resulta de escasa utilidad práctica.

En resumen, el cálculo de cualquier efecto viene dado por:

$$\text{Efecto} = \text{prom } y_+ - \text{prom } y_- = \bar{y}_+ - \bar{y}_-$$

En donde: $\text{prom } y_+$ es el promedio aritmético de todas las respuestas con signo positivo cuando la columna de los resultados obtenidos en la experimentación se multiplica por el vector columna de los signos del efecto L y $\text{prom } y_-$ es el promedio de todas las respuestas con signo negativo cuando la columna de los resultados obtenidos en la experimentación se multiplica por el vector columna de los signos del efecto L (JURAN y GRAYNA, 1997, cap. 26).

- b. Algoritmo de Yates: Este procedimiento requiere un menor número de operaciones que el anterior y, al mismo tiempo, es el más adecuado para implementarlo en una hoja de cálculo o cualquier otro programa informático. Para la aplicación de este algoritmo de Yates se procede como sigue (JURAN y GRAYNA, 1997, cap. 26):

1. Listar los datos en el orden de Yates o estándar.

2. Crear $k+2$ columnas, en la primera columna colocar el diseño experimental, es decir, las combinaciones de los 2^k tratamientos en el orden estándar de Yates (ver tabla 4.15).
3. En la segunda columna, colocar las observaciones o respuestas a cada combinación de tratamientos. Si el plan es replicado indicar el total de cada combinación.
4. En la primera mitad de la tercera columna, colocar las sumas de los pares consecutivos de datos correspondientes, es decir el primer valor será la suma del 1º y 2º dato, el segundo valor la suma del 3º y 4º dato y así sucesivamente. En la segunda mitad indicar por orden las diferencias de los pares consecutivos de datos (primer valor resta del 2º menos el 1º dato, segundo valor resta del 4º menos el 3º dato, y así sucesivamente).
5. Calcular las columnas 4, 5, ..., $k+2$ de la misma forma que la columna 3, es decir, obteniendo en cada caso las sumas y diferencias de los pares de datos de la columna precedente.
6. Los datos de la columna $k+2$, se designan con los códigos $g(T)$, $g(A)$, $g(AB)$, etc. Las letras entre paréntesis corresponden a los signos más de la notación geométrica, es decir, de la combinación de tratamientos de la fila correspondiente del diseño experimental. El primer valor $g(T)$ se divide entre N para obtener el promedio global. Las estimaciones de los efectos principales y de interacción se obtienen dividiendo cada valor $g(\dots)$ por $N/2$.

(Las siguientes etapas son controles)

7. La suma de las respuestas individuales en la segunda columna, ha de ser igual al primer dato de la columna de las $g(\dots)$, ya que $g(T)$ es el total general.
8. La suma de los cuadrados de las cantidades de la columna dos, ha de ser igual a la suma de los cuadrados de los datos de la columna $k+2$, dividida entre 2^k .
9. Cada $g(\dots)$, de la última columna ha de ser igual a la sumatoria de las observaciones con signo más (+) menos la sumatoria de las observaciones con signo menos (-).

4.6.5 SIGNIFICACION DE EFECTOS EN EL DISEÑO 2^k

Una vez calculados los efectos, la primera tarea, mediante la técnica estadística más adecuada, será distinguir cuáles de ellos son significativamente diferentes de cero⁴ y cuáles no lo son. Hay dos situaciones de partida distintas al abordar este problema: cuando se ha replicado el experimento, y cuando cada experimento elemental se ha llevado a cabo una sola vez (por razones obvias de economía experimental, la segunda situación es la más habitual).

Existen varias formas estadísticamente válidas para evaluar la significación de los efectos en un diseño 2^k , entre ellas están:

- a) La prueba t-Student de significación de los efectos
- b) La prueba de análisis de varianza
- c) Pruebas gráficas de significación de efectos

A continuación, en las secciones de la 4.6.5.1 a la 4.6.5.2, se presenta una descripción de cada una de ellas.

4.6.5.1 PRUEBA T-STUDENT DE SIGNIFICACIÓN DE EFECTOS

Para poder determinar la significación de los efectos por la prueba de t-Student, es necesario disponer de una medida de la variación que el error experimental ha inducido en la estimación de los mismos. El cálculo de dicha medida de variabilidad depende si el diseño es replicado o no.

Si el diseño es replicado y no se tiene igual número de réplicas en las distintas combinaciones de tratamientos, la varianza del error puede calcularse a partir de (PRAT et al., 2000, cap. 7):

$$s_r^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_c - 1)s_c^2}{n_1 + n_2 + \dots + n_c - C}$$

⁴ Que la estimación de un efecto, hallada a través del experimento, sea distinta de cero no implica que el verdadero valor del mismo lo sea. Es decir, no implica que afecte en forma detectable a la respuesta y que, por tanto, corresponda a un término que deba ser incluido en el modelo.

con $n_1 + n_2 + \dots + n_c - C$ grados de libertad, donde C es el número de condiciones experimentales distintas ($C = 2^k$) y n_i el número de réplicas de la i -ésima condición experimental.

Si se tiene una sola réplica para cada combinación, la varianza del error experimental puede estimarse utilizando la suma de los cuadrados correspondientes a las interacciones de tres o más factores. La suma de los cuadrados con un grado de libertad correspondiente al efecto factorial de dos niveles viene dada por la fórmula (JURAN y GRZYNA, 1997, cap. 26):

$$\frac{N \sum_{i=1}^e (\text{efecto de interacciones de orden tres o mayor})^2}{4}$$

Para obtener la estimación de la varianza del error experimental, debe dividirse la suma de cuadrados de la ecuación anterior entre e , siendo e el número de los efectos de interacción que contribuyen al total. En un factorial 2^k , el número de interacciones de orden tres o mayor serán $2^k - (k^2 + k + 2)/2$.

Finalmente, si no se dispone de réplicas de las distintas combinaciones, pero el diseño es aumentado con puntos centrales (ver sección 4.6.9), la varianza del error experimental puede estimarse como la varianza de las "q" repeticiones de los puntos centrales. Los grados de libertad en este caso son $n_i - 1$, en donde n_i es el número de repeticiones de los puntos centrales.

A continuación se describe el procedimiento general para la prueba t-Student de significación de los efectos (JURAN y GRZYNA, 1997, cap. 26):

1. Elegir el nivel de significación α .
2. Estimar la varianza del error experimental (de acuerdo a lo anteriormente expuesto).
3. Buscar $t_{\alpha/2}$ para los grados de libertad correspondientes, según sea la estimación de la varianza del error experimental.
4. El error tipo de cada efecto estimado viene dado por:

$$ET(\text{efecto estimado}) = ET(\bar{y}_+ - \bar{y}_-) = \frac{2s}{\sqrt{N/2}}$$

en donde N es el número total de observaciones.

5. Los límites de confianza para el $100(1-\alpha)\%$ de cada efecto son por tanto:

$$\text{Efecto} \pm t_{\alpha/2} [ET(\text{efecto estimado})]$$

6. Para cualquier estimación de los efectos principales o de interacción, si el valor absoluto de uno de los efectos es mayor que $t_{\alpha/2} [ET(\text{efecto estimado})]$, concluir que la diferencia del verdadero efecto a partir de cero es estadísticamente, significativa. Otra forma de evaluar la significación de los efectos es a través del intervalo de confianza creado para cada uno de ellos, ya que puede concluirse que el efecto es significativo si el intervalo no contiene al valor de cero.

En la práctica por razones de comodidad y rapidez, se recurre al cálculo de intervalos de confianza aproximados, en la forma (PRAT et al., 2000, cap.

7) :

$$\text{efecto} \pm 2 \text{ desviacion es tipo de los efectos}$$

Y se consideran significativos aquellos cuyo intervalo no contiene al cero.

4.6.5.2 ANÁLISIS DE VARIANZA PARA LA SIGNIFICACIÓN DE EFECTOS

En los diseños 2^k replicados la prueba de significación de los efectos puede realizarse también a partir de análisis de varianza, lo que implica estimar y calcular las sumas de cuadrados para cada uno de ellos. En general, las sumas de cuadrados de los efectos vendrán dadas, por (JURAN y GRAYNA, 1997, cap. 26):

$$SC_{\text{Efecto}} = \frac{1}{N} (y_+ - y_-)^2 = \frac{N}{4} (\bar{y}_+ - \bar{y}_-)^2$$

En donde N es el número total de observaciones que para un diseño con n réplicas viene dado por $n \times 2^k$.

En la tabla 4.16, se presenta el análisis de varianza para el diseño 2^k . Obsérvese que la suma de cuadrados medios se excluye de ésta tabla ya que la suma de cuadrados de los efectos principales y de interacción son iguales, a excepción del cuadrado medio del error $CM_{\text{Error}} = SC_{\text{Error}} / (2^k(n-1))$ (MONTGOMERY, 1991, cap. 9).

Las razones SCA / CM_{Error} , SCB / CM_{Error} , etc., tienen una distribución F con un grado de libertad en el numerador y $2^k(n-1)$ grados de libertad en el denominador. Las regiones críticas corresponden al extremo superior de la distribución F. En otras palabras, la hipótesis

nula para un nivel de significación α se rechaza cuando $F_{\text{factor}} > F_{\alpha, (1, 2^k (n-1))}$; donde F_{factor} es el F calculado para cada factor (MONTGOMERY, 1991, cap. 9).

Si se tiene una sola réplica el análisis de varianza puede realizarse tomando como estimación de la varianza del error la presentada para la prueba t de una sola réplica de la sección 4.6.5.1.

4.6.5.3 PRUEBAS GRAFICAS DE SIGNIFICACION DE EFECTOS

En el trabajo experimental, muchas veces no se tendrá la disponibilidad de realizar replicas. Sin embargo, existen métodos sencillos de tipo gráfico que permiten detectar la presencia de efectos significativos de forma eficaz, dos de los cuales son: (a) Los gráficos en escala probabilística normal de los efectos, y (b) el gráfico de Pareto. A continuación se describe cada uno de estos métodos.

(a) GRAFICOS DE LOS EFECTOS EN ESCALA PROBABILISTICA NORMAL:

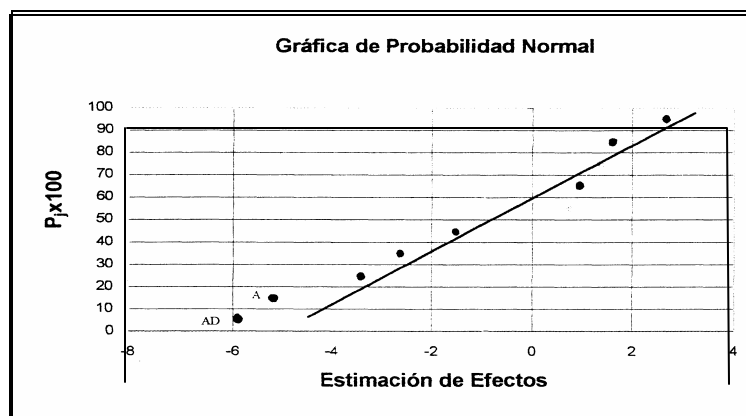
La respuesta experimental esta sometida a variación aleatoria. Esta variación seguirá, en general, una ley normal con media cero y una desviación típica que refleja el error experimental. Aún el caso infrecuente en que no se sigue la ley normal, los efectos son combinaciones lineales de la respuesta, y por el *teorema central del límite* tendrán tendencia a seguirla (PRAT et al., 2000, cap. 7).

Supóngase que se tiene un experimento en el que todos los efectos sean no significativos, los efectos hallados oscilan alrededor de cero, siguiendo una ley normal ($N(0, \sigma_{\text{efectos}})$). Al representarlos en una escala probabilística normal aparecerán todos ellos alineados, la pendiente de la recta es una indicación del tamaño del error experimental. Una recta más horizontal indicaría una mayor variabilidad (PRAT et al., 2000, cap. 7).

Por lo contrario, cuando al menos uno de los efectos es significativo, si bien tiene la misma desviación tipo, la media varía dependiendo del tamaño del efecto. De forma que cada efecto se distribuye como $N(\mu_{\text{efecto}}, \sigma_{\text{efectos}})$, donde μ_{efecto} es distinta en cada caso. En la figura 4.7, se presenta un ejemplo del comportamiento de los efectos en escala probabilística normal (PRAT et al., 2000, cap. 7).

TABLA 4.16 ANÁLISIS DE VARIANZA PARA EL DISEÑO 2^k (MONTGOMERY, 1991, cap. 9)

Fuente de variación		Suma de Cuadrados	Grados de libertad
K efectos principales			
	A	SCA	1
	B	SCB	1
	·	·	·
	·	·	·
	K	SCK	1
$\binom{k}{2}$	Interacción de 2 Factores		
	AB	SCAB	1
	AC	SCAC	1
	·	·	·
	·	·	·
	JK	SCJK	1
$\binom{k}{3}$	Interacción de 3 Factores		
	ABC	SCABC	1
	ABD	SCABD	1
	·	·	·
	·	·	·
	IJK	SCI JK	1
$\binom{k}{k}$	Interacción de k Factores		
	ABC...K	SCABC...K	1
	Error	SC _{Error}	$2^k(n - 1)$
	Total	SC _{Total}	$n2^k - 1$


FIGURA 4.7 GRAFICA DE PROBABILIDAD NORMAL DE LOS EFECTOS

Las gráficas en escala probabilística normal de los efectos se construyen de la siguiente forma (LAWSON et. al., 1992, cap. 1):

1. Se ordenan las estimaciones de los efectos en sentido creciente (de menor a mayor).
2. Se crean percentiles $p_i = (i - 0.5)/(\text{número de efectos})$ para cada uno de los efectos ordenados, en donde i indica el orden del efecto.
3. Se calculan los estadísticos de orden z asociados a cada p_i (z-score), obtenidos de la tabla del anexo A1. Si se desea puede utilizarse la aproximación:

$$z = 5.05 \left[p_i^{0.135} - (1 - p_i)^{0.135} \right]$$

4. Representar gráficamente las parejas (efecto, $p \times 100$) o bien (orden de efecto, z score).
5. Interpretar. Requiere cierta práctica trazar correctamente las rectas e interpretar los resultados. Es obvio que la recta debe ajustarse a los puntos centrales y no a los de los extremos, que son los efectos potencialmente significativos. Un error común es considerar como significativos efectos que se apartan de la recta por estar demasiado próximos a cero (esto ocurre con cierta frecuencia, ya que al extraer muestras de tamaño reducido de una ley normal los extremos tienden a estar sobrepresentados). Ver figura 4.8.

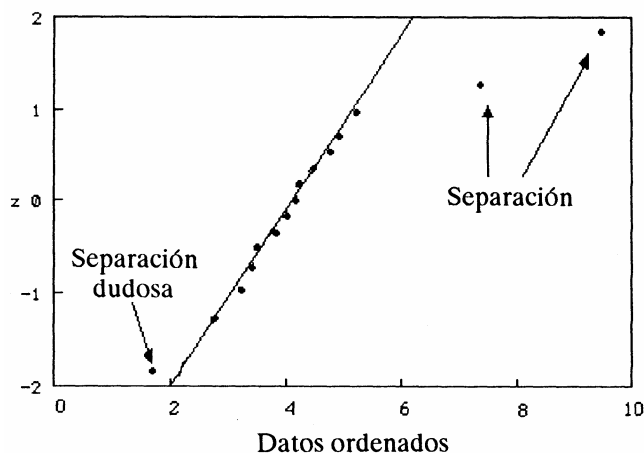


FIGURA 4.8 INTERPRETACION DE SIGNIFICACIÓN DE LOS EFECTOS

(b) GRAFICOS DE PARETO DE LOS EFECTOS:

El análisis de Pareto constituye otro recurso gráfico muy útil para representar la significación de los efectos y consiste en elaborar un histograma ordenado en el que la longitud de cada

rectángulo representa la magnitud de los efectos estimados. El análisis de significación consiste en comparar la altura de cada rectángulo con una línea trazada sobre el valor del efecto normal calculado con un nivel de significación α , para el σ_{efectos} correspondiente, es decir: $z_{\alpha} = \mu_{\text{efecto}}/\sigma_{\text{efectos}}$ en donde μ_{efecto} es el promedio de los valores de los efectos obtenidos. En la figura 4.9 se presenta un análisis de Pareto para la significación de efectos.

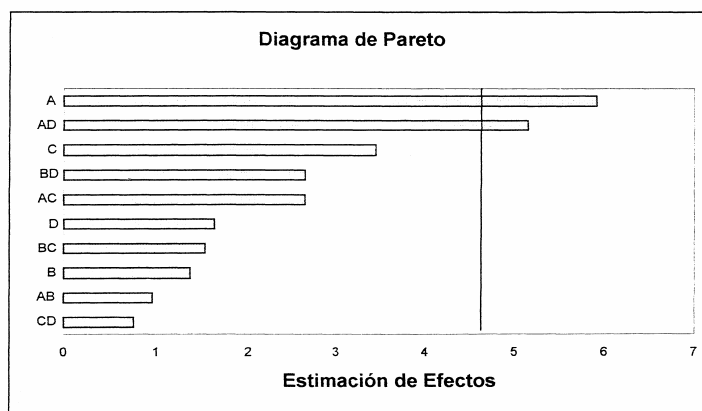


FIGURA 4.9 GRAFICO DE PARETO PARA EL ANALISIS DE SIGNIFICACIÓN DE LOS EFECTOS

4.6.6 INTERPRETACION DE RESULTADOS

Una vez que se ha determinado cuales de los efectos son significativos un paso importante es interpretar físicamente esos resultados. Una forma de interpretarlos es observando si los efectos producen un aumento o decremento en la respuesta. Un recurso bastante útil consiste en elaborar gráficos de los niveles de los efectos con respecto a las respuestas obtenidas tanto para los efectos principales, como para los de interacción (ver figura 4.6). Por supuesto es imprescindible hacer estas interpretaciones a la luz de los conocimientos previos que sobre el problema se pudiesen tener.

Por otra parte, si lo que se quiere lograr con la experimentación es construir un modelo, la interpretación de los resultados puede venir enfocada sobre la búsqueda de una función que represente aproximadamente la relación entre la respuesta y los factores, que resulte útil en la zona experimental (es importante recalcar que no es conveniente extrapolar variables más allá de los niveles que se han experimentado).

Mediante diseños 2^k se pueden estimar los coeficientes de modelos polinómicos sin términos cuadráticos. El modelo a estimar es del tipo:

$$y = \beta_0 + \sum_{i=1}^k \beta_{ii} x_i + \sum_i \sum_j \beta_{ij} x_i x_j + \dots + (\text{hasta } 2^k - (k^2 + k + 2)/2 \text{ interacciones de orden tres o mayor})$$

donde:

- β_0 Término independiente del modelo. Corresponde a la media.
- $\beta_{1...}, \beta_{123}...$ Coeficientes. Son la mitad de los efectos calculados. Esto es debido a que el efecto principal de un factor representa el cambio en la respuesta al pasar del nivel bajo (-) al alto (+). Mientras que el coeficiente representa el cambio en la respuesta al cambiar el factor una unidad (del 0 al 1, o del -1 al 0).
- x_i, x_j, \dots Valores que toma cada uno de los factores.
- $x_i x_j, x_i x_j x_l$ Productos de los valores que toman los factores que se indican. En el modelo representan los efectos de las interacciones.
- y Valor que toma la respuesta para los valores dados de los factores.

En el modelo sólo deben incluirse los términos correspondientes a aquellos factores que resultaron significativos. La superficie creada en función de los factores codificados permite interpretar los efectos de los factores. La decodificación de las variables puede lograrse a partir de:

$$X_c = \frac{X_o - \frac{\text{nivel sup.} + \text{nivel inf.}}{2}}{\frac{\text{nivel sup.} + \text{nivel inf.}}{2}}$$

Donde el subíndice C indica unidades codificadas y el O originales. Las ecuaciones con las variables en las unidades originales deben ser tratadas con cautela, ya que su interpretación es más compleja. Son útiles para predecir, pero no para interpretar los efectos de los factores (PRAT et al., 2000, cap. 7).

4.6.7 DIAGNOSIS DEL MODELO

Es fácil calcular los residuos de un diseño 2^k por medio del modelo de regresión creado. En general los residuos vienen dados por:

$$\varepsilon = y - \hat{y}$$

En donde \hat{y} es el valor previsto para cada condición experimental. Los residuos calculados permiten comprobar si se cumplen las hipótesis del modelo, que son *independencia*, *normalidad* y *varianza constante*, y que pueden ser comprobadas mediante las técnicas gráficas para el análisis de residuos presentadas en la sección 4.2.4.

4.6.8 CALCULO DEL NUMERO DE REPLICAS EN EL DISEÑO 2^k

Como se vio en la sección 4.6.3, las repeticiones son necesarias para lograr la precisión deseada cuando el sistema es muy variable. Si δ identifica la magnitud máxima de la diferencia práctica⁵ entre las medias muestrales que se está dispuesto a aceptar, y σ es la desviación estándar, el número de experimentos necesarios para lograr una precisión deseada es (LAWSON et al., 1992, cap. 2):

$$N = \left(\frac{8\sigma}{\delta} \right)^2$$

Por ejemplo, si se requiere un 2^4 y en función de la precisión el $N = 32$, el experimento debe replicarse. Si N no es un número entero el valor debe aproximarse al múltiplo más cercano a 2^k ($r2^k$). Lo anterior garantizará que el error máximo sea menor al 10% y permitirá probar con un 95% de confianza, la significación de los efectos (LAWSON et al., 1992, cap. 2).

A medida que δ disminuye (mayor precisión) el número de experimentos necesarios (N), será mayor. A menudo no es necesario que δ sea mucho menor que σ (a menos que los experimentos sean baratos). La precisión será la misma independientemente del tamaño del factorial. La precisión depende del número de combinaciones totales ejecutadas y no del número de factores (LAWSON et al., 1992, cap. 2).

Al iniciar un programa experimental es recomendable empezar determinando N , ya que es útil para determinar el costo del programa experimental, con la precisión deseada. En la práctica es útil definir a $\delta = 8\sigma/\sqrt{n}$, ya que así se indica la precisión que es posible comprar con N datos (presupuesto disponible para el estudio). Si δ no es adecuado se debe proponer

⁵ La diferencia práctica puede tener una significación estadística o empírica, se dice que es estadísticamente significativa si su valor es muy grande con respecto a la media de cuadrados del error. Por otra parte, es empíricamente significativa cuando esta diferencia es lo suficientemente grande como para producir beneficios económicos o en la operación del proceso a largo plazo. Se pueden tener estudios en los que se encontrará significación estadística pero no práctica o viceversa.

aumentar N ó no ejecutar el programa (o tratar de disminuir la desviación estándar σ) (LAWSON et al., 1992, cap. 2).

4.6.9 ADICION DE PUNTOS CENTRALES EN EL DISEÑO 2^k

Los puntos centrales son respuestas obtenidas cuando todos los factores experimentales se encuentran a un nivel intermedio a los niveles inferior (-) y superior (+) del diseño 2^k . La repetición del experimento en los puntos centrales es importante en los factoriales a dos niveles, debido a dos razones (LAWSON et al., 1992, cap. 2):

- Proporcionan un estimador del error, con base en el cual se pueden medir los efectos: La repetición de puntos centrales permite determinar un estimador "puro" del error ya que son puntos repetidos.
- Producen una prueba para el modelo: Un factorial se basa en que los cambios en la variable dependiente, se pueden estimar bien con un modelo lineal. La validez de dicha hipótesis no se puede verificar ya que sólo se tienen dos niveles para cada factor. En el caso de existir efectos no lineales (curvatura), las predicciones no son buenas excepto en los puntos cercanos a los puntos factoriales. El uso de puntos centrales permite determinar la no linealidad analizando la magnitud de la diferencia entre los promedios del punto central y los puntos factoriales (LAWSON et al., 1992, cap. 2):

$$\text{Efecto de curvatura} = C = \bar{y}_{\text{factorial}} - \bar{y}_{\text{puntoscentrales}}$$

En donde, $\bar{y}_{\text{factorial}}$ es el promedio de las respuestas de las $2^k \times n$ corridas factoriales y $\bar{y}_{\text{puntoscentrales}}$ es el promedio de las respuestas asociadas a los puntos centrales.

El análisis de significación del efecto de curvatura puede lograrse mediante la prueba t para el estadístico (LAWSON et al., 1992, cap. 2):

$$t_c = \frac{C}{S_c} \text{ en donde } S_c \text{ viene dado por } S_c = s \sqrt{\frac{1}{N_f} + \frac{1}{n_{p.c}}}$$

(en el cual s es la estimación de la desviación estándar de y , N_f es el número de puntos factoriales y $n_{p.c.}$ es el número de puntos centrales).

El efecto de curvatura resulta significativo si t_{ϵ} es mayor que el valor de una t de Student con v grados de libertad, en donde v son los grados de libertad con los que se ha estimado σ a partir de s (PRAT et al., 2000, cap. 9).

4.6.10 BLOQUES EN DISEÑOS FACTORIALES 2^k

El experimento factorial puede sesgarse sustancialmente, si existe la posibilidad que los niveles de ciertas variables latentes (factores no controlables) cambien, es decir, si hay motivos que provoquen que el experimento no se lleve a cabo bajo condiciones homogéneas. Al bloquear es posible dividir el global de experimentos en grupos o bloques dentro de los cuales se piensa que las condiciones bajo las que se va a experimentar son homogéneas. Para conseguir esas divisiones, se confunden interacciones, de las que en principio se piensa que no son importantes (en general, tres o más factores), con los factores de bloqueo. Los principios para crear variables de bloque son los que se aplican a los diseños factoriales fraccionarios, que pueden estudiarse con más detalle en PRAT et. al.(2000).

4.7 PROBLEMAS DE DISEÑO DE EXPERIMENTOS

EJEMPLO 4.7.1 DISEÑO UNIFACTORIAL DE EFECTOS FIJOS

Se quiere comparar el trabajo de los cuatro analistas de un laboratorio en el ensayo de determinación del porcentaje de alcohol metílico en muestras de un producto químico mediante la técnica de cromatografía líquida de alta resolución (HPLC). Cada uno efectúa tres determinaciones. Para asegurar la distribución aleatoria de las muestras se procedió de la siguiente forma:

Los analistas se enumeraron del 1 al 4 y las muestras del producto químico del 1 al 12, seguidamente se tomaron al azar un analista y una muestra, hasta completar la distribución. La distribución resultante de los análisis se presenta a continuación:

Analista	Muestra		
	1	9	1
2	3	10	12
3	4	6	8
4	5	2	7

Los resultados reportados por los analistas fueron los siguientes:

Analista	Porcentaje de alcohol metílico		
	1	84.99	84.38
2	85.15	84.88	85.13
3	84.72	85.15	84.48
4	84.20	84.55	84.10
Total			

En función de los resultados obtenidos determine si el trabajo de los analistas difiere significativamente. Use un nivel de significación (α) de 0.05.

Solución:

Objetivo del problema: Comparar el trabajo de los cuatro analistas de un laboratorio, en el análisis de metanol en un producto químico por HPLC, como parte del proceso de aseguramiento de la calidad del laboratorio.

1. El modelo estadístico lineal para el análisis de datos en el modelo unifactorial de efectos fijos es el siguiente:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

2. Contraste de hipótesis:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_1: \mu_i \neq \mu_j \quad (\text{en al menos un par})$$

3. Criterio de decisión:

$$F_o = \frac{SC_{\text{Tratamientos}} / (k - 1)}{SC_{\text{Error}} / (N - k)} = \frac{MC_{\text{Tratamientos}}}{MC_{\text{Error}}}$$

Debe rechazarse H_0 si $F_o > F_{\alpha, k-1, N-k}$

4. Si se rechaza H_0 , estimar los efectos.

5. Cálculos:

- En la tabla E4.1.1 se presentan los totales y los promedios obtenidos de los resultados de cada analista:

TABLA E4.1.1

Analista	Porcentaje de alcohol metílico			Total	Promedio
1	84.99	84.38	84.02	253.39	84.46
2	85.15	84.88	85.13	255.16	85.05
3	84.72	85.15	84.48	254.36	84.79
4	84.20	84.55	84.10	252.85	84.28
Total				1015.76	84.65

- Suma de cuadrados de tratamientos $SC_{\text{Tratamientos}}$

$$SC_{\text{Tratamientos}} = \sum_{i=1}^k \frac{y_i^2}{n_i} - \frac{Y^2}{N} = \frac{(253.39)^2 + (255.16)^2 + (254.36)^2 + (252.85)^2}{3} - \frac{(1015.76)^2}{N}$$

$$SC_{\text{Tratamientos}} = 8598175 - 8598070 = 1.05$$

- Suma de cuadrados del error y suma de cuadrados total (SC_{Error} y SC_{Total})

$$SC_{\text{Error}} = SC_{\text{Total}} - SC_{\text{Tratamientos}}$$

$$SC_{\text{Total}} = \sum_{i=1}^4 \sum_{j=1}^3 y_{ij}^2 - \frac{Y^2}{N} = 85982.62 - 85980.7 = 1.92$$

$$SC_{\text{Error}} = 1.92 - 1.05 = 0.87$$

- En la tabla E.4.1.2 se presenta el análisis de la varianza:

TABLA E4.1.2

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	F_0
Entre tratamientos	1.05	3	0.35	3.21
Dentro tratamientos	0.87	8	0.109	
Total	1.92	11		

$$F_{0.05,3,8} = 4.06 > F_0$$

Conclusión:

No existe evidencia significativa al 5% como para asegurar que los analistas del laboratorio trabajan diferente.

Prueba de idoneidad del modelo

En general, los residuos para el modelo unifactorial vienen dados por:

$$\varepsilon_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \bar{y}_i$$

A continuación se presenta el cálculo de los residuos para el ejemplo 4.1.

Las estimaciones de los y_j vienen dadas por:

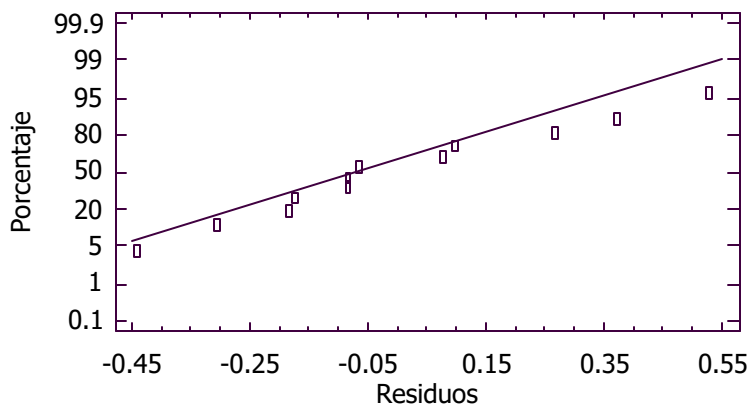
Analista	1	2	3	4	Media
$\mu_i = \bar{y}_i = \hat{y}_{ij}$	84.4633	85.0533	84.7867	84.2833	84.6467

Por tanto, los residuos son:

Analista	1	1	1	2	2	2
y_j	84.99	84.02	84.38	85.15	85.13	84.88
$\varepsilon_{ij} = y_{ij} - \hat{y}_{ij}$	0.5267	-0.4433	-0.0833	0.0967	0.0767	-0.1733
Analista	3	3	3	4	4	4
y_j	84.72	84.48	85.16	84.2	84.1	84.55
$\varepsilon_{ij} = y_{ij} - \hat{y}_{ij}$	-0.0667	-0.3067	0.3733	-0.0833	-0.1833	0.2667

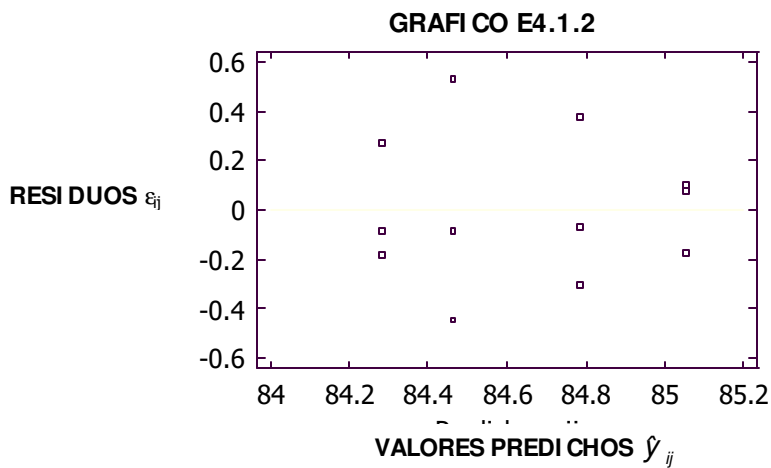
- *Prueba de normalidad de los residuos:* según el gráfico E4.1.1, para esta prueba la variabilidad de los residuos puede considerarse normal, por lo que se cumplen hipótesis, que justifican el análisis de la varianza.

GRAFICO E4.1.1



Por otra parte, puede considerarse a través de una prueba de hipótesis que la media de los residuos no es significativamente diferente de cero a un nivel de significación del 5%.

- *Prueba de homogeneidad de la varianza de los residuos:* Según el gráfico E4.1.2, hay una dispersión aleatoria de los puntos, por lo tanto puede decirse que existe homocedasticidad.



- Según la prueba de Barlett, puede considerarse que la varianza para cada analista es la misma.

Prueba de Barlett:

Las hipótesis a probar son:

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_k^2$$

$$H_1: \text{lo anterior no es cierto al menos para una } \sigma_j^2$$

La estadística de prueba (ver sección 4.2.4b) es: $\chi_o^2 = 2.3026 \frac{q}{c}$

donde:

$$q = (N - k) \log_{10} S_P^2 - \sum_{i=1}^k (n_i - 1) \log_{10} S_i^2$$

$$c = 1 + \frac{1}{3(k-1)} \left(\sum_{i=1}^k (n_i - 1)^{-1} - (N - k)^{-1} \right)$$

$$S_P^2 = \frac{\sum_{i=1}^k (n_i - 1) S_i^2}{N - k}$$

y S_i^2 es la varianza muestral de la i -ésima población.

Las varianzas muestrales (S_i^2) para cada analista, son las siguientes:

Analista	Varianza muestral
1	0.240433
2	0.0226333
3	0.118933
4	0.0558333

En donde cada S_i^2 viene dada por: $S_i^2 = \sum_{j=1}^k \frac{(y_{ji} - \bar{y}_i)^2}{n_i - 1}$

Por tanto:

$$S_p^2 = \frac{2(0.240433) + 2(0.0226333) + 2(0.118933) + 2(0.0558333)}{8}$$

$$c = 1 + \frac{1}{3(3)} \left(\frac{4}{2} - \frac{1}{8} \right) = 1.20833$$

$$q = 8 \log_{10}(0.10945815) - [2 \log_{10}(0.240433) + \dots + 2 \log_{10}(0.0558333)] = 1.198109576$$

$$\chi_o^2 = 2.3026 \frac{q}{c} = 2.3026 \frac{1.19811}{1.20833} = 2.283$$

Dado que $\chi_{\alpha, k-1}^2 = \chi_{0.05, 3}^2 = 7.815$, y es mayor que $\chi_o^2 = 2.283$, se concluye mediante la prueba de Barlett que existe homogeneidad de las varianzas entre los tratamientos, para un nivel de significancia del 5%.

Cálculo del tamaño del número de repeticiones por tratamiento (cálculo del tamaño de muestra):

Como no existe evidencia significativa al 5% (ni al 10%), como para asegurar que los analistas del laboratorio trabajan diferente, vale la pena analizar el número de repeticiones por tratamiento y la potencia de la prueba, adecuados a esta investigación.

Los valores de las medias y de σ^2 , son:

Analista	Media muestral μ_i
1	84.4633
2	85.0533
3	84.7867
4	84.2833
Media(μ)	84.6467

$$S^2 = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \mu)^2}{N - 1} = 0.1752$$

Utilizando las tablas de potencia del anexo D4 y tomando $\alpha = 0.01$, se procede a calcular el número de mediciones más apropiado por analista.

Los resultados se muestran a continuación:

n	$v_1 = k - 1$	$v_2 = N - k$	$\Phi^2 = 0.5n$	Φ	$1 - \beta$
3	3	8	1.5	1.22	0.10
5	3	16	2.5	1.58	0.35
10	3	36	5	2.24	0.86
11	3	40	5.5	2.34	0.90

Por lo tanto, para obtener resultados más concluyentes sobre si existe o no diferencia en el trabajo de los analistas, es recomendable realizar un nuevo experimento en el que el número de réplicas por tratamiento sea de 11, ya que sólo así se logra satisfacer el requisito mínimo de potencia de la prueba que es de 0.9, como puede observarse en la tabla anterior.

EJEMPLO 4.7.2 DISEÑO UNIFACTORIAL DE EFECTOS ALEATORIOS

Los siguientes resultados muestran el porcentaje de agua intersticial total recuperada al centrifugar muestras de piedras areniscas tomadas de diferentes profundidades:

Profundidad de la muestra (m)	Agua recuperada (%)
7	33.3, 33.3, 35.7, 38.1, 31.0, 33.3
8	43.6, 45.2, 47.7, 45.4, 43.8, 46.5
16	73.2, 68.7, 73.6, 70.9, 72.5, 74.5
23	72.5, 70.4, 65.2, 66.7, 77.6, 69.8

Demuestre que el porcentaje de agua recuperada difiere significativamente a las diferentes profundidades.

SOLUCION:

1. El modelo estadístico lineal es: $y_j = \mu + \tau_i + \varepsilon_j$

Por tratarse de un modelo de efectos aleatorios la varianza de cualquier observación vendrá dada por:

$$V(y_j) = \sigma_\tau^2 + \sigma^2 \text{ (componentes de varianza)}$$

2. Contraste de hipótesis:

$$H_0: \sigma_\tau^2 = 0$$

$$H_1: \sigma_\tau^2 > 0$$

3. Criterio de decisión:

Debe rechazarse H_0 si $F_0 = CM_{\text{Tratamientos}} / CM_{\text{Error}} > F_{\alpha, k-1, N-k}$

4. Cálculos:

Profundidad de la muestra (m)	Agua recuperada (%)	Y_i	\bar{y}_i
7	33.3, 33.3, 35.7, 38.1, 31.0, 33.3	204.7	34.12
8	43.6, 45.2, 47.7, 45.4, 43.8, 46.5	272.2	45.37
16	73.2, 68.7, 73.6, 70.9, 72.5, 74.5	433.4	72.23
23	72.5, 70.4, 65.2, 66.7, 77.6, 69.8	422.2	70.37
Totales		1332.5	55.55

SUMA DE CUADRADOS:

$$SC_{\text{Tratamientos}} = \sum_{i=1}^k \frac{y_i^2}{n_i} - \frac{Y^2}{N} = [(204)^2 + (272.2)^2 + (433.4)^2 + (422.2)^2] / 6 - [(1332.15)^2 / 24]$$

$$SC_{\text{Tratamientos}} = 6365.7$$

$$SC_{\text{Error}} = SC_{\text{Total}} - SC_{\text{Tratamientos}}$$

$$SC_{\text{Total}} = \sum_{i=1}^4 \sum_{j=1}^3 y_{ij}^2 - \frac{Y^2}{N}$$

$$SC_{\text{Total}} = (33.3^2 + \dots + 69.8^2) - ((1332)^2 / 24) = 80509.25 - 73981.51 = 6527.74$$

$$SC_{\text{Error}} = 6527.74 - 6365.7 = 162.04$$

TABLA DE ANALISIS DE VARIANZA

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	F_0
Entre tratamientos	6365.7	3	2121.9	261.898
Dentro de tratamientos	162.04	20	8.102	
Total	6527.74	23		

$$F_{0.05, 3, 20} = 3.098, F_{0.1, 3, 20} = 2.38$$

CONCLUSIÓN

Existe efecto significativo al 5% y al 10% de la profundidad sobre el contenido del agua intersticial de las piedras areniscas. En otras palabras, puede concluirse que el contenido de agua intersticial en las piedras areniscas varía con la profundidad de la que fueron tomadas, para los niveles de significación del 5 y el 10%.

Calculo de los componentes de varianza:

Dado que se ha constatado el efecto del factor aleatorio profundidad (m), se procede a estimar qué parte de la variabilidad en el contenido de agua intersticial de las piedras es debida a este factor.

Las estimaciones de los componentes de varianza son:

$$CM_{Error} = \sigma^2 = 8.102$$

$$\sigma_{\tau}^2 = (CM_{tratamientos} - CM_{Error})/n$$

$$\sigma_{\tau}^2 = (2121.9 - 8.102)/ 6 = 352.3$$

La estimación de la varianza de cualquier estimación de la muestra es: $8.102 + 352.3 = 360.4$. La mayor parte de la variabilidad se atribuye a diferencias en la profundidad. La variabilidad debida a la profundidad constituye el 97%.

Los datos para calcular la potencia de la prueba para este análisis son:

$$\lambda = \sqrt{1 + \frac{\sigma_{\tau}^2}{\sigma^2}} = \sqrt{1 + 6 \frac{352.3}{8.102}} = 16.18$$

$$v_1 = k - 1 = 3; v_2 = N - k = 20$$

Al utilizar las tablas del anexo D4, se obtiene una potencia de la prueba de más del 99%.

Prueba de idoneidad del modelo:

Para la prueba de idoneidad del modelo es necesario estimar los parámetros del modelo lineal para luego determinar los valores predichos (ajustados) y sus respectivos residuos, como se muestra a continuación:

Las estimaciones de los efectos para las distintas profundidades son:

Dados: $\mu = \bar{y} = 55.55$ y $\tau_i = \bar{y}_i - \mu$, los valores de los efectos para cada nivel son:

$$\tau_1 = 34.12 - 55.55 = -21.43$$

$$\tau_2 = 45.37 - 55.55 = -10.18$$

$$\tau_3 = 72.23 - 55.55 = 16.68$$

$$\tau_4 = 70.37 - 55.55 = 14.82$$

Por tanto, los valores predichos dados por $\hat{y}_{ij} = 55.55 + \tau_i$, son:

Profundidad	Valor predicho
7	34.12
8	45.37
16	72.23
23	70.37

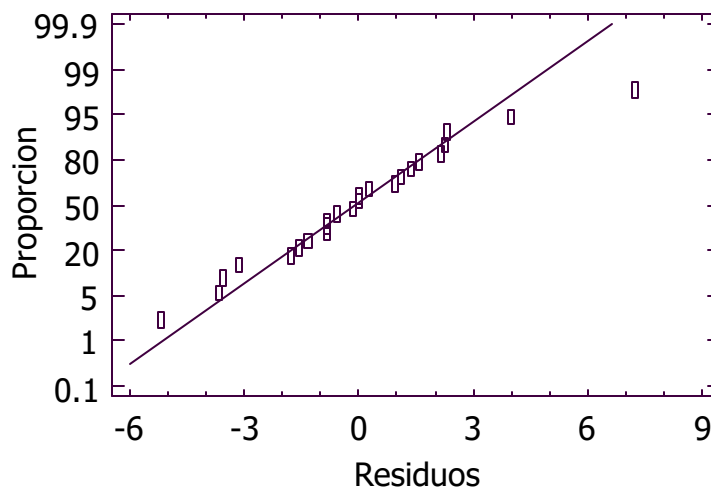
Y los residuos, $\varepsilon_{ij} = y_{ij} - \hat{y}_{ij}$, son entonces:

Profundidad de la muestra (m)	$\varepsilon_{ij} = y_{ij} - \hat{y}_{ij}$
7	-0.82, -0.82, 1.45, 3.98, -3.12, -0.82
8	-1.77, -0.17, 2.33, 0.03, -1.57, 1.13
16	0.97, -3.53, 1.37, -1.33, 0.27, 2.27
23	2.13, 0.03, -5.17, -3.67, 7.23, -0.57

En el gráfico E4.2.1, se presenta la distribución de variabilidad de los residuos en escala probabilística normal:

GRAFICO E4.2.1

Probabilidad normal para los residuos

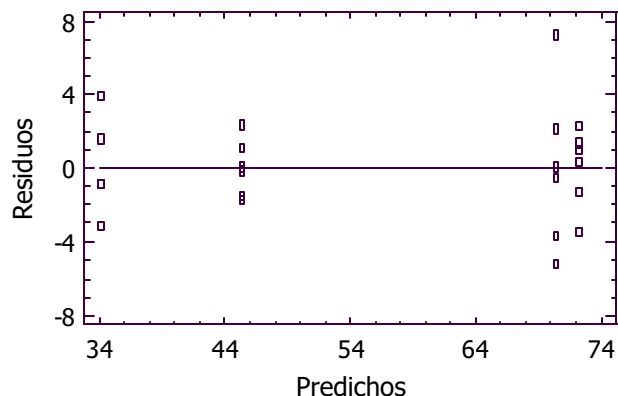


Como se puede observar la distribución de variabilidad de los residuos puede considerarse normal.

En función del análisis del gráfico E4.2.2, que describe el comportamiento de los residuos contra los valores predichos, puede concluirse que la varianza de los residuos es homogénea, los puntos en el gráfico están dispersos sin sugerir ninguna tendencia.

GRAFI CO E4.2.2

Grafico de residuos para agua recuperada (%)

**EJEMPLO 4.7.3 DI SEÑO DE BLOQUES ALEATORIOS**

En un experimento de colaboración entre laboratorios para la determinación de arsénico en carbón (ver capítulo 6), se enviaron muestras de carbón de tres regiones diferentes a tres laboratorios. Cada laboratorio realizó un análisis de cada muestra y para la distribución aleatoria se etiquetaron las muestras como (A1, A2, A3, B1, B2, B3, C1, C2, C3) y los laboratorios como (1, 2, 3). La asignación aleatoria muestra – laboratorio para el ensayo fue la siguiente:

Laboratorio	1	2	3
muestra	A2, B3, C1	A1, B2, C3	A3, B1, C2

Los resultados obtenidos (medidos en $\mu\text{g/g}$) se muestran a continuación:

Muestra	Laboratorio		
	1	2	3
A	5.10	5.35	5.20
B	5.60	5.65	5.35
C	6.40	6.65	6.45

Determinar si existen diferencias significativas en el trabajo de los tres laboratorios.

SOLUCION:

1. El modelo estadístico lineal para el análisis de datos es el siguiente:

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

es un modelo de BLOQUES ALEATORIOS, porque existe heterogeneidad en el material experimental.

2. Contraste de hipótesis:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$$H_1: \mu_i \neq \mu_j \text{ (en al menos un par)}$$

Nivel de significación $\alpha = 0.05$

3. Criterio de decisión:

$$F_o = \frac{SC_{Tratamientos}/(k-1)}{SC_{Error}/(N-k-b+1)} = \frac{CM_{Tratamientos}}{CM_{Error}}$$

Debe rechazarse H_0 si $F_o > F_{\alpha, k-1, N-k}$. Si se rechaza H_0 , estimar los efectos.

4. Cálculos:

MUESTRA	1	2	3	y_{bj}	\bar{y}_{bj}
A	5.1	5.35	5.20	15.65	5.23
B	5.6	5.65	5.35	16.6	5.53
C	6.4	6.65	6.45	19.5	6.5
y_i	17.1	17.65	17.00	51.75	
\bar{y}_i	5.7	5.88	5.67		5.75

CALCULO DE SUMAS DE CUADRADOS

Suma de cuadrados de tratamientos:

$$SC_{Tratamientos} = \sum_{i=1}^k \frac{y_i^2}{n_i} - \frac{Y^2}{N} = (17.1^2 + 17.65^2 + 17^2)/3 - (51.75)^2/9 = 0.08167$$

Suma de cuadrados de bloques:

$$SC_{Bloques} = \sum_{j=1}^b \frac{y_{bj}^2}{k} - \frac{Y^2}{bk} = (15.65^2 + 16.6^2 + 19.5^2)/3 - (53.75)^2/2 = 2.6817$$

Suma de cuadrados total:

$$SC_{Total} = \sum_{i=1}^4 \sum_{j=1}^3 y_{ij}^2 - \frac{Y^2}{N} = (5.1^2 + \dots + 6.45^2) - (51.75)^2/9 = 2.8$$

Suma de cuadrados del error:

$$SC_{Error} = SC_{Total} - SC_{Tratamientos} - SC_{Bloques} = 0.0367$$

Tabla de análisis de varianza

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	F _o
Laboratorio	0.08167	2	0.0408	4.45
Muestra	2.6817	2	1.34	146.27
Error	0.0367	4	0.00917	
Total	2.8	8		

$$F_{0.05, 2, 4} = 6.94$$

CONCLUSIÓN

No existe evidencia significativa como para afirmar que los laboratorios están trabajando de forma diferente, con un grado de confianza del 95%.

Prueba de idoneidad del modelo

Recuerde que los residuos en este modelo son:

$$\varepsilon_{ij} = y_{ij} - \bar{y}_{ki} - \bar{y}_{bj} + \bar{y}$$

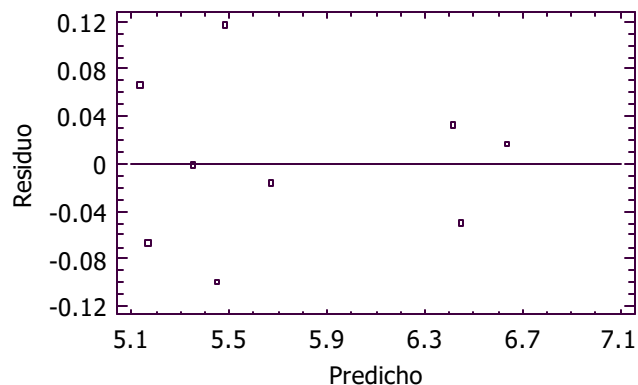
En la tabla siguiente se presentan los residuos calculados:

MUESTRA	1	2	3
A	-0.08	-0.01	0.05
B	0.12	-0.01	-0.1
C	-0.05	0.02	0.03

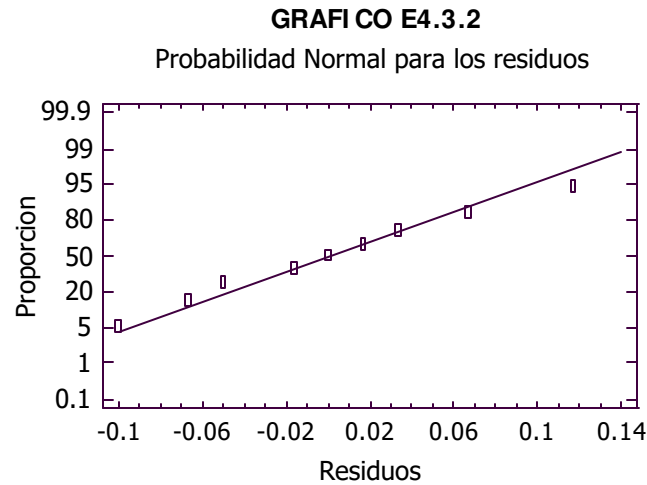
En el gráfico E4.3.1, se presenta el comportamiento de los residuos versus los valores predichos: Como puede observarse del gráfico, hay una distribución aleatoria de los puntos, por lo que se cumple la hipótesis de homogeneidad de la varianza.

GRAFICO E4.3.1

Grafico de residuos para cantidad de Arsenico



La prueba de normalidad de los residuos viene dada por la grafica en papel probabilístico normal y como puede observarse en el grafico E4.3.2 los residuos se distribuyen en forma normal.



EJEMPLO 4.7.4 APLICACIÓN DEL DISEÑO BIFACTORIAL DE EFECTOS FIJOS

En una planta de fabricación de jugos se quiere decidir si la concentración de vitamina C en el concentrado de jugo de naranja varía en función del tiempo de almacenaje y del proveedor. Los resultados, en miligramos de ácido ascórbico por litro se registraron de la siguiente manera:

Marca	Tiempo (días)					
	0		3		7	
XXX	52.6	54.2	49.4	49.2	42.7	48.8
	49.8	46.5	42.8	53.2	40.4	47.6
YYY	56.0	48.0	48.8	44.0	49.2	44.0
	49.6	48.4	44.0	42.4	42.0	43.2
ZZZ	52.5	52.0	48.0	47.0	48.5	43.4
	51.8	53.6	48.2	49.6	45.2	47.6

Utilizando un nivel de significación de 0.05 probar si:

- El nivel de variación en el contenido de ácido ascórbico en el concentrado de jugo de naranja depende del tipo de marca;
- El contenido de ácido ascórbico en el concentrado de jugo de naranja varía con el tiempo de almacenaje;

- c) Las marcas de concentrado de jugo de naranja y el número de días que transcurren desde su almacenamiento no interactúan.

SOLUCION:

Para dar respuestas a los tres literales; hay que probar las siguientes hipótesis:
Para el factor marca (factor A):

$$H_o : \tau_i^A = 0$$

$$H_1 : \tau_j^A \neq 0 \text{ al menos para un } i$$

Para el factor tiempo de almacenaje (factor B):

$$H_o : \tau_j^B = 0$$

$$H_1 : \tau_j^B \neq 0 \text{ al menos para un } j$$

Para el factor de interacción AB:

$$H_o : \tau_{ij}^{AB} = 0$$

$$H_1 : \tau_{ij}^{AB} \neq 0 \text{ al menos para un par } ij$$

De esta forma la hipótesis nula para el factor A se rechaza si:

$$F_A = CM_A/CM_{Error} > F_{\alpha, a-1, ab(n-1)}$$

Para el factor B la hipótesis nula se rechaza si:

$$F_B = CM_B/CM_{Error} > F_{\alpha, b-1, ab(n-1)}$$

Y finalmente, la hipótesis nula para el factor de interacción se rechaza si:

$$F_{AB} = CM_{AB}/CM_{Error} > F_{\alpha, (a-1)(b-1), ab(n-1)}$$

CALCULO DE SUMAS DE CUADRADOS (ver tabla E4.4.1):

Suma de cuadrados total:

$$SC_{Total} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 - \frac{nY^2}{ab} = (52.6^2 + 54.2^2 + \dots + 45.2^2 + 47.6^2) - \frac{4*(431.05)^2}{9} = 531.08$$

Las sumas de cuadrados para los efectos principales son:

$$SC_{MARCA} = \sum_{i=1}^a \frac{ny_i^2}{b} - \frac{nY^2}{ab} = \frac{4 * (144.3^2 + 139.9^2 + 146.85^2)}{3} - \frac{4 * (431.05)^2}{9} = 32.96$$

$$SC_{TIEMPO} = \sum_{j=1}^b \frac{ny_{.j}^2}{a} - \frac{nY^2}{ab} = \frac{4 * (153.75^2 + 141.65^2 + 135.65^2)}{3} - \frac{4 * (431.05)^2}{9} = 226.68$$

$$SC_{Subtotales} = \sum_{i=1}^a \sum_{j=1}^b \frac{y_{ij}^2}{n} - \frac{nY^2}{ab} = \frac{203.1^2 + 194.6^2 + \dots + 192.8^2 + 184.7^2}{4} - \frac{4 * (431.05)^2}{9} = 276.94$$

$$SC_{Interacciones} = SC_{Subtotales} - SC_{MARCA} - SC_{TIEMPO} = 276.94 - 32.97 - 226.68 = 17.29$$

Mientras que la suma de cuadrados del error se calcula por diferencia:

$$SC_{Error} = SC_{Total} - SC_{Subtotales} = 531.08 - 276.94 = 254.14$$

**TABLA E4.4.1 TOTALES Y PROMEDIOS
TIEMPO**

MARCA	0	3	7	TOTALES	ΣPROM.
XXX	52.6 54.2 49.8 46.5 <u>203.1</u> <u>50.78</u>	49.4 49.2 42.8 53.2 <u>194.6</u> <u>48.65</u>	42.7 48.8 40.4 47.6 <u>179.5</u> <u>44.88</u>	577.2	144.3
YYY	56.0 48.0 49.6 48.4 <u>202</u> <u>50.5</u>	48.8 44.0 44.0 42.4 <u>179.2</u> <u>44.8</u>	49.2 44.0 42.0 43.2 <u>178.4</u> <u>44.6</u>	559.6	139.9
ZZZ	52.5 52.0 51.8 53.6 <u>209.9</u> <u>52.48</u>	48.0 47.0 48.2 49.6 <u>192.8</u> <u>48.2</u>	48.5 43.4 45.2 47.6 <u>184.7</u> <u>46.18</u>	587.4	146.85
TOTALES	615	566.6	542.6	1724.2	47.89
ΣPROM.	153.75	141.65	135.65		

A continuación se presenta la tabla E4.4.2 de análisis de varianza:

TABLA E4.4.2 ANÁLISIS DE VARIANZA PARA EL EJEMPLO 4.4

Fuente de variación	Suma de Cuadrados	Grados de libertad	Media de cuadrados	F _o
Factor A	32.97	2	16.48	1.75
Factor B	226.68	2	113.34	12.04
Interacción AB	17.30	4	4.32	0.46
Error	254.14	27	9.41	
Total	531.08	35		

Para ambos factores A ó B, $F_{0.05(2,27)} = 3.35$ y para la interacción AB, $F_{0.05(4,27)} = 2.73$.

Por lo tanto, el único efecto considerado significativo es el efecto producido por el factor tiempo (factor B); es decir, que para un nivel de significación del 5%, la variación en el contenido de vitamina C del concentrado de jugo de naranja sólo depende del tiempo de almacenamiento.

Al analizar los datos en un paquete estadístico, los resultados de la columna P valor indican que la hipótesis nula para el factor A se rechaza para un $\alpha = 0.191$; por consiguiente la hipótesis nula no se rechaza para $\alpha = 0.05$, y la hipótesis nula para el factor B sí se rechaza para cualquier α .

La hipótesis nula para la interacción AB no se rechaza para $\alpha = 0.05$ ya que requiere de un α mínimo $\alpha = 0.767$.

Prueba de idoneidad del modelo

El elemento principal para la prueba de idoneidad del modelo es el análisis de los residuos; los residuos para el modelo factorial de dos factores son:

$$\varepsilon_{ij} = y_{ijk} - \hat{y}_{ijk} = y_{ijk} - \bar{y}_{ij}$$

Los residuos obtenidos para este ejemplo se presentan en la tabla E4.4.3.

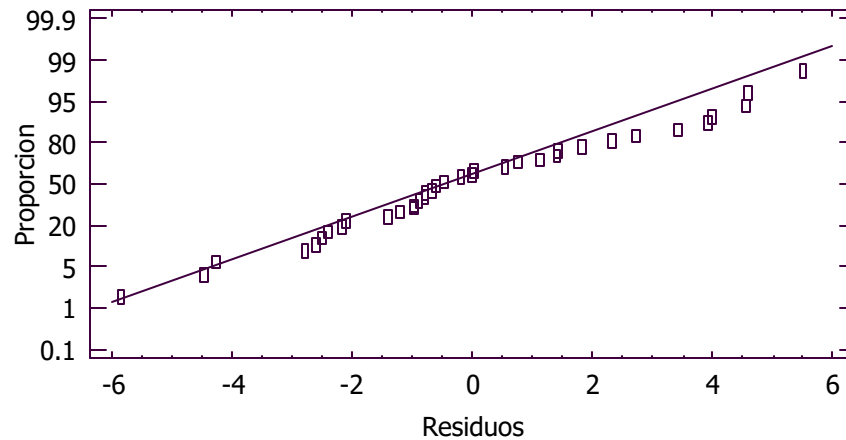
TABLA E4.4.3

MARCA	0	3	7
XXX	1.82 3.42 -0.98 -4.28 <u>50.78</u>	0.75 0.55 -5.85 4.55 <u>48.65</u>	-2.18 3.92 -4.48 2.72 <u>44.88</u>
YYY	5.5 -2.5 -0.9 -2.1 <u>50.5</u>	4.0 -0.8 -0.8 -2.4 <u>44.8</u>	4.6 -0.6 -2.6 -1.4 <u>44.6</u>
ZZZ	0.02 -0.48 -0.68 1.12 <u>52.48</u>	-0.2 -1.2 0 1.4 <u>48.2</u>	2.32 -2.78 -0.98 1.42 <u>46.18</u>

La distribución de los residuos en escala probabilística, se presenta en el gráfico E4.4.1, según ésta la variabilidad de los residuos es prácticamente normal, por lo que se cumple la hipótesis de normalidad.

GRAFICO E4.4.1

Probabilidad normal para los residuos



De las gráficas E4.4.2 y E4.4.3, se puede observar que los errores no siguen ningún modelo matemático específico y además la varianza permanece relativamente constante. Por lo que se puede concluir que el modelo es estadísticamente válido para aplicar el análisis de varianza.

GRAFICO E4.4.2

Gráfica de residuos para Acido

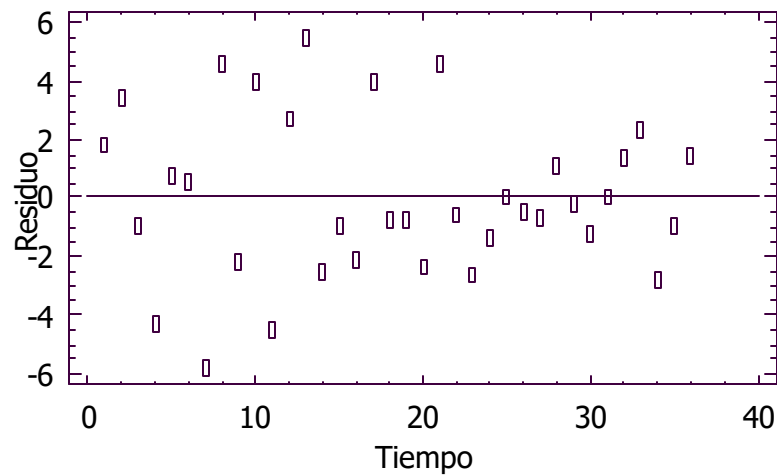
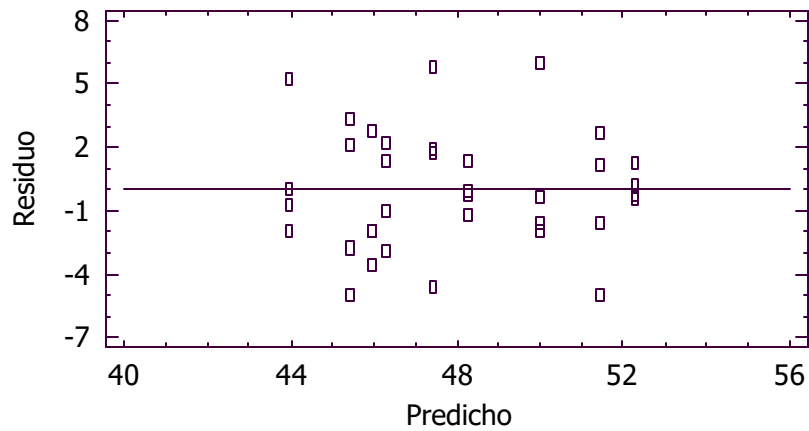


GRAFICO E4.4.3

Grafico de residuos para el Acido



EJEMPLO 4.7.5 DETERMINACION DEL TAMAÑO DE MUESTRA PARA UN DISEÑO BI FACTORIAL

En el ejemplo 4.4, se tomó como tamaño de muestra 4 observaciones por tratamiento, a continuación se explica un procedimiento a seguir para calcularlo.

SOLUCION:

Se obtiene σ^2 a través de una prueba piloto; en este caso supongamos que $\sigma^2 = 0.7$; además se decide rechazar la hipótesis nula con una probabilidad alta ($1-\beta > 90\%$) si la diferencia en el nivel medio del ácido ascórbico fuera por lo menos igual a 5 ($D=5$). Se tiene que:

$$\Phi^2 = \frac{nD^2}{2\sigma^2[(a-1)(b-1)+1]}$$

$$\Rightarrow \Phi^2 = \frac{5^2}{2(0.7)(5)}n$$

$$\Rightarrow \Phi^2 = 3.57n$$

Se partirá de una prueba con un $\alpha = 0.01$ y $1-\beta > 0.9$. Para el cálculo del tamaño de la muestra se hace uso de las curvas características, obteniéndose la tabla E4.5.1:

TABLA E4.5.1

n	Φ^2	Φ	v_1	v_2	$1-\beta$
2	7.14	2.67	2	9	0.30
3	10.71	3.27	2	18	0.70
4	14.28	3.78	2	27	0.94

Luego se debe tomar un tamaño de muestra de 4 unidades experimentales por tratamiento para lograr una prueba con potencia de al menos 0.9.

Nota: Si se toma cualquiera de las ecuaciones:

$$\Phi^2 = \frac{nbD^2}{2a\sigma^2} \quad \Phi^2 = \frac{naD^2}{2b\sigma^2}$$

el tamaño apropiado es de 2 para lograr que $1-\beta = 0.9$. Por lo que se tomó el máximo tamaño de muestra para garantizar potencia en la prueba de hipótesis de la interacción; es decir 4 observaciones por tratamiento.

EJEMPLO 4.7.6 APLICACIÓN DEL DISEÑO 2^k

Actualmente en un laboratorio la determinación de compuestos fenólicos en aguas residuales, se logra por la destilación de la muestra y posterior reacción de los fenoles con 4 amono antipirina en un pH de 10.0, en presencia del ferro cianato. La intensidad del color del compuesto es medida espectrofotométricamente.

Se desea conocer si la etapa de la destilación podría ser eliminada, sustituyéndola por la filtración de la muestra y posterior determinación de los compuestos fenólicos en la suspensión formada. Los factores que se considera afectan al proceso de filtración para la separación de los compuestos fenólicos son:

FACTOR
pH (A)
TEMPERATURA (B)
TIPO DE PAPEL FILTRO (C)

Se decide planificar un experimento 2^3 para estudiar el efecto de los distintos factores sobre la etapa de filtración. Los niveles de los factores se presentan a continuación:

FACTOR	NIVEL (+)	NIVEL (-)
pH	Medio básico	Medio ácido
TEMPERATURA	80 °C	20 °C
TIPO DE PAPEL FILTRO	papel filtro cuantitativo	papel filtro semi-cuantitativo

La variable respuesta a considerar fue el contenido fenólico determinado en muestras procesadas con el método modificado de medición. Normalmente el valor obtenido de la concentración de fenoles, en ese tipo de muestras, a través del método tradicional del análisis gira alrededor de 4.1.

Los resultados obtenidos a través de la planificación de la experiencia se muestran en la siguiente tabla:

EXPERIENCIAS	(1)	a	b	ab	c	ac	bc	abc
Y_{ijk}	4.05	6.92	4.14	5.74	4.06	7.21	4.23	5.79

Compruebe qué efectos afectan más significativamente a la etapa de filtración.

SOLUCION:

Estimación de efectos

En la tabla E4.6.1 se presenta la cuantificación de los efectos de los factores por el método de Yates (ver sección 2.6.4). El número de columnas a crear son $k + 2 = 3 + 2 = 5$.

TABLA E4.6.1 APLICACIÓN DEL METODO DE YATES

Combinaciones	y_{ijk}	(1)	(2)	(3)	efectos	
- - - (1)	4.05	10.97	20.85	42.14	g(T)	5.268
+ - - (a)	6.92	9.88	21.29	9.18	g(A)	2.295
- + - (b)	4.14	11.27	4.47	-2.34	g(B)	-0.585
++ - (ab)	5.74	10.02	4.71	-2.86	g(AB)	-0.715
- - + (c)	4.06	2.87	-1.09	0.44	g(C)	0.11
+ + (ac)	7.21	1.6	-1.25	0.24	g(AC)	0.06
- + + (bc)	4.23	3.15	-1.27	-0.16	g(BC)	-0.04
+++ (abc)	5.79	1.56	-1.59	-0.32	g(ABC)	-0.08

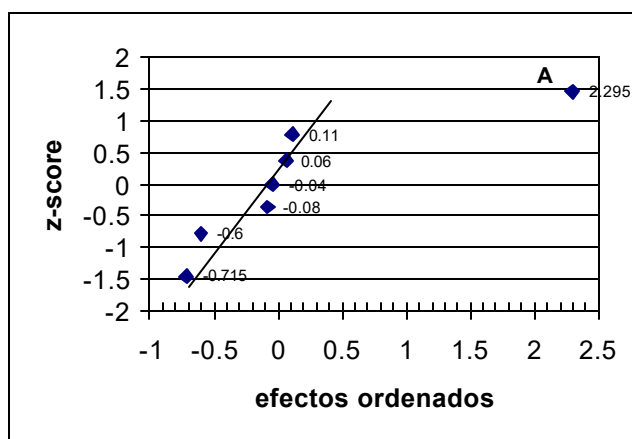
Prueba de significación de efectos

La prueba de significación de los efectos se hará a través de la representación gráfica de los efectos en escala probabilística. En la tabla E4.6.2 se presentan los datos para la construcción de la gráfica E4.6.1.

TABLA E4.6.2

	Efectos	Orden(i)	Percentil $p_i = (i - 0.5) / 7$	z-score $z = 5.05 \left[p_i^{0.135} - (1 - p_i)^{0.135} \right]$
g(A)	2.295	7	0.9286	1.4635
g(B)	-0.585	2	0.2143	-0.7864
g(AB)	-0.715	1	0.0714	-1.4635
g(C)	0.11	6	0.7857	0.7864
g(AC)	0.06	5	0.6428	0.3628
g(BC)	-0.04	4	0.5000	0
g(ABC)	-0.08	3	0.3571	-0.3630

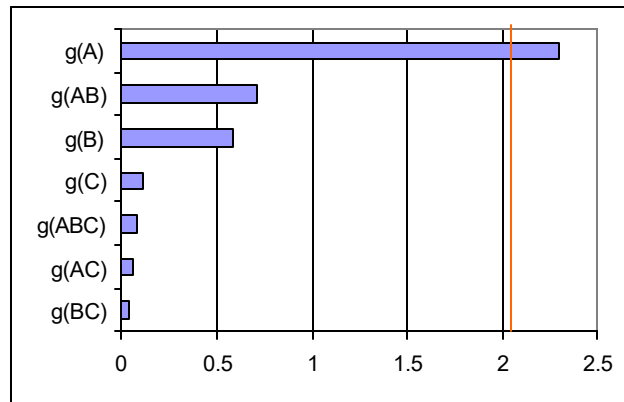
GRAFICA E4.6.1 EFECTOS ORDENADOS



En la gráfica E4.6.2 se presenta el diagrama de Pareto de los efectos. La variable normal para una distribución con la media de los efectos de 0.555 y con la desviación estándar de los mismos de 0.8154, correspondiente al área de probabilidad de 0.95, puede calcularse aplicando la función `DISTR.NORM.INV(0.95,0.555,0.8152)` de Excel, siendo su valor de 1.896.

Como puede observarse de las figuras E4.6.1 y E4.6.2, el único factor significativo es el factor A. Es decir, que la variación en la medición del contenido fenólico de las muestras usando la etapa de filtración sólo es función de la regulación del pH, tendiéndose al uso de pH bajo. Si la prueba de idoneidad lo justifica, sería conveniente proceder con estudios para determinar cuál es el valor de pH óptimo a emplear en la etapa de filtración y luego de comparación de las técnicas (modificada y no modificada), de tal manera que se asegure que la técnica de medición modificada permite obtener resultados confiables.

GRAFICO E4.6.2 GRAFICO DE PARETO DE EFECTOS



Prueba de idoneidad del modelo

Para calcular los residuos $\varepsilon_{ij} = y_{ijk} - \hat{y}_{ijk}$, primero se calcula el valor estimado (\hat{y}_{ijk}) lo cual se hace por medio de un modelo de regresión ajustado, de la forma:

$$\hat{y}_{ijk} = \bar{y} + \frac{g(A)}{2}x_1 + \frac{g(B)}{2}x_2 + \frac{g(AB)}{2}x_1x_2 + \dots + (\text{sobre todos los efectos significativos})$$

donde x_1 y x_2 son los valores codificados (-1), (+1), nivel bajo y alto de cada factor. En este ejemplo, el modelo es el siguiente:

$$\hat{y}_{ijk} = \bar{y} + \frac{g(A)}{2}x_1 = 5.268 + 1.148x_1$$

El valor predicho para ambos factores a nivel bajo, es por ejemplo:

$$\hat{y}_{ijk} = 5.268 + 1.148(-1) = 4.12$$

El cálculo de los residuos se presenta en la tabla E4.6.3.

TABLA E4.6.3 CALCULO DE RESIDUOS

Experimento	y_{ijk}	\hat{y}_{ijk}	$\varepsilon_{ij} = y_{ijk} - \hat{y}_{ijk}$
(1)	4.05	4.12	-0.07
A	6.92	6.416	0.504
B	4.14	4.12	0.02
AB	5.74	6.416	-0.676
C	4.06	4.12	-0.06
AC	7.21	6.416	0.794
BC	4.23	4.12	0.11
ABC	5.79	6.416	-0.626

La prueba de homogeneidad de la varianza viene dada por el gráfico de residuos versus los valores predichos, gráfico E4.6.3. Mientras que la prueba de normalidad de los residuos se presenta en el gráfico E4.6.4.

GRAFICO E4.6.3

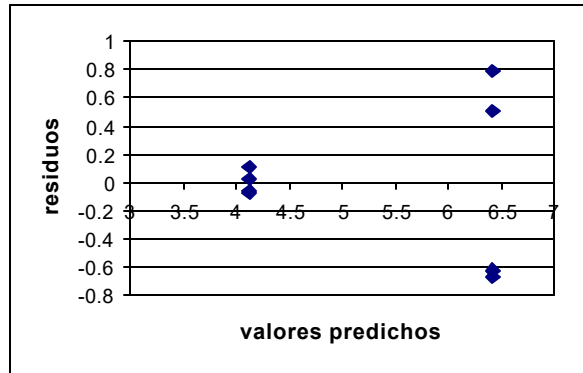
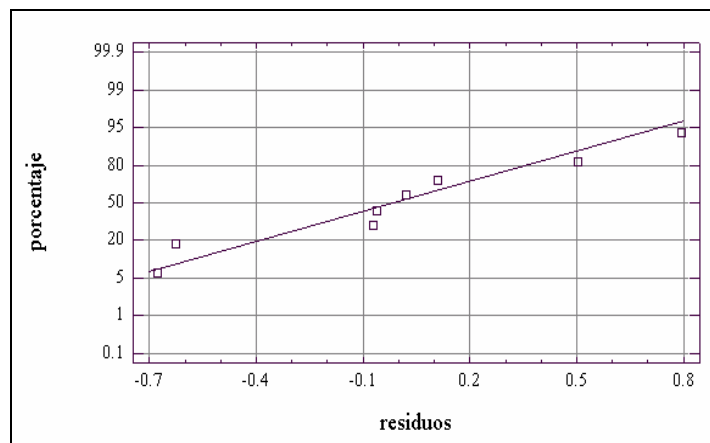


GRAFICO E4.6.3



Como puede observarse no existe ninguna tendencia que evidencie la falta de homogeneidad de la varianza (gráfico E4.6.3); y la distribución de variabilidad de los residuos puede considerarse normal (gráfico E4.6.4).

En el archivo [Yates k=3](#), se presenta una hoja para el calculo de efectos por el algoritmo de Yates para un modelo 2^3 .



Yates k=3

REFERENCIAS BIBLIOGRÁFICAS

- **Box, G.E.P, Hunter, N.G y Hunter, J.S (1989).** Estadística para investigadores. España: Reverté.
- **Ferré, J. y Rius, F.X (2000).** Introducción al diseño estadístico de experimentos. Tarragona: Departamento de Química Analítica y Química Orgánica. Universitat Rovira i Virgili.
- **Juran, J.M y Gryna, F. M. (1998).** Manual de control de calidad. (2 tomos). España: Mcgraw-Hill.
- **Kenett, R. S. y Zacks, S. (2000).** Estadística industrial moderna. Diseño y control de la calidad y la confiabilidad. México: International Thomson Editores.
- **Kuehl, R. O. (2001).** Diseño de experimentos. Principios estadísticos de diseño y análisis de investigación. México:International Thomson Editores.
- **Lawson, J., Madrigal, J.L. y Erjavec, J. (1992).** Estrategias experimentales para el mejoramiento de la calidad en la industria. México: Grupo Editorial Iberoamérica.
- **Miller, J. C. y Miller, J. N. (1993).** Estadística para química analítica. Segunda edición. Wilmington, Delaware, E.U.A: Addison-Wesley Iberoamericana, S.A.
- **Montgomery, D. C. (1991).** Diseño y análisis de experimentos. México: Grupo Editorial Iberoamérica.
- **Prat, A., Tort-Martorell, X., Grima, P. y Pozueta, L. (2000).** Métodos estadísticos. Control y mejora de la calidad. México: Alfaomega Grupo Editor.
- **Walpole, R. E. y Myers, R. H. (1992).** Probabilidad y estadística. Cuarta edición. México: Mcgraw-Hill.
- **Wayne, W. D. (1992).** Bioestadística. Base para el análisis de las ciencias de la salud. Cuarta edición. México: Limusa.

ANEXO A4 TABLA DE NUMEROS ALEATORIOS
(MC DANIELS y GATES, 1999, pág. A-36)

Números aleatorios

63271	59986	71744	51102	15141	80714	58683	93108	13554	79945
88547	09896	95436	79115	08303	01041	20030	63754	08459	28364
55957	57243	83865	09911	19761	66535	40102	26646	60147	15702
46276	87453	44790	64122	45573	84358	21625	16999	13385	22782
55363	07449	34835	15290	76616	67191	12777	21861	68689	03263
69393	92785	49902	58447	42048	30378	87618	26933	40640	16281
13186	29431	88190	04588	38733	81290	89541	70290	40113	08243
17726	28652	56836	78351	47327	18518	92222	55201	27340	10493
36520	64465	05550	30157	82242	29520	69753	72602	23756	54935
81628	36100	39254	56835	37636	02421	98063	89641	64953	99337
84649	48968	75215	75498	49539	74240	03466	49292	36401	45525
63291	11618	12613	75055	43915	26488	41116	64531	56827	30825
70502	53225	03655	05915	37140	57051	48393	91322	25653	06543
06426	24771	59935	49801	11082	66762	94477	02494	88215	27191
20711	55609	29430	70165	45406	78484	31639	52009	18873	96927
41990	70538	77191	25860	55204	73417	83920	69468	74972	38712
72452	36618	76298	26678	89334	33938	95567	29380	75906	91807
37042	40318	57099	10528	09925	89773	41335	96244	29002	46453
53766	52875	15987	46962	67342	77592	57651	95508	80033	69828
90585	58955	53122	16025	84299	53310	67380	84249	25348	04332
32001	96293	37203	64516	51530	37069	40261	61374	05815	06714
62606	64324	46354	72157	67248	20135	49804	09226	64419	29457
10078	28073	85389	50324	14500	15562	64165	06125	71353	77669
91561	46145	24177	15294	10061	98124	75732	00815	83452	97355
13091	98112	53959	79607	52244	63303	10413	63839	74762	50289
73864	83014	72457	22682	03033	61714	88173	90835	00634	85169
66668	25467	48894	51043	02365	91726	09365	63167	95264	45643
84745	41042	29493	01836	09044	51926	43630	63470	76508	14194
48068	26805	94595	47907	13357	38412	33318	26098	82782	42851
54310	96175	97594	88616	42035	38093	36745	56702	40644	83514
14877	33095	10924	58013	61439	21882	42059	24177	58739	60170
78295	23179	02771	43464	59061	71411	05697	67194	30495	21157
67524	02865	39593	54278	04237	92441	26602	63835	38032	94770
58268	57219	68124	73455	83236	08710	04284	55005	84171	42596
97158	28672	50685	01181	24262	19427	52106	34308	73685	74246
04230	16831	69085	30802	65559	09205	71829	06489	85650	38707
94879	56606	30401	02602	57658	70091	54986	41394	60437	03195
71446	15232	66715	26385	91518	70566	02888	79941	39684	54315
32886	05644	79316	09819	00813	88407	17461	73925	53037	91904
62048	33711	25290	21526	02223	75947	66466	06332	10913	75336
84534	42351	21628	53669	81352	95152	08107	98814	72743	12849
84707	15885	84710	35866	06446	86311	32648	88141	73902	69981
19409	40868	64220	80861	13860	68493	52908	26374	63297	45052
57978	48015	25973	66777	45924	56144	24742	96702	88200	66162
57295	98298	11199	96510	75228	41600	47192	43267	35973	23152
94044	83785	93388	07833	38216	31413	70555	03023	54147	06647
30014	25879	71763	96679	90603	99396	74557	74224	18211	91637
07265	69563	64268	88802	72264	66540	01782	08396	19251	83613
84404	88642	30263	80310	11522	57810	27627	78376	36240	48952
21778	02085	27762	46097	43324	34354	09369	14966	10158	76089

ANEXO B4 VALORES CRITICOS PARA LA PRUEBA DE DUNNETT DE COMPARACIONES BI LATERALES DE TRATAMIENTOS CON UN CONTROL

$d_{0.05}(k, \nu)$, k número de tratamientos sin incluir el control
(KUEHL, 2001, pág. 598)

$\alpha = 0.05$ (de dos lados)													
ν, k	2	3	4	5	6	7	8	9	10	11	12	15	20
5	3.03	3.29	3.48	3.62	3.73	3.82	3.90	3.97	4.03	4.09	4.14	4.26	4.42
6	2.86	3.10	3.26	3.39	3.49	3.57	3.64	3.71	3.76	3.81	3.86	3.97	4.11
7	2.75	2.97	3.12	3.24	3.33	3.41	3.47	3.53	3.58	3.63	3.67	3.78	3.91
8	2.67	2.88	3.02	3.13	3.22	3.29	3.35	3.41	3.46	3.50	3.54	3.64	3.76
9	2.61	2.81	2.95	3.05	3.14	3.20	3.26	3.32	3.36	3.40	3.44	3.53	3.65
10	2.57	2.76	2.89	2.99	3.07	3.14	3.19	3.24	3.29	3.33	3.36	3.45	3.57
11	2.53	2.72	2.84	2.94	3.02	3.08	3.14	3.19	3.23	3.27	3.30	3.39	3.50
12	2.50	2.68	2.81	2.90	2.98	3.04	3.09	3.14	3.18	3.22	3.25	3.34	3.45
13	2.48	2.65	2.78	2.87	2.94	3.00	3.06	3.10	3.14	3.18	3.21	3.29	3.40
14	2.46	2.63	2.75	2.84	2.91	2.97	3.02	3.07	3.11	3.14	3.18	3.26	3.36
15	2.44	2.61	2.73	2.82	2.89	2.95	3.00	3.04	3.08	3.12	3.15	3.23	3.33
16	2.42	2.59	2.71	2.80	2.87	2.92	2.97	3.02	3.06	3.09	3.12	3.20	3.30
17	2.41	2.58	2.69	2.78	2.85	2.90	2.95	3.00	3.03	3.07	3.10	3.18	3.27
18	2.40	2.56	2.68	2.76	2.83	2.89	2.94	2.98	3.01	3.05	3.08	3.16	3.25
19	2.39	2.55	2.66	2.75	2.81	2.87	2.92	2.96	3.00	3.03	3.06	3.14	3.23
20	2.38	2.54	2.65	2.73	2.80	2.86	2.90	2.95	2.98	3.02	3.05	3.12	3.22
24	2.35	2.51	2.61	2.70	2.76	2.81	2.86	2.90	2.94	2.97	3.00	3.07	3.16
30	2.32	2.47	2.58	2.66	2.72	2.77	2.82	2.86	2.89	2.92	2.95	3.02	3.11
40	2.29	2.44	2.54	2.62	2.68	2.73	2.77	2.81	2.85	2.87	2.90	2.97	3.06
60	2.27	2.41	2.51	2.58	2.64	2.69	2.73	2.77	2.80	2.83	2.86	2.92	3.00
120	2.24	2.38	2.47	2.55	2.60	2.65	2.69	2.73	2.76	2.79	2.81	2.87	2.95
∞	2.21	2.35	2.44	2.51	2.57	2.61	2.65	2.69	2.72	2.74	2.77	2.83	2.91

$d_{0.01}(a, f)$, k número de tratamientos sin incluir el control
(MONTGOMERY, 1991, pág. 560)

$a - 1 =$ número de medias de tratamientos (excluyendo el control)									
f	1	2	3	4	5	6	7	8	9
5	4.03	4.63	4.98	5.22	5.41	5.56	5.69	5.80	5.89
6	3.71	4.21	4.51	4.71	4.87	5.00	5.10	5.20	5.28
7	3.50	3.95	4.21	4.39	4.53	4.64	4.74	4.82	4.89
8	3.36	3.77	4.00	4.17	4.29	4.40	4.48	4.56	4.62
9	3.25	3.63	3.85	4.01	4.12	4.22	4.30	4.37	4.43
10	3.17	3.53	3.74	3.88	3.99	4.08	4.16	4.22	4.28
11	3.11	3.45	3.65	3.79	3.89	3.98	4.05	4.11	4.16
12	3.05	3.39	3.58	3.71	3.81	3.89	3.96	4.02	4.07
13	3.01	3.33	3.52	3.65	3.74	3.82	3.89	3.94	3.99
14	2.98	3.29	3.47	3.59	3.69	3.76	3.83	3.88	3.93
15	2.95	3.25	3.43	3.55	3.64	3.71	3.78	3.83	3.88
16	2.92	3.22	3.39	3.51	3.60	3.67	3.73	3.78	3.83
17	2.90	3.19	3.36	3.47	3.56	3.63	3.69	3.74	3.79
18	2.88	3.17	3.33	3.44	3.53	3.60	3.66	3.71	3.75
19	2.86	3.15	3.31	3.42	3.50	3.57	3.63	3.68	3.72
20	2.85	3.13	3.29	3.40	3.48	3.55	3.60	3.65	3.69
24	2.80	3.07	3.22	3.32	3.40	3.47	3.52	3.57	3.61
30	2.75	3.01	3.15	3.25	3.33	3.39	3.44	3.49	3.52
40	2.70	2.95	3.09	3.19	3.26	3.32	3.37	3.41	3.44
60	2.66	2.90	3.03	3.12	3.19	3.25	3.29	3.33	3.37
120	2.62	2.85	2.97	3.06	3.12	3.18	3.22	3.26	3.29
∞	2.58	2.79	2.92	3.00	3.06	3.11	3.15	3.19	3.22

**ANEXO C4 INTERVALOS SIGNIFICATIVOS PARA LA PRUEBA DE
INTERVALOS MÚLTIPLES DE DUNCAN**
(MONTGOMERY, 1991, pág. 555-556)

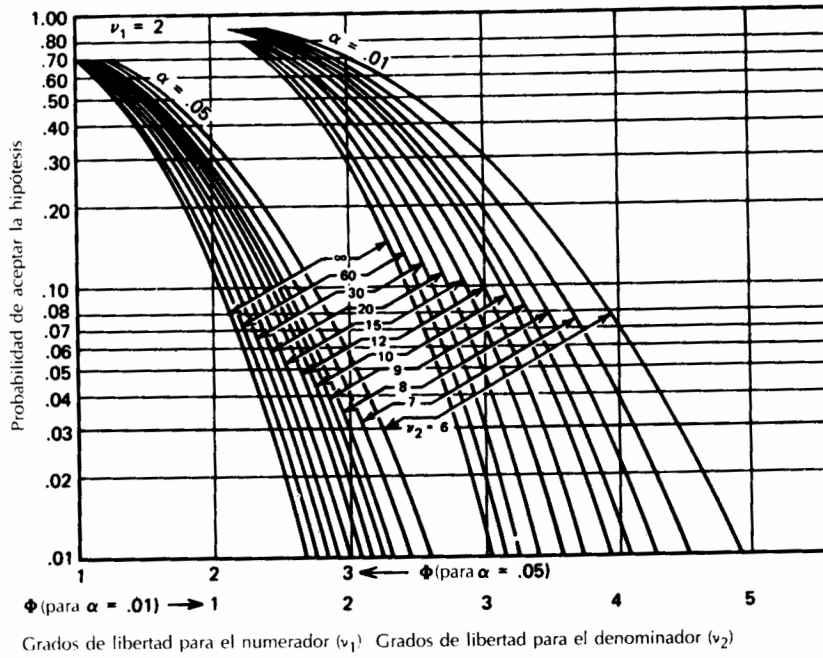
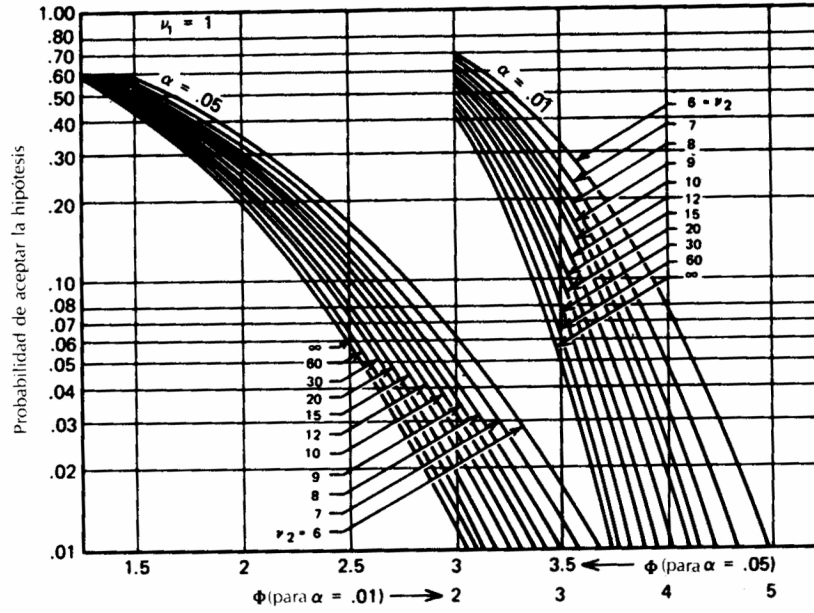
$r_{.01}(p, f)$

f	p											
	2	3	4	5	6	7	8	9	10	20	50	100
1	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0
2	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0
3	8.26	8.5	8.6	8.7	8.8	8.9	8.9	9.0	9.0	9.3	9.3	9.3
4	6.51	6.8	6.9	7.0	7.1	7.1	7.2	7.2	7.3	7.5	7.5	7.5
5	5.70	5.96	6.11	6.18	6.26	6.33	6.40	6.44	6.5	6.8	6.8	6.8
6	5.24	5.51	5.65	5.73	5.81	5.88	5.95	6.00	6.0	6.3	6.3	6.3
7	4.95	5.22	5.37	5.45	5.53	5.61	5.69	5.73	5.8	6.0	6.0	6.0
8	4.74	5.00	5.14	5.23	5.32	5.40	5.47	5.51	5.5	5.8	5.8	5.8
9	4.60	4.86	4.99	5.08	5.17	5.25	5.32	5.36	5.4	5.7	5.7	5.7
10	4.48	4.73	4.88	4.96	5.06	5.13	5.20	5.24	5.28	5.55	5.55	5.55
11	4.39	4.63	4.77	4.86	4.94	5.01	5.06	5.12	5.15	5.39	5.39	5.39
12	4.32	4.55	4.68	4.76	4.84	4.92	4.96	5.02	5.07	5.26	5.26	5.26
13	4.26	4.48	4.62	4.69	4.74	4.84	4.88	4.94	4.98	5.15	5.15	5.15
14	4.21	4.42	4.55	4.63	4.70	4.78	4.83	4.87	4.91	5.07	5.07	5.07
15	4.17	4.37	4.50	4.58	4.64	4.72	4.77	4.81	4.84	5.00	5.00	5.00
16	4.13	4.34	4.45	4.54	4.60	4.67	4.72	4.76	4.79	4.94	4.94	4.94
17	4.10	4.30	4.41	4.50	4.56	4.63	4.68	4.73	4.75	4.89	4.89	4.89
18	4.07	4.27	4.38	4.46	4.53	4.59	4.64	4.68	4.71	4.85	4.85	4.85
19	4.05	4.24	4.35	4.43	4.50	4.56	4.61	4.64	4.67	4.82	4.82	4.82
20	4.02	4.22	4.33	4.40	4.47	4.53	4.58	4.61	4.65	4.79	4.79	4.79
30	3.89	4.06	4.16	4.22	4.32	4.36	4.41	4.45	4.48	4.65	4.71	4.71
40	3.82	3.99	4.10	4.17	4.24	4.30	4.34	4.37	4.41	4.59	4.69	4.69
60	3.76	3.92	4.03	4.12	4.17	4.23	4.27	4.31	4.34	4.53	4.66	4.66
100	3.71	3.86	3.98	4.06	4.11	4.17	4.21	4.25	4.29	4.48	4.64	4.65
∞	3.64	3.80	3.90	3.98	4.04	4.09	4.14	4.17	4.20	4.41	4.60	4.68

$r_{.05}(p, f)$

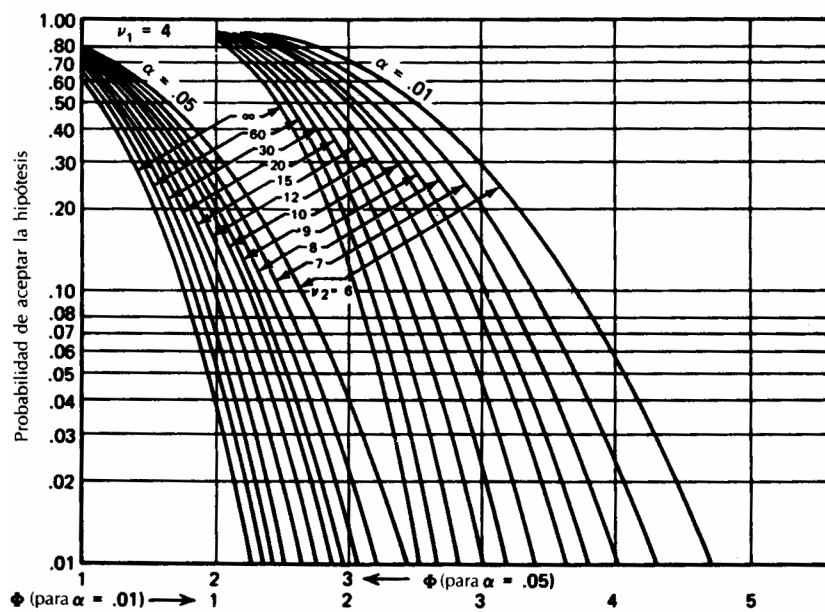
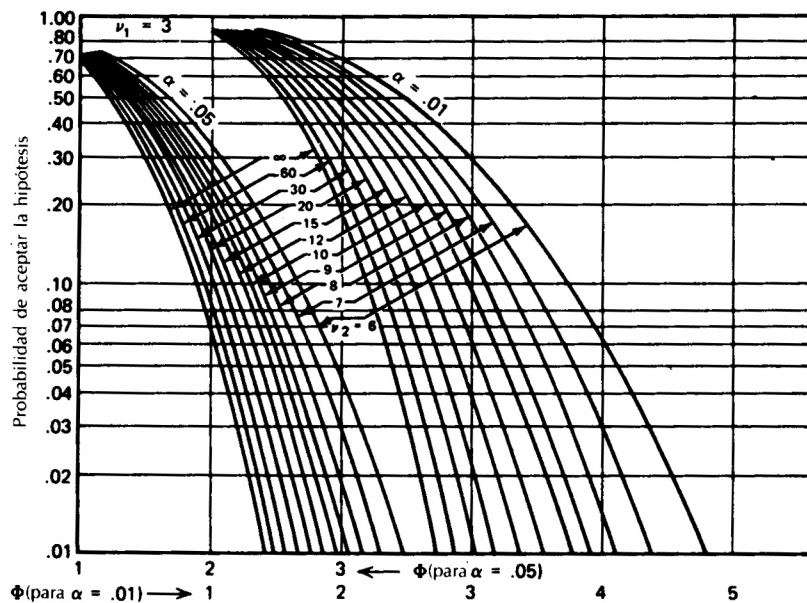
f	p											
	2	3	4	5	6	7	8	9	10	20	50	100
1	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
2	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
3	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50
4	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
5	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83
6	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68
7	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61	3.61
8	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56	3.56
9	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52	3.52
10	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.48	3.48	3.48
11	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.48	3.48	3.48
12	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.46	3.48	3.48	3.48
13	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.47	3.47	3.47
14	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.47	3.47	3.47
15	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.47	3.47	3.47
16	3.00	3.15	3.23	3.30	3.34	3.37	3.39	3.41	3.43	3.47	3.47	3.47
17	2.98	3.13	3.22	3.28	3.33	3.36	3.38	3.40	3.42	3.47	3.47	3.47
18	2.97	3.12	3.21	3.27	3.32	3.35	3.37	3.39	3.41	3.47	3.47	3.47
19	2.96	3.11	3.19	3.26	3.31	3.35	3.37	3.39	3.41	3.47	3.47	3.47
20	2.95	3.10	3.18	3.25	3.30	3.34	3.36	3.38	3.40	3.47	3.47	3.47
30	2.89	3.04	3.12	3.20	3.25	3.29	3.32	3.35	3.37	3.47	3.47	3.47
40	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.47	3.47	3.47
60	2.83	2.98	3.08	3.14	3.20	3.24	3.28	3.31	3.33	3.47	3.48	3.48
100	2.80	2.95	3.05	3.12	3.18	3.22	3.26	3.29	3.32	3.47	3.53	3.53
∞	2.77	2.92	3.02	3.09	3.15	3.19	3.23	3.26	3.29	3.47	3.61	3.67

ANEXO D4 CURVAS CARACTERÍSTICAS DE OPERACIÓN PARA EL ANÁLISIS DE VARIANZA DE EFECTOS FIJOS
(MONTGOMERY, 1991, pág. 547 – 550)

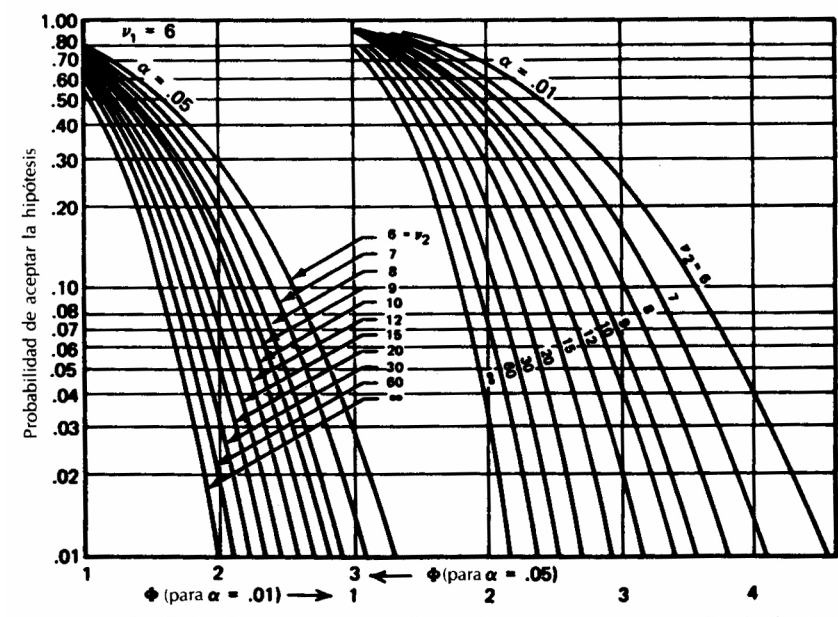
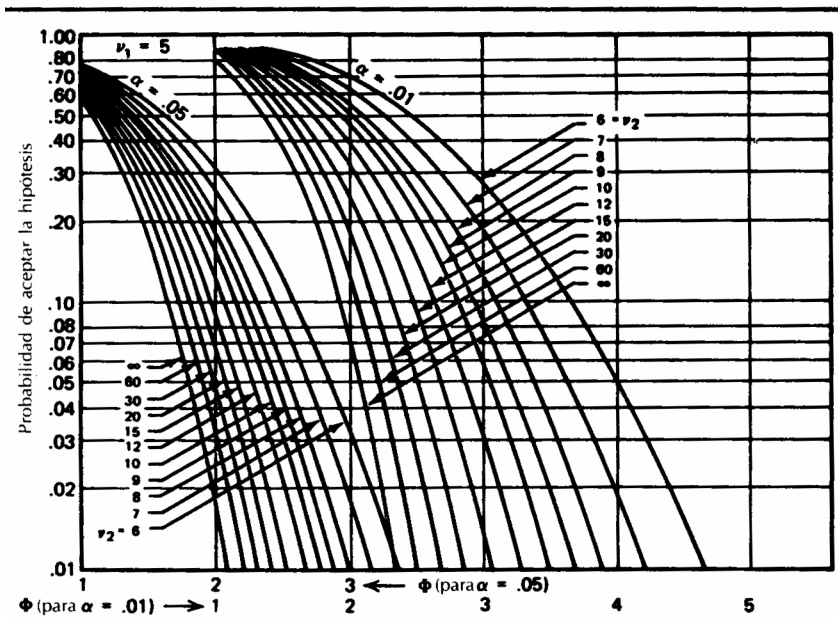


Grados de libertad para el numerador (ν_1) Grados de libertad para el denominador (ν_2)

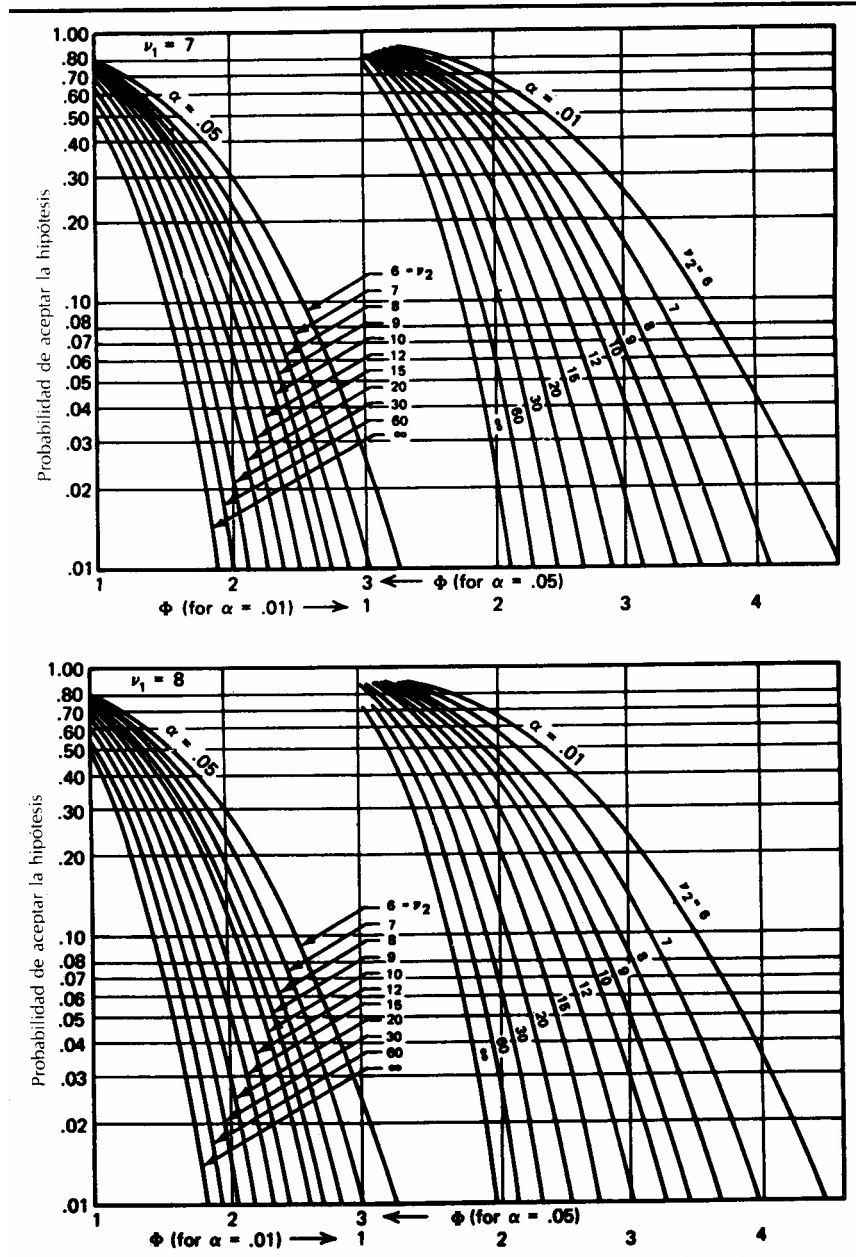
ANEXO D4 CURVAS CARACTERÍSTICAS DE OPERACIÓN PARA EL ANÁLISIS DE VARIANZA DE EFECTOS FIJOS (continuación)
(MONTGOMERY, 1991, pág. 547 – 550)



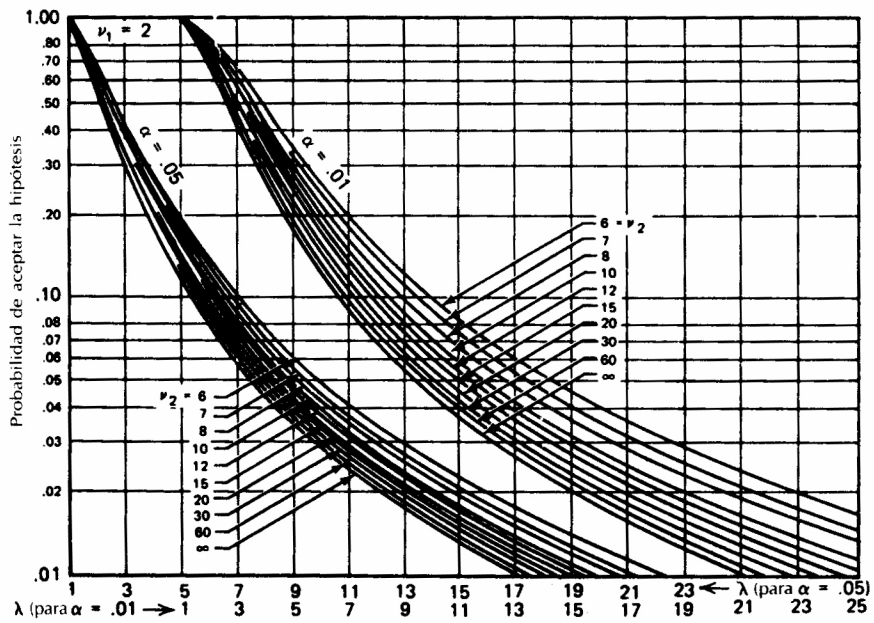
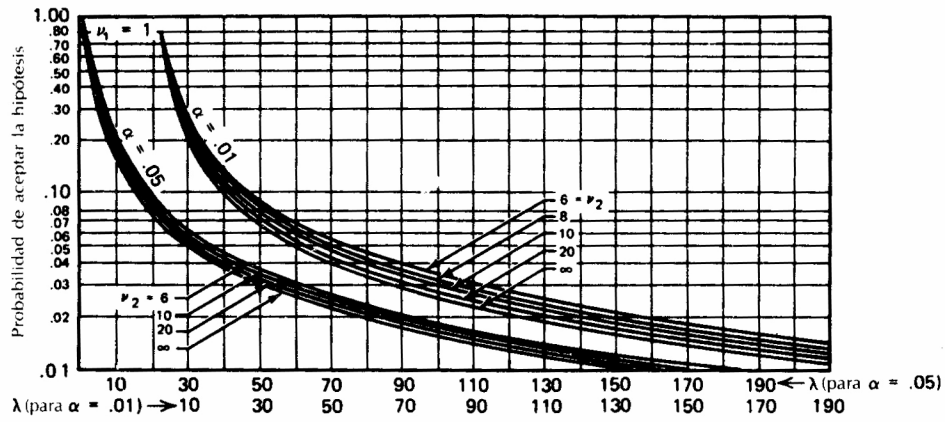
ANEXO D4 CURVAS CARACTERÍSTICAS DE OPERACIÓN PARA EL ANÁLISIS DE VARIANZA DE EFECTOS FIJOS (continuación)
(MONTGOMERY, 1991, pág. 547 – 550)



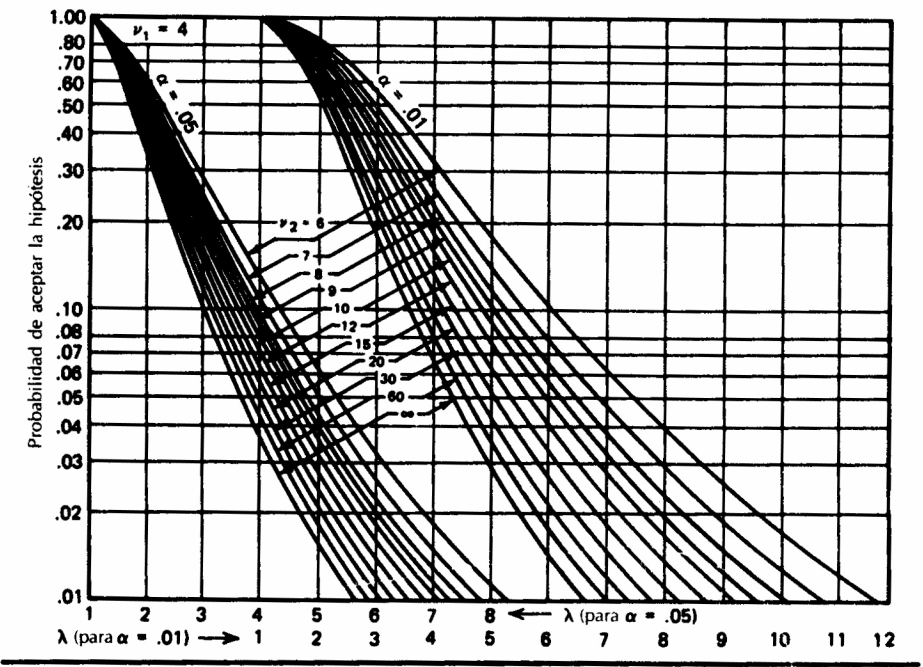
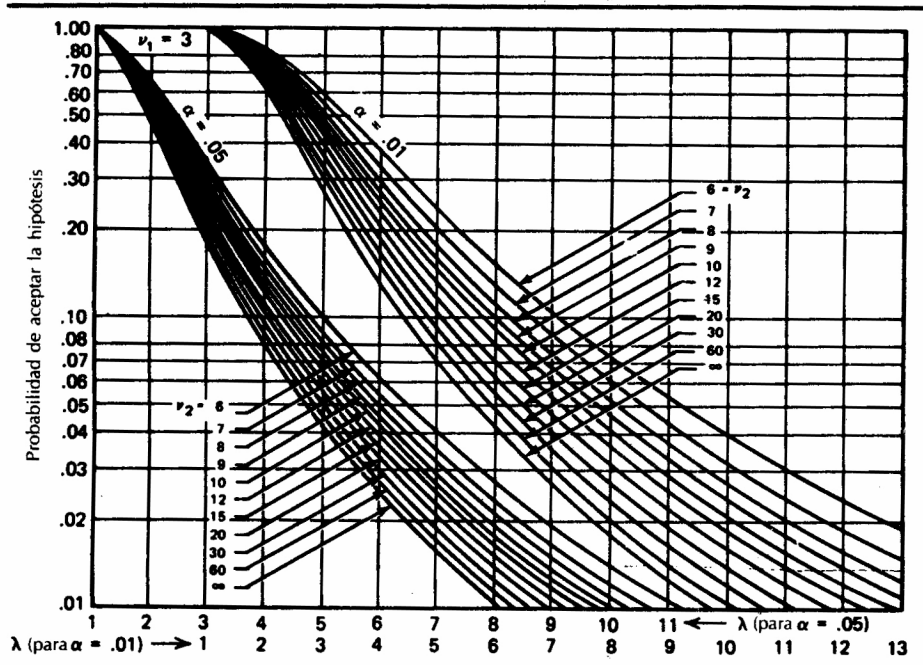
**ANEXO D4 CURVAS CARACTERÍSTICAS DE OPERACIÓN PARA EL
ANÁLISIS DE VARIANZA DE EFECTOS FIJOS (continuación)**
(MONTGOMERY, 1991, pág. 547 – 550)



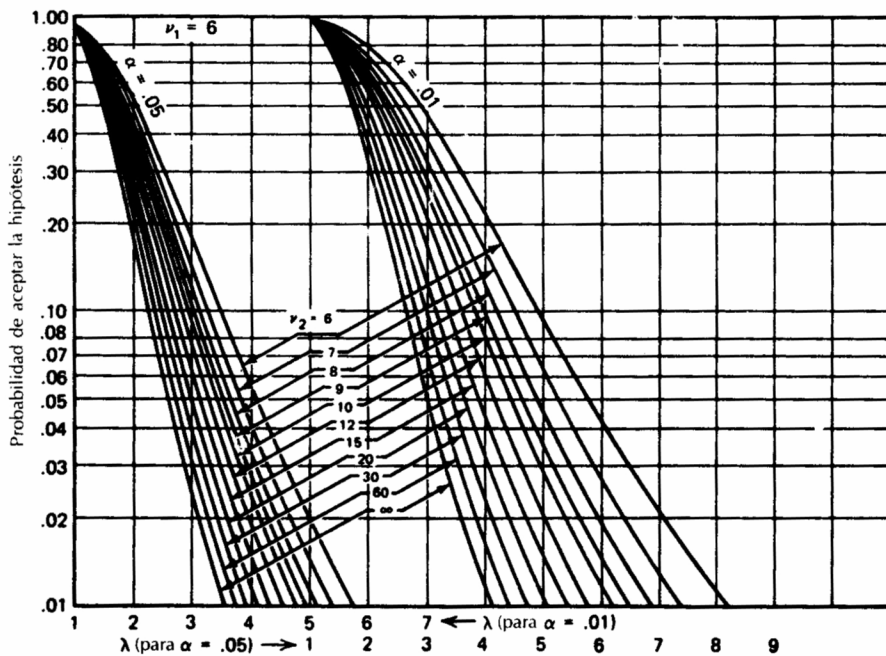
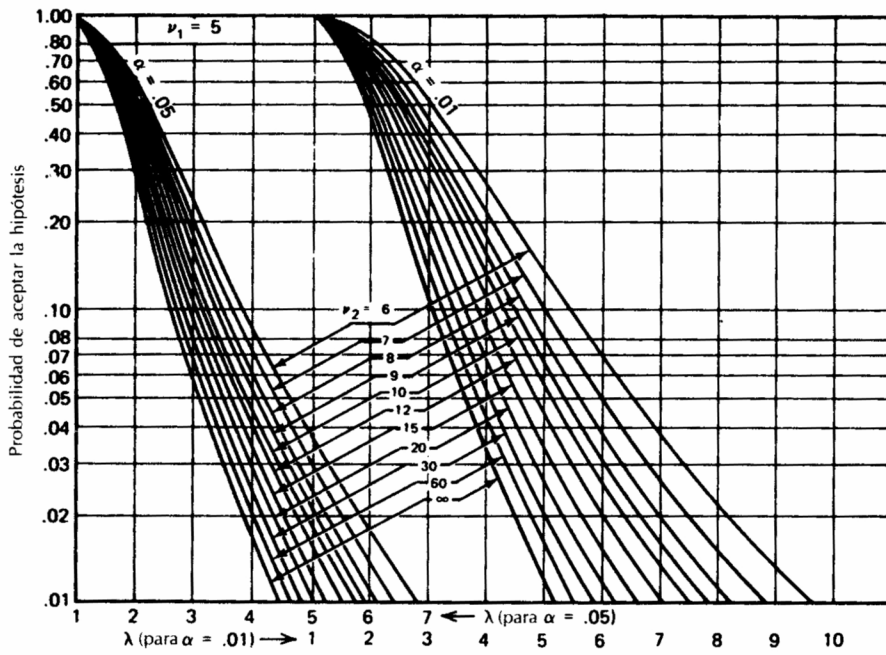
ANEXO D4 CURVAS CARACTERÍSTICAS DE OPERACIÓN PARA EL ANÁLISIS DE VARIANZA DE EFECTOS ALEATORIOS
(MONTGOMERY, 1991, pág. 551 – 554)



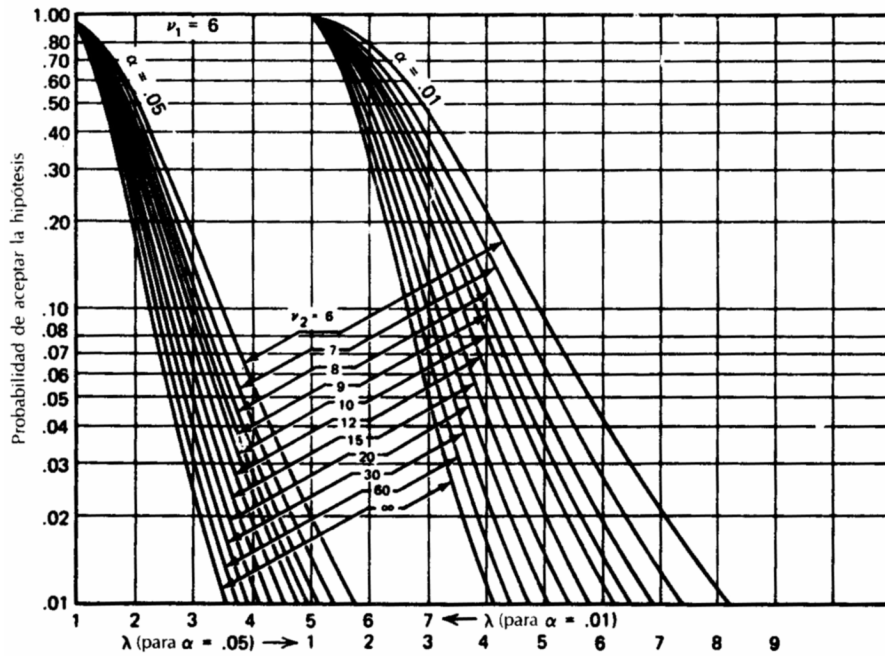
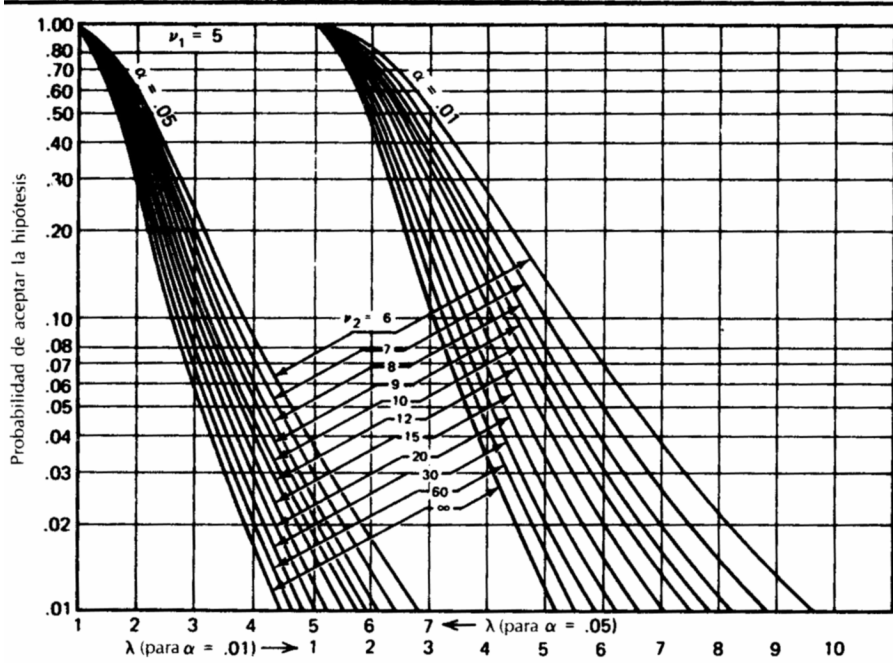
ANEXO D4 CURVAS CARACTERÍSTICAS DE OPERACIÓN PARA EL ANÁLISIS DE VARIANZA DE EFECTOS ALEATORIOS (continuación)
(MONTGOMERY, 1991, pág. 551 – 554)



ANEXO D4 CURVAS CARACTERÍSTICAS DE OPERACIÓN PARA EL ANÁLISIS DE VARIANZA DE EFECTOS ALEATORIOS (continuación)
(MONTGOMERY, 1991, pág. 551 – 554)



ANEXO D4 CURVAS CARACTERÍSTICAS DE OPERACIÓN PARA EL ANÁLISIS DE VARIANZA DE EFECTOS ALEATORIOS (continuación)
(MONTGOMERY, 1991, pág. 551 – 554)



CAPITULO V

ANÁLISIS DE REGRESIÓN APLICADO AL ANÁLISIS QUÍMICO

CONTENIDO

	Pág.
5.0 Análisis de regresión	5-2
5.1 Regresión lineal simple	5-4
5.2 Aplicaciones de la regresión lineal simple en el análisis químico	5-13
5.3 Regresión lineal múltiple	5-22
5.4 Regresión no lineal	5-33
5.5 Ejemplos de aplicación de análisis de regresión	5-34
REFERENCIAS BIBLIOGRAFICAS	5-56



5.0 ANALISIS DE REGRESION

Muchos problemas de investigación requieren la estimación de las relaciones existentes entre la pauta de variabilidad de una variable aleatoria y los valores de una o más variables (aleatorias o no) de las que la primera depende o puede depender. Por ejemplo, ¿Cómo varía la densidad con la temperatura?

El análisis de regresión es una técnica estadística para la estimación de los parámetros de una ecuación que relaciona una determinada variable con un conjunto de variables. El análisis se lleva a cabo mediante el establecimiento de "un modelo de regresión" cuyos parámetros recogen y cuantifican los efectos que se pretende estudiar, tal como el que se esquematiza en la figura 5.1. Dichos parámetros se estiman a partir de los datos disponibles. Algunos autores se refieren al análisis de regresión como TECNICA DE LOS MINIMOS CUADRADOS O DE AJUSTE DE CURVAS (JURAN y GRUNA, 1997, cap. 23).

Observaciones o datos del problema	Variables a explicar (aleatoria)	Variables o factores explicativos (aleatorios o no)
	Y ₁ ---- Y _j ---- Y _J	X ₁₁ ...X _{1i} ...X _{1i}

MODELO FUNDAMENTAL $E(Y_j) = f(X_{1j} \dots X_{ij} \dots X_{ij})$

$$\sigma^2(Y_j) = \text{CONSTANTE}$$

**Técnicas sofisticadas también permiten
Estudiar el caso en el que además**

$$\sigma^2(Y_j) = g(X_{1j} \dots X_{ij} \dots X_{ij})$$

FIGURA 5.1 MODELOS DE REGRESION MULTIPLE
(ROMERO y ZUNICA, 1993, cap. 11)

La razón básica por la cual se construye un modelo de regresión es describir la naturaleza de una relación en forma cuantitativa. Sin embargo, los objetivos son con frecuencia más

específicos. Por ejemplo, para un proceso en el que la humedad es controlable, el objetivo podría ser hallar el valor particular de la humedad que minimiza inestabilidad de un instrumento, o alguna función de costos basada en dicha inestabilidad. O bien determinar variables independientes importantes de un proceso. Por ejemplo, ver si la humedad, presión y temperatura afectan a una característica de calidad de un producto.

En problemas, donde la variable (o las variables) dependientes no son controlables, el objetivo puede ser predecir el valor de dicha variable dependiente (o respuesta). Esto puede ser interesante cuando la variable independiente es más fácil de medir que la variable dependiente o bien cuando la variable independiente esta disponible antes que la dependiente. En otros casos puede permitir sustituir un ensayo destructivo por uno no destructivo (JURAN y GRAYNA, 1997, cap. 23).

En síntesis la utilidad de los modelos de regresión puede ser la siguiente (JURAN y GRAYNA, 1997, cap. 23):

1. Proyecto y predicción
2. Descripción cuantitativa entre un conjunto de variables
3. Interpretación de los valores de la función
4. Determinación de variables independientes importantes
5. Descubrimiento de las condiciones de funcionamiento óptimas
6. Selección entre modelos alternativos
7. Estimación de coeficientes de regresión particulares

Las fases en un estudio mediante modelos de regresión se presentan en la figura 5.2.

En general, el análisis de regresión puede clasificarse en: *a) Regresión lineal y b) Regresión no lineal. Ambos a su vez se clasifican en regresión lineal (no lineal) simple y múltiple.*

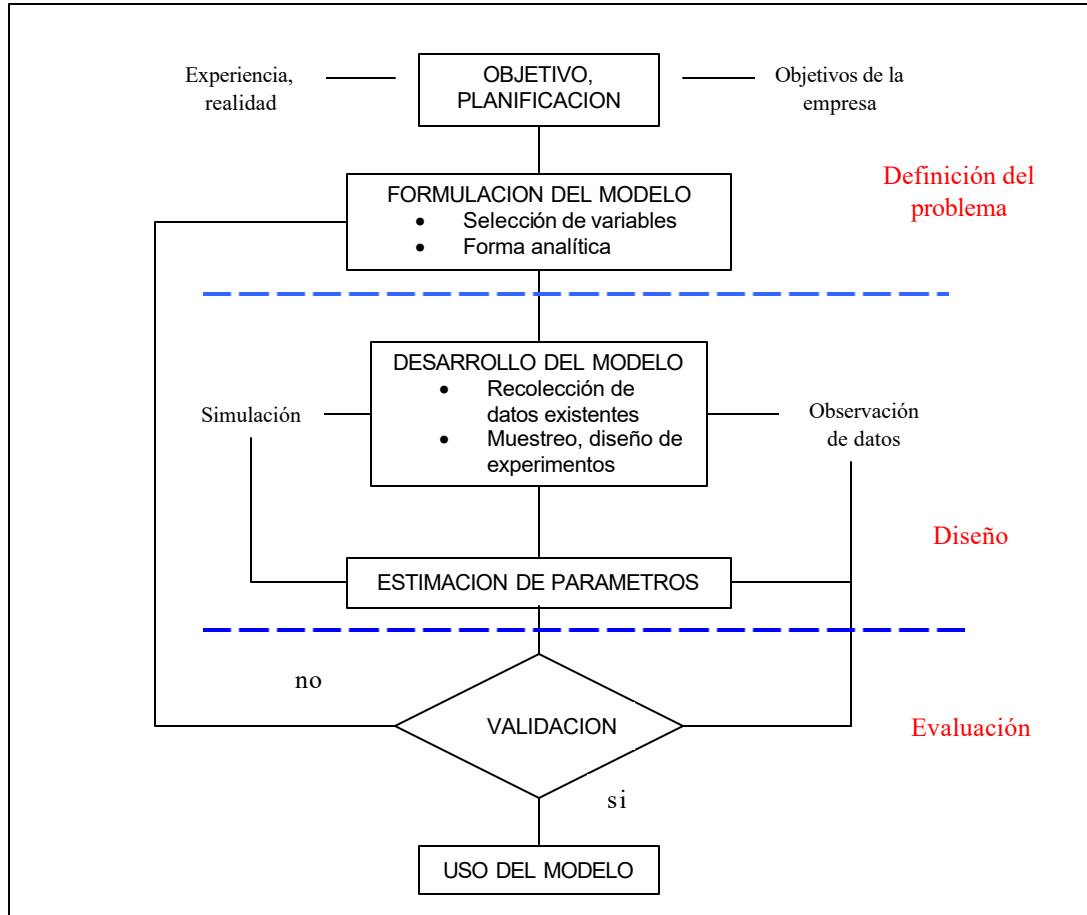


FIGURA 5.2 FASES EN EL ANALISIS DE REGRESION

5.1 REGRESION LINEAL SIMPLE

En este tipo de regresión se desea caracterizar el efecto lineal de una única variable explicativa sobre la variable respuesta.

Los pasos para efectuar un análisis son los siguientes (JURAN y GRANA, 1997, cap. 23):

1. Representación gráfica de datos
2. Planteamiento del modelo
3. Estimación de la ecuación de predicción
4. Examen de la adecuación del modelo lineal
5. Intervalos de confianza para la estimación

5.1.1 REPRESENTACION GRAFICA DE DATOS

El primer paso para el estudio de relaciones entre variables, consiste en trazar una gráfica de los datos (corrientemente llamado diagrama de dispersión). La variable respuesta se indica en el eje vertical y la variable explicativa en el horizontal. En la figura 5.3 se presentan ejemplos de gráficos de dispersión.

Los gráficos de dispersión son útiles debido a los siguientes aspectos (JURAN y GRYNA, 1997, cap. 23):

- Facilita información sobre la relación existente entre las variables
- Permite sugerir modelos posibles para los datos o transformaciones de datos.
- Puede señalar la existencia de observaciones anómalas
- Puede facilitar una indicación de y para x fija. Además puede mostrar que esa variabilidad permanece constante para todos los valores de x o que cambia con x .

5.1.2 PLANTEAMIENTO DEL MODELO

La ecuación que relaciona los datos para una regresión lineal simple es la siguiente (MONTGOMERY, 1991, cap. 15):

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (\text{ec. 5.1})$$

Donde: β_0 y β_1 son respectivamente la ordenada en el origen y el nivel de variación de y por cada unidad de variación de x (desconocidas de la recta de regresión); x variable explicativa o una transformación de esta (ej: $\log x$); ε es el error aleatorio con media cero y varianza constante (σ^2). También se supone que las $\{\varepsilon\}$ constituyen un conjunto de variables aleatorias independientes. El error aleatorio puede deberse a errores en la medición de y y/o a efectos de "variables no incluidas en el modelo". Se denota como *error de la ecuación*.

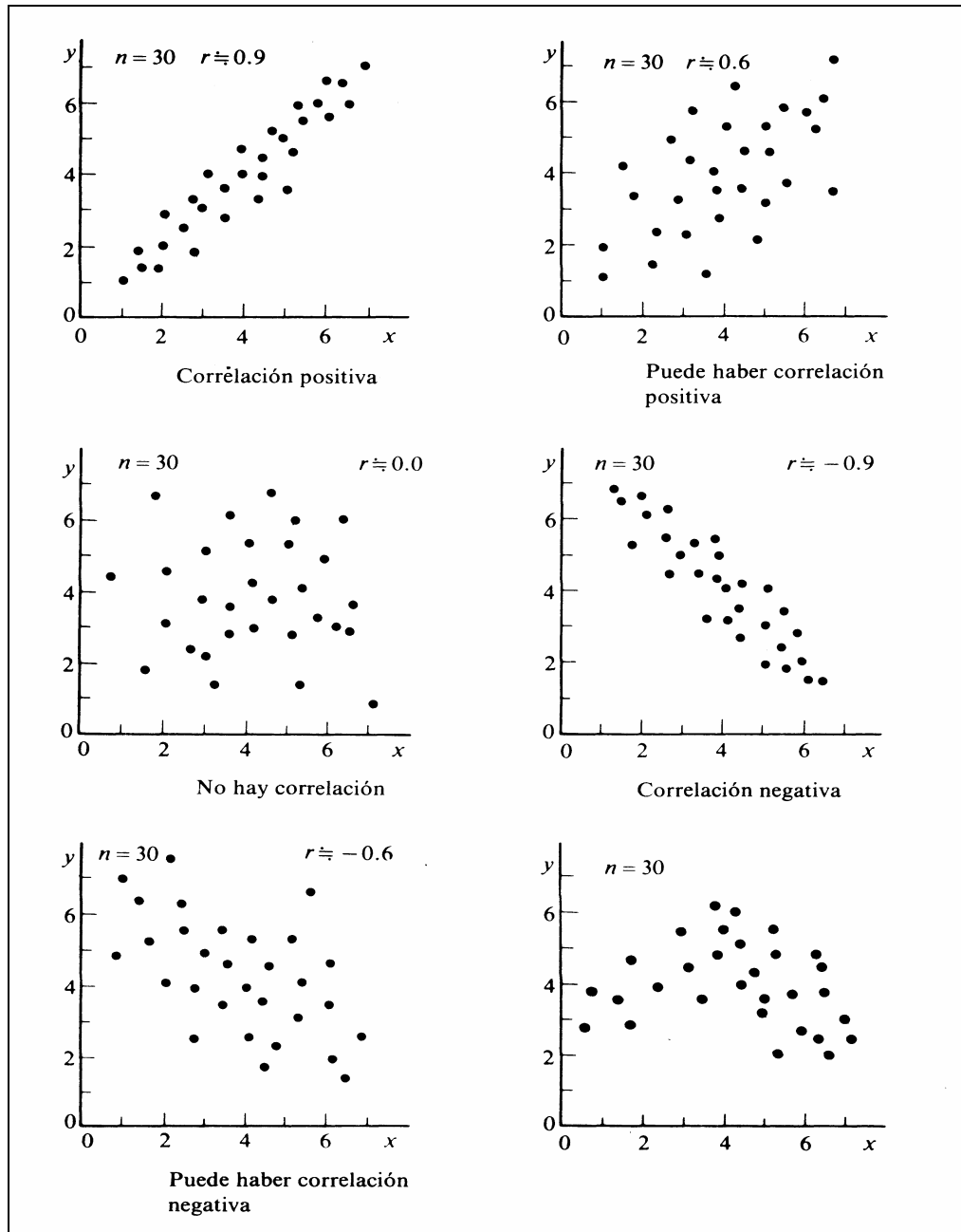


FIGURA 5.3 GRAFICOS DE DISPERSION (KUME, 1997, cap. 6)

5.1.3 ESTIMACION DE LA ECUACION DE PREDICCION

El objetivo de esta etapa es hallar estimadores (b_0 , b_1) de los parámetros desconocidos (β_0 , β_1), y obtener la ecuación de predicción en base a los datos (JURAN y GRZYNA, 1997, cap. 23):

$$\hat{y} = b_0 + b_1 x \quad (\text{ec. 5.2})$$

donde:

\hat{y} es el valor de y pronosticado por el modelo para un valor de x .

El método que proporciona los estimadores más eficientes, lineales no sesgados, b_0 y b_1 a partir de un conjunto de n observaciones $(x_1, y_1) \dots (x_n, y_n)$, es el método de los mínimos cuadrados, b_0 y b_1 se denominan por tanto estimadores mínimocuadráticos, porque son los que minimizan la suma de desviaciones elevadas al cuadrado entre los valores observados y los valores pronosticados por el modelo para la variable respuesta (JURAN y GRYNA, 1997, cap. 23):

$$\sum (y_m - \hat{y}_m)^2 = \sum (y_m - b_0 - b_1 x_m)^2 \quad (\text{ec. 5.3})$$

Las condiciones para la aplicación del método de los mínimos cuadrados para estimar los parámetros β_0 , β_1 , son las siguientes (JURAN y GRYNA, 1997, cap. 23):

1. Las observaciones son independientes
2. La varianza de los errores es constante para estas observaciones
3. El modelo lineal promulgado es correcto. Entonces los estimadores de mínimos cuadrados son "los mejores estimadores lineales no sesgados"

En la figura 5.4, se presenta el significado de la recta de regresión lineal.

Los estimadores mínimos cuadrados para el modelo lineal simple se estiman como sigue (JURAN y GRYNA, 1997, cap. 23):

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$b_1 = \frac{\sum (x_m - \bar{x})(y_m - \bar{y})}{\sum (x_m - \bar{x})^2} = \frac{S_{XY}}{S_{XX}} \quad (\text{ec. 5.4})$$

Unas formulas que permiten facilitar para el cálculo de S_{xy} y S_{xx} son las siguientes (MONTGOMERY, 1991, cap. 15):

$$S_{xy} = \sum_{j=1}^n x_j y_j - \frac{\left(\sum_{i=1}^n x_i \right) \left(\sum_{j=1}^n y_j \right)}{n} \quad (\text{ec. 5.5})$$

$$S_{xx} = \sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i \right)^2}{n}$$

Todas las sumatorias desde i ó $j = 1$ hasta i ó $j = n$.

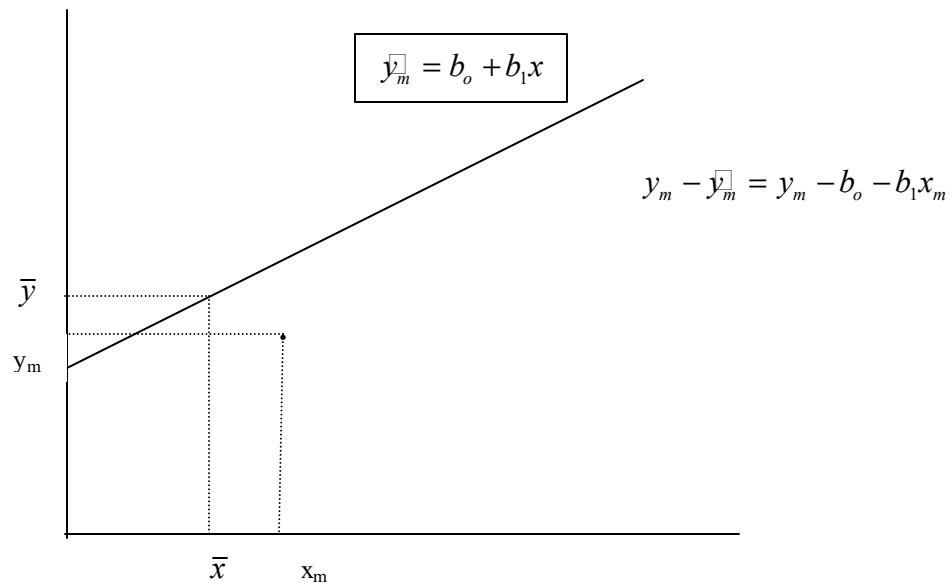


FIGURA 5.4 SIGNIFICADO DE LA RECTA DE REGRESION (JURAN y GRINA, 1997, cap. 23)

Por otra parte, \bar{y} y \bar{x} vienen dados por:

$$\bar{y} = \frac{\sum_{j=1}^n y_j}{n} \quad (\text{ec. 5.6})$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

5.1.4 EXAMEN DE ADECUACION DEL MODELO

El examen preliminar de adecuación del modelo, consiste en dibujar la ecuación. Aproximadamente la mitad de los puntos deben quedar por encima de la recta y la otra

mitad por debajo. Además la ecuación debe pasar exactamente por (\bar{x}, \bar{y}) (JURAN y GRZYNA, 1997, cap. 23).

A continuación del examen preliminar debe determinarse la significación global del ajuste, lo cual se logra a través del cálculo del coeficiente de regresión o determinación R^2 (JURAN y GRZYNA, 1997, cap. 23):

$$R^2 = \frac{SC(\text{regresion})}{S_{yy}} = \frac{\sum_{j=1}^n (\hat{y}_j - \bar{y})^2}{\sum_{j=1}^n (y_j - \bar{y})^2} = \frac{b_1 \sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (y_j - \bar{y})^2} \quad (\text{ec. 5.7})$$

Donde el numerador de la ecuación se refiere a la variación explicada por el modelo, y el denominador se refiere a la variación total. R^2 debe oscilar entre 0 y 1. Cuanto más cercano a 1, la mayor parte de la variabilidad constatada de "y" estará asociada a las variables explicativas incluidas en el modelo (MONTGOMERY, 1991, cap. 15).

Cuando la variable explicativa es aleatoria, la raíz cuadrada del coeficiente de determinación R^2 , es el coeficiente de correlación entre x e y (MONTGOMERY, 1991, cap. 15).

Para estudiar la hipótesis de significación del efecto de la variable explicativa (a nivel poblacional), se utiliza el resultado que se explica a continuación. Dado el modelo:

$$y = \beta_0 + \beta_1 x$$

La variable x influye en y si y solo si β_1 es significativamente distinto de cero.

Se puede probar el grado de significación de la regresión, es decir, se puede contrastar la hipótesis estadística (JURAN y GRZYNA, 1991, cap. 23):

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

La prueba para contrastar las hipótesis puede efectuarse de dos formas equivalentes:

a) FORMA DE PRUEBA 1 (ROMERO y ZUNICA, 1993, cap. 11):

$$\text{Si } \beta_1 = 0 \quad \Rightarrow \quad \frac{b_1}{S_{b1}} \rightarrow t_{n-2, \alpha/2}$$

$$\text{Si } \beta_1 \neq 0 \quad \left| \frac{b_1}{S_{b1}} \right| > t_{n-2, \alpha/2}$$

Y se deduce si x influye sobre y . S_{b1} es la varianza del estimador que se calcula como:

$$S_{b1} = S \sqrt{\frac{1}{S_{xx}}} \quad \text{con } S \text{ calculado como } S = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-2}} \quad (\text{ec. 5.8})$$

donde la sumatoria es de $i = 1$ hasta $i = n$.

La prueba de significación para b_0 ($H_0: \beta_0 = 0$ contra $H_1: \beta_0 \neq 0$) viene dada como:

$$\left| \frac{b_0}{S_{b0}} \right| \rightarrow t_{n-2, \alpha/2}, \text{ si se cumple } H_0 \text{ y } \left| \frac{b_0}{S_{b0}} \right| > t_{n-2, \alpha/2}, \text{ si } H_0 \text{ es falsa, con}$$

$$S_{b0} = S \sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)}.$$

En la tabla 5.1 se presenta un resumen para esta prueba.

TABLA 5.1 RESUMEN PRUEBA 1 DE SIGNIFICANCIA DE COEFICIENTES

Coeficiente	valor	desv. Estándar del coeficiente	t_o
β_0	b_0	$S_{b0} = S \sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)}$	$\frac{b_0}{S_{b0}}$
β_1	b_1	$S_{b1} = S \sqrt{\frac{1}{S_{xx}}}$	$\frac{b_1}{S_{b1}}$

b) FORMA DE PRUEBA 2 (ROMERO y ZUNICA, 1993, cap. 11):

Sean:

$CM_{residual}$, el cuadrado medio residual del ajuste con la variable x con $n - 2$ grados de libertad, y calculado como:

$$CM_{residual} = \frac{SC_{residual}}{n-2} = \frac{(S_{yy} - b_1 S_{xy})}{n-2} \quad (\text{ec. 5.9})$$

Consideramos el valor del cuadrado medio explicado ($CM_{regresion}$), por el modelo incluyendo x , con un grado de libertad y calculado como:

$$CM_{regresion} = \frac{SC_{regresion}}{1} = \frac{b_1 S_{xy}}{1} \quad (\text{ec. 5.10})$$

$$\text{Si } \beta_1 = 0 \quad \Leftrightarrow \quad \frac{CM_{regresion}}{CM_{residual}} \rightarrow F_{\alpha,1,n-2}$$

$$\text{Si } \beta_1 \neq 0 \quad \frac{CM_{regresion}}{CM_{residual}} > F_{\alpha,1,n-2}$$

Y se deduce si x influye sobre y . Con:

$$SC_{regresion} = SC_{total} - SC_{residual}$$

En la tabla 5.2 se presenta un resumen para esta prueba.

TABLA 5.2 RESUMEN DE PRUEBA 2 DE SIGNIFICACION DEL MODELO DE REGRESION

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	F_o
regresión	$b_1 S_{xy}$	1	$CM_{regresion}$	$\frac{CM_{regresion}}{CM_{residual}}$
residual	$S_{yy} - b_1 S_{xy}$	n - 2	$CM_{residual}$	
total	S_{yy}	n - 1		

5.1.4.1 DESVIACIONES Y FALTA DE AJUSTE. FORMA GRAFICA

Las desviaciones y falta de ajuste del modelo lineal pueden determinarse a partir del análisis de los residuos (MONTGOMERY, 1991, cap. 15).

Si en los gráficos de los residuos contra los valores predichos y contra la variable de regresión, se observa una dispersión aleatoria de puntos por encima y por debajo de la línea de desviación cero, puede concluirse que la falta de ajuste no es significativa y por lo

tanto adoptar el modelo lineal planteado. Por otra parte, si en el gráfico de residuos contra valores predichos por el modelo lineal se observan tendencias como curvaturas, tendencias de aumento o disminución de los valores de los residuos al aumentar los valores predichos, etc., puede concluirse que existe una falta de ajuste significativa y debe ajustarse otro modelo, como un modelo simple no lineal, o bien un modelo múltiple. Otro gráfico útil es el de normalidad de los residuos (ROMERO y ZUNICA, 1993, cap. 11).

5.1.5 CREACION DE INTERVALOS DE CONFIANZA

Los intervalos de confianza sobre los parámetros proporcionan otra medida de la fiabilidad de los estimadores y de las estimaciones (JURAN y GRZYNA, 1997, cap.23).

El intervalo de confianza para β_1 , viene dado a partir de (JURAN y GRZYNA, 1997, cap.23) (ROMERO y ZUNICA, 1993, cap. 11):

$$b_1 \pm \frac{tS}{\sqrt{\sum (x_i - \bar{x})^2}} \equiv b_1 \pm tS_{b_1} \quad (\text{ec. 5.11})$$

Donde $t_{\alpha/2, n-2}$, el término $\frac{S}{\sqrt{\sum (x_m - \bar{x})^2}}$ se denomina error del coeficiente de regresión

(ó S_{b_1}). Similarmente un intervalo de confianza del $100(1 - \alpha)\%$ para β_0 es dado por (JURAN y GRZYNA, 1997, cap.23) (ROMERO y ZUNICA, 1993, cap. 11):

$$b_0 \pm tS \sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)} \equiv b_0 \pm tS_{b_0} \quad (\text{ec. 5.12})$$

El intervalo de confianza para el valor medio de y correspondiente a un valor dado de x viene dado por (JURAN y GRZYNA, 1997, cap.23):

$$b_0 + b_1x \pm tS \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{S_{xx}}} \quad (\text{ec. 5.13})$$

Donde x es el valor para el que se determina el intervalo de confianza, y t tiene $n-2$ grados de libertad.

5.1.2 CONFIABILIDAD DE UNA NUEVA PREDICCIÓN

El intervalo de predicción del 100(1 - α)% para la media de k observaciones futuras en x_o es (JURAN y GRZYNA, 1997, cap.23) (MILLER y MILLER, 1993, cap. 5):

$$b_o + b_1 x_o \pm tS \sqrt{\frac{1}{k} + \frac{1}{n} + \frac{(x_o - \bar{x})^2}{S_{xx}}} \quad \text{con } t_{\alpha/2, n-2} \quad (\text{ec. 5.14})$$

5.2 APLICACIONES DE LA REGRESIÓN LINEAL SIMPLE EN EL ANÁLISIS QUÍMICO

Dentro de las aplicaciones de las rectas de regresión al análisis químico se encuentran:

- a) Gráficas de calibración en análisis instrumental.
- b) Determinación de la incertidumbre asociada al cálculo de una concentración.
- c) Método de las adiciones estándar.
- d) Rectas de regresión ponderadas.
- e) Uso de rectas de regresión para comparar métodos analíticos.

A continuación se presenta un resumen de estas aplicaciones.

A) GRÁFICAS DE CALIBRACIÓN EN ANÁLISIS INSTRUMENTAL

La capacidad de las técnicas de análisis instrumental para manejar un intervalo amplio de concentraciones de analito significa que se calculan los resultados, y se evalúan los errores aleatorios de manera concreta, que difiere de la que se utiliza cuando se repite una sola medición varias veces. El procedimiento habitual es como sigue: El analista toma una serie de muestras (normalmente tres o cuatro, o posiblemente más) en las que se conoce la concentración del analito, valores "x". Estas calibraciones estándar se miden en el instrumento analítico en las mismas condiciones que las usadas para la muestra problema (es decir, las "desconocidas"), obteniéndose las mediciones "y". Una vez que se ha establecido la ecuación de calibración, se puede obtener la concentración del analito en

cualquier muestra problema por predicción. Este procedimiento general plantea varias preguntas estadísticas importantes (MILLER y MILLER, 1993, cap. 5):

- I. ¿Es lineal la gráfica de calibración? Si es curva ¿Qué forma tiene?
- II. Teniendo en cuenta que cada uno de los puntos de la línea de calibración está sujeto a errores, ¿cuál es la mejor línea recta (o curva) que pasa por esos puntos?
- III. Suponiendo que la calibración es realmente lineal, ¿Cuáles son los errores estimados y los límites de confianza para la pendiente y la ordenada en el origen de la recta?
- IV. Cuando la gráfica de calibración se usa para el análisis de una muestra problema, ¿Cuáles son los errores y los límites de confianza para la concentración determinada?
- V. ¿Cuál es el límite de detección del método? Esto es, ¿Cuál es la menor concentración de analito que se puede detectar con un nivel de confianza predeterminado?

Todas las preguntas anteriores pueden responderse mediante el análisis de regresión simple, que se ha expuesto en el tema 5.1. En los ejercicios de la sección 5.4.4, se presenta esta aplicación al análisis químico.

B) EL METODO DE LAS ADICIONES PATRON (O DE ADICION ESTANDAR)

En el método de adiciones estándar, se prepara una curva de calibración de forma que todos los patrones contengan muestra, y la contengan en la misma cantidad. El primer punto de la curva de calibración, es una disolución de la muestra a la que no se ha añadido patrón, y por tanto, contiene la cantidad de analito original de la muestra. El segundo punto de la curva de calibración contiene esa misma cantidad de analito más una cantidad conocida de patrón que se le ha añadido. El tercer punto de la curva contiene de nuevo el analito original de la muestra, mas una cantidad mayor de patrón, y así sucesivamente. Por otra parte, debe prepararse y medirse cuidadosamente un blanco. Luego se obtienen las lecturas en el instrumento (datos de calibrado). Con los datos de

calibrado se hace la representación gráfica de la figura 5.5, y se obtiene la correspondiente recta de regresión (RAMIS, 1997, tema 5).

Es muy importante tener en cuenta que el origen de las coordenadas debe colocarse en la lectura del blanco, o alternatively, la señal del blanco debe restarse de todos los demás valores de la señal. Si se utiliza la señal dada por el disolvente puro, o lo que es peor, el valor cero de la señal dada por el instrumento, se comete un error sistemático que puede ser muy grande. El resultado buscado, x_E (E valor extrapolado) la concentración del analito en la muestra problema, se puede leer sobre la propia gráfica, extrapolando el eje de las abscisas. También puede calcularse como $x_E = b_0/b_1$, esto es, la concentración de analito en la muestra es igual a la razón entre la ordenada en el origen y la pendiente. La desviación estándar del resultado se calcula mediante (RAMIS, 1997, tema 5):

$$s_E = \frac{S}{b_1} \left[\frac{1}{n} + \frac{\bar{y}^2}{b_1^2 \sum (x_i - \bar{x})^2} \right]^{1/2} \quad (\text{ec. 5.15})$$

Donde n es el número de puntos de la recta de adición estándar. De la expresión anterior puede concluirse lo siguiente (RAMIS, 1997, tema 5):

- Igual que en el caso de la curva de calibrado ordinario, la precisión se mejora si los puntos están poco dispersos (s_{xy} es pequeño) y si b_1 es grande.
- Al aumentar el valor de n mejora la precisión, pero no se obtendrán mejoras importantes cuando ya se tengan al menos 6 puntos. Con frecuencia se trabaja con 3 o 4 puntos.
- La sumatoria del denominador del último término indica que la precisión se mejora cuando los valores de x_i se alejan del centroide, esto es, cuando se aumenta el intervalo de valores de concentración abarcado. Sin embargo, no es conveniente cubrir un intervalo excesivamente grande, pues pueden aparecer desviaciones de la linealidad, lo que producirá un error sistemático.

El método de adición estándar se utiliza para descubrir y corregir la presencia de efecto de matriz. Sin embargo, por sí solo, el método de adición estándar no sirve para validar un método, puesto que (RAMIS, 1997, tema 5):

- a) No es capaz de descubrir ni corregir las interferencias.
- b) No siempre corrige el efecto de matriz en un grado suficientemente satisfactorio.
- c) Mediante la sola aplicación del método de adición estándar no se puede establecer si las interferencias y el efecto matriz han sido corregidos satisfactoriamente.

Existe efecto de matriz cuando la recta obtenida por el método de adición estándar no tiene la misma pendiente que la recta obtenida por calibración externa.

Las limitaciones del método de adiciones estándar se pueden resumir como sigue (RAMIS, 1997, tema 5):

- a) Es un método de extrapolación, y por lo tanto es menos preciso que el calibrado por interpolación (la medida del analito se realiza muy lejos del centroide). Si la curva de calibrado no es lineal, el método de adición estándar, que está basado en una extrapolación lineal, no puede aplicarse.
- b) Se gasta más muestra, lo que puede ser un problema cuando se dispone de pequeñas cantidades.
- c) Es necesario repetir el calibrado para cada muestra o tipo de muestras, o en cualquier circunstancia en que se sospeche que ha habido un cambio del efecto matriz.
- d) Como ya se ha indicado, no es capaz de descubrir ni corregir las interferencias, y no siempre corrige el efecto matriz en un grado suficiente.

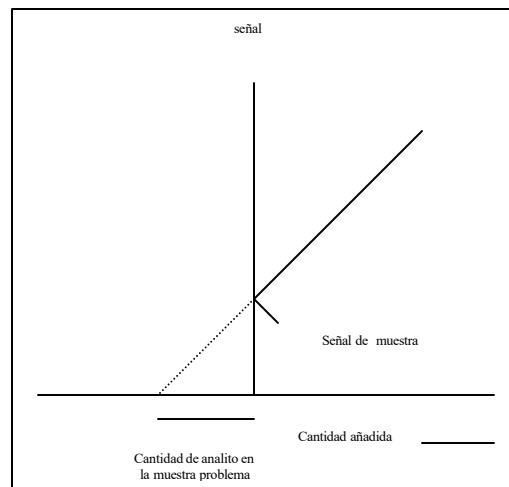


FIGURA 5.5 METODO DE LAS ADICIONES ESTANDAR

C) USO DE RECTAS DE REGRESION PARA COMPARAR METODOS ANALITICOS

Supóngase que se ha desarrollado un método nuevo para determinar un analito en una matriz dada. Para validar la incertidumbre y la exactitud del método este se aplica a un material de referencia certificado. Finalmente, se realizarán pruebas estadísticas para saber si el método tiene mayor incertidumbre que la deseada, o si existe error sistemático respecto al valor certificado (RAMIS, 1997, tema 5).

Sin embargo, el proceso de validación cambia cuando se quiere validar un método a varios niveles de concentración, y no a uno solo. Para ello, será necesario utilizar varios materiales de referencia, conteniendo distintos niveles de concentración de analito, o también, aplicar el método a un conjunto de muestras ya estudiadas por otro procedimiento estándar o de prestigio.

Cuando se comparan dos métodos a diferentes concentraciones de analito, se adopta normalmente el procedimiento que ilustra la figura 5.6. Uno de los ejes de la gráfica de regresión se utiliza para los resultados obtenidos con el nuevo método, y el otro para los que se obtienen aplicando el de comparación o referencia a las mismas muestras. Así cada punto de la gráfica representa una única muestra analizada por dos métodos separados (RAMIS, 1997, tema 5).

Si cada muestra conduce a un resultado idéntico en ambos métodos para todos los niveles de concentración estudiados, la recta de regresión tendrá una ordenada en el origen igual a cero, y una pendiente y un coeficiente de correlación iguales a 1. Cualquier resultado que no sea este, implica error sistemático de un método con respecto al otro, a algunos o a todos los niveles de concentración. En la práctica, debido a la presencia de los errores aleatorios, nunca serán exactamente cero la ordenada en el origen ni uno la pendiente. Sin embargo, la sola observación del gráfico dará una idea clara acerca del mantenimiento o no de la exactitud a lo largo del intervalo de concentración estudiado. Se distinguirán varios casos (RAMIS, 1997, tema 5):

- a) La recta de regresión tiene una pendiente de 1, pero la ordenada en el origen no es cero. Este resultado implica que al menos uno de los métodos sufre una interferencia, o lo que es igual, la señal de fondo de uno de los métodos no está

bien medida, esto es, al menos uno de los blancos o los dos no esta bien construido.

- b) La pendiente no es la unidad, o también, la representación gráfica de la curva. Entonces existe error sistemático de un método, y error se agrava cuanto más se alejan los puntos de la recta $y = x$.
- c) Los errores a) y b) suceden simultáneamente.

Para validar un método con respecto al otro (o con respecto a los materiales de referencia utilizados) se establecen las dos hipótesis nulas siguientes: $\beta_0 = 0$ y $\beta_1=1$. Para comprobar estas hipótesis se aplican los siguientes ensayos estadísticos: se establecen los límites de confianza para b_0 y b_1 , para un nivel de confianza del 95% y se comprueba si el cero y el uno están dentro de los correspondientes intervalos de confianza. El valor del coeficiente de determinación tiene menos importancia, no dice nada respecto a la exactitud de un método respecto al otro, si los puntos están dispersos ($R^2 < 0.99$) es indicativo que al menos uno de los dos métodos tiene una precisión pobre (RAMIS, 1997, tema 5).

Las limitaciones de este método de validación de la exactitud son las siguientes (RAMIS, 1997, tema 5):

- a) Este método constituye un ejemplo típico de los casos en los que no está justificada la utilización de una recta de regresión de Y sobre X, puesto que los errores en la dirección de X pueden ser tan importantes o más que en la dirección Y. Sin embargo, los resultados pueden ser confiables si se toman las precauciones siguientes:
 1. El método más preciso se representa en el eje X. Para ello se debe establecer previamente la precisión de ambos métodos.
 2. Utilizar un número razonable de puntos (al menos 10). Ya que los límites de confianza se basan en $n - 2$ grados de libertad, es importante evitar valores pequeños de n .
 3. Los puntos experimentales deberían cubrir el intervalo de concentraciones de interés de una manera aproximadamente uniforme, esto es, debería haber tantos puntos a concentraciones bajas como a intermedias y altas,

dentro del intervalo estudiado. Este requisito debe ser difícil de cumplir cuando se trabaja con muestras reales.

- b) Si no existe homocedasticidad debería utilizarse la recta de regresión ponderada. La falta de homocedasticidad es muy frecuente en la práctica: para muchos procedimientos analíticos, la desviación estándar relativa (coeficiente de variación) es aproximadamente constante en un intervalo de concentraciones, por tanto, la desviación estándar en valor absoluto aumenta con la concentración de analito (en vez de tener el mismo valor para todas las concentraciones).

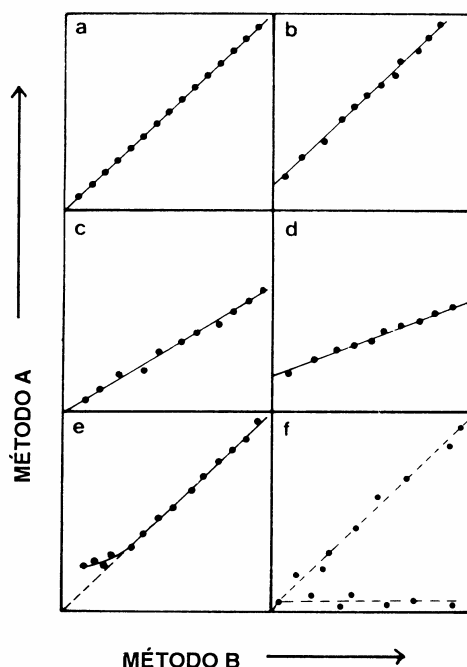


FIGURA 5.6 USO DE RECTAS DE REGRESI ON PARA COMPARAR METODOS ANALI TI COS: a) MUESTRA UNA CONCORDANCIA PERFECTA ENTRE LOS DOS METODOS PARA TODAS LAS MUESTRAS b) a f) ILUSTRA LOS RESULTADOS DE VARI S TI POS DE ERRORES SI STEMATI COS (MILLER y MILLER, 1993, cap. 5)

D) RECTAS DE REGRESI ON PONDERADAS

Los cálculos de regresión ponderada se utilizan para la determinación de la concentración de un único analito y no para comparar dos métodos separados (MILLER y MILLER, 1993, cap. 5).

Los cálculos de regresión ponderada requieren más información (estimaciones de la desviación estándar a varios niveles de concentración) y son más complejos de ejecutar. Por otra parte, aparentemente proporcionan datos muy similares a los obtenidos a partir del método, mucho más simple, de regresión no ponderada. Dichas consideraciones de hecho pueden justificar que en la práctica se desprecien los cálculos de regresión ponderada (RAMIS, 1997, tema 5).

Pero un químico analítico que utiliza métodos instrumentales no emplea cálculos de regresión simplemente para determinar la pendiente y la ordenada en el origen de una grafica de calibración y las concentraciones de las muestras problema. También quiere obtener estimaciones de los errores o límites de confianza de esas concentraciones, y en este contexto el método de regresión ponderada ofrece una gran ventaja respecto al método no ponderado si los puntos experimentales son claramente heterocedásticos: las estimaciones de las desviaciones estándar de la pendiente y la ordenada en el origen son mucho más realistas (RAMIS, 1997, tema 5).

Supóngase que se lee una concentración sobre una recta de regresión que tiene un error relativo constante. Puesto que la desviación estándar absoluta aumenta con la concentración, no es lógico que se obtenga un error constante a cualquier nivel de concentración, sino que la lectura de concentraciones altas debe dar errores mayores que la lectura de concentraciones bajas. Si se ha utilizado una recta de regresión ponderada, el error de la lectura $s_{x_0,w}$ variará dependiendo de la concentración. Las expresiones que permiten calcular $s_{x_0,w}$ son (RAMIS, 1997, tema 5):

$$s_{xy,w} = \left[\frac{\sum (w_i y_i^2 - n \bar{y}_w^2) - b_1^2 (\sum w_i x_i^2 - n \bar{x}_w^2)}{n - 2} \right]^{1/2} \quad (\text{ec. 5.16})$$

$$s_{x_0,w} = \frac{s_{xy,w}}{b_1} \left[\frac{1}{w_0} + \frac{1}{n} + \frac{(y_0 - \bar{y}_w)^2}{b_1^2 (\sum w_i x_i^2 - n \bar{x}_w^2)} \right]^{1/2}$$

En donde:

$$\begin{aligned}
 w_i &= \frac{s_i^{-2}}{\left(\sum s_i^{-2} / n\right)} \\
 b_1 &= \frac{\sum w_i x_i y_i - n \bar{x}_w \bar{y}_w}{\sum w_i x_i^2 - n \bar{x}_w^2} \quad (\text{ec. 5.17}) \\
 \bar{x}_w &= \frac{\sum w_i x_i}{n} \\
 \bar{y}_w &= \frac{\sum w_i y_i}{n}
 \end{aligned}$$

la expresión para $s_{x_0, w}$ es en todo similar a su análoga para la regresión no ponderada. Ello confirma que los puntos cercanos al origen, donde las ponderaciones son más altas, y los cercanos al centro de gravedad, donde $y_0 - \bar{y}_w$ es pequeño, tendrán límites de confianza pequeños. La principal diferencia entre las dos ecuaciones es la introducción del término $1/w_0$ (donde w_0 es la ponderación de y_0), que aumenta rápidamente cuando aumenta y . Por tanto se obtienen límites de confianza mayores, a valores altos de concentración. Esto concuerda mucho más con el experimento de calibración real que los resultados obtenidos al utilizar la regresión no ponderada. Además los métodos de regresión ponderados pueden ser absolutamente necesarios cuando se obtenga una recta por transformaciones algebraicas de una representación intrínsecamente curva (RAMIS, 1997, tema 5).

E) INCERTIDUMBRE ASOCIADA AL USO DE UNA RECTA DE REGRESION

De las ecuaciones que permiten calcular las desviaciones estándar y los límites de confianza de: 1) La pendiente, 2) La ordenada en el origen, y 3) la confiabilidad de una nueva predicción, se extraerán algunas consecuencias de orden práctico referentes a la forma como debe prepararse y utilizarse una curva de calibrado (RAMIS, 1997, tema 5). Supóngase que se dispone ahora de un valor experimental de "y" (o un promedio de k replicas) al que se denominará y_0 . Se lee este valor, y se hace uso de la ecuación de regresión para obtener el correspondiente valor de "x", x_0 . La desviación estándar de x_0 , s_{x_0} , viene dada por (RAMIS, 1997, tema 5):

$$s_{x_o} = \frac{S}{b_1} \left[\frac{1}{k} + \frac{1}{n} + \frac{(y_o - \bar{y})^2}{b_1^2 \sum (x_i - \bar{x})^2} \right]^{1/2} \quad (\text{ec. 5.18})$$

5.3 REGRESION LINEAL MULTIPLE

Aunque hay muchos problemas que atañen a variables explicativas únicas, es muy frecuente que aparezcan varias. El problema de regresión lineal múltiple consiste en ajustar el modelo general (JURAN y GRZYNA, 1997, cap. 23):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (\text{ec. 5.19})$$

Donde los β son los parámetros desconocidos, llamados coeficientes de regresión, y ε es el error aleatorio.

Las variables x_1, x_2, \dots, x_k pueden ser :

- La variable original cuantitativa.
- Transformaciones de la variable original: polinómicas, logarítmicas, etc.

Ejemplo: $\bar{y} = a + bx_1 + cx_1^2$, puede definirse una nueva variable $x_2 = x_1^2$ y se obtiene el modelo lineal:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

- Variables cualitativas con p variantes, deben incluirse en este caso lo que se conoce como variable "dummy" (p -1) cuyos únicos valores probables son 0 y 1. Generalmente se simbolizan como z_i . Por ejemplo, para incluir el efecto lineal de una variable cuantitativa x_1 y de otra cualitativa x_2 con dos posibles variantes, el modelo sería:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 z_1 + \beta_3 z_2$$

- Inclusión de interacciones: es posible que se den efectos interactivos entre las variables explicativas. Este efecto debe incluirse como una nueva variable que sea

el producto de las variables implicadas en dicha interacción. Por ejemplo, una interacción entre los efectos lineales de dos variables cuantitativas x_1 y x_2 implica crear un nuevo término: $x_1x_2 = x_3$.

En general en el modelo lineal, se asume que los errores son normales e idénticamente distribuidos con media cero y varianza constante es decir: $\varepsilon \sim N(0, \sigma^2)$.

La interpretación de los parámetros del modelo es la siguiente (ROMERO y ZUNICA, 1993, cap. 11):

β_0 valor medio de "y" cuando $x_1 = x_2 = \dots = x_k = 0$

β_i incremento en el valor medio de "y" cuando " x_i " aumenta en una unidad, si x_i es una variable cuantitativa, manteniéndose constantes las restantes variables explicativas.

En general, los pasos para el análisis de regresión lineal múltiple son los siguientes (JURAN y GRZYNA, 1997, cap. 23):

1. Planteamiento del modelo.
2. Proceso de estimación.
3. Examen de adecuación del modelo.
4. Estimación de intervalos de confianza.

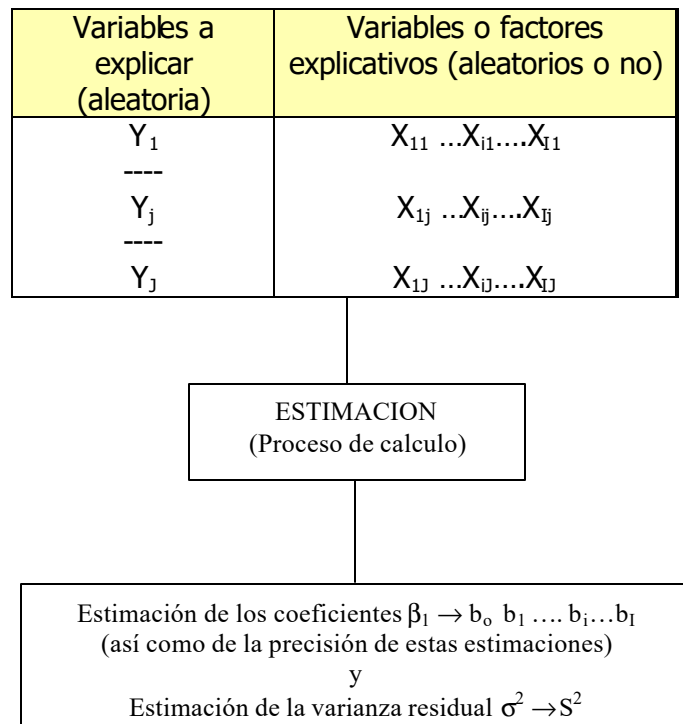
5.3.1 ESTIMACION EN MODELOS DE REGRESION

El esquema general del proceso de estimación se presenta a continuación (ROMERO y ZUNICA, 1993, cap. 11):

Dado el modelo:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

Y dados los datos:



Son conocidas para cada observación j ($j = 1, \dots, n$)

y_j : valor de la variable dependiente

$x_{1j}, \dots, x_{ij}, \dots, x_{lj}$: valores de las variables explicativas

El residuo ε_j que se obtendría para los posibles valores b_0, \dots, b_l de los coeficientes se define como (ROMERO y ZUNICA, 1993, cap. 11):

$$\varepsilon_j = y_j - (b_0 + b_1 x_1 + \dots + b_l x_{lj})$$

Se demuestra que los estimadores b_0, b_1, \dots, b_k óptimos desde el punto de vista de sus propiedades estadísticas son los que conducen a un valor mínimo de $\sum \varepsilon_j^2$ (ROMERO y ZUNICA, 1993, cap. 11), lo cual puede lograrse por el método de los mínimos cuadrados.

Resultados matemáticos: De los datos (ROMERO y ZUNICA, 1993, cap. 11):

$$\bar{y} = \begin{Bmatrix} y_1 \\ \vdots \\ y_j \\ \vdots \end{Bmatrix} \quad \bar{x} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1n} \\ \vdots & \vdots & & \vdots \\ 1 & x_{1j} & \cdots & x_{jn} \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

La estimación de los coeficientes de regresión viene dada por (ROMERO y ZUNICA, 1993, cap. 11):

$$\bar{b} = (\bar{x}'\bar{x})^{-1}\bar{x}'\bar{y} \quad (\text{ec. 5.20})$$

La estimación de la varianza residual es obtenida por (ROMERO y ZUNICA, 1993, cap. 11):

$$S^2 = \frac{\sum_1^N \varepsilon_j^2}{n - (k + 1)} = \frac{\sum_1^N (y_j - (b_o + b_1x_1 + \dots + b_kx_{kj}))^2}{n - (k + 1)} \quad (\text{ec. 5.21})$$

En donde k representa el número de variables explicativas y n el número de puntos para el análisis.

La estimación de la varianza de los coeficientes de regresión viene dada por (ROMERO y ZUNICA, 1993, cap. 11):

$$S_{b_i} = S\sqrt{c_{ii}} \quad \text{donde } c_{ii} \text{ es el elemento } i, i \text{ de la diagonal principal } (\bar{x}'\bar{x})^{-1}$$

5.3.2 SIGNIFICACION GLOBAL DEL AJUSTE

La variabilidad total de la variable dependiente "y" en el conjunto de las n observaciones, viene dada por (ROMERO y ZUNICA, 1993, cap. 11):

$$SC_{TOTAL} = \sum_{j=1}^{j=n} (y_j - \bar{y})^2 \quad \text{con } n - 1 \text{ grados de libertad} \quad (\text{ec. 5.22})$$

parte de la variabilidad es debida (o al menos esta asociada) a las variables explicativas x_1, \dots, x_k . Esta parte explicada tiene k grados de libertad.

El resto de la variabilidad estará recogida en los residuos, viniendo medida su magnitud por la suma de cuadrados residual (ROMERO y ZUNICA, 1993, cap. 11):

$$SC_{RESIDUAL} = \sum_{j=1}^n \epsilon_j^2 \quad \text{con } (n-1) - k \text{ grados de libertad} \quad (\text{ec. 5.23})$$

La diferencia:

$$SC_{EXPLICADA} = SC_{TOTAL} - SC_{RESIDUAL} \quad (\text{ec. 5.24})$$

es la parte de la variabilidad de "y" asociada a las variables explicativas.

Se define el coeficiente de determinación R^2 como (ROMERO y ZUNICA, 1993, cap. 11):

$$R^2 = \frac{SC_{EXPLICADA}}{SC_{TOTAL}} \quad (\text{ec. 5.25})$$

Cuyos valores están comprendidos entre 0 y 1. Cuanto más cercano a 1 sea R^2 , mayor parte de la variabilidad constatada de "y" estará asociada a las variables explicativas incluidas en el modelo (ROMERO y ZUNICA, 1993, cap. 11).

Para estudiar la hipótesis nula de si al menos una de las variables explicativas estudiadas tiene efecto poblacional o si al menos un β_i es diferente de cero, se usa el siguiente resultado (ROMERO y ZUNICA, 1993, cap. 11):

Si:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

Entonces,

$$\frac{SC_{EXPLICADA}/k}{SC_{RESIDUAL}/(n-1-k)} = \frac{CM_{EXPLICADA}}{CM_{RESIDUAL}} \sim F_{\alpha, k, (n-1-k)}$$

H_0 se rechaza si $F_0 > F_{\alpha, k, (n-1-k)}$; H_1 al menos un $\beta_i \neq 0$

5.3.3 SIGNIFICACION DEL EFECTO DE UNA VARIABLE X_i

Dado el modelo:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon$$

La variable x_i no influye sobre y si y solo sí $\beta_i = 0$

La prueba para contrastar las hipótesis:

$$H_0: \beta_i = 0$$

y

$$H_1: \beta_i \neq 0$$

que implica un efecto poblacional de x_i sobre y , puede realizarse de dos formas equivalentes (ROMERO y ZUNICA, 1993, cap. 11).

a) *FORMA DE PRUEBA 1* (ROMERO y ZUNICA, 1993, cap. 11):

$$\text{Si } \beta_i = 0 \text{ entonces } \frac{b_i}{Sb_i} \rightarrow t_{\alpha/2, n-1-k}$$

Por tanto si $\left| \frac{b_i}{Sb_i} \right| > t_{\alpha/2, n-1-k}$ debe rechazarse H_0 , y se deduce que x_i influye sobre y .

b) *FORMA DE PRUEBA 2* (ROMERO y ZUNICA, 1993, cap. 11):

Sean:

$CM_{residual}$, el cuadrado medio residual del ajuste con todas las variables con $n-1-k$ grados de libertad. Calculado a partir de:

$$CM_{residual} = \frac{SC_{residual}}{n-1-k} = \frac{\bar{y}'\bar{y} - b'x'y}{n-1-k} \quad (\text{ec. 5.26})$$

$\Delta SC_{residual}$, el incremento de la suma de cuadrados residual si se repite el ajuste sin incluir a x_i entre las variables explicativas.

Si $\beta_i = 0$, entonces $\Delta SC_{residual} / CM_{residual}$ se distribuye como una $F_{\alpha, 1, n-1-k}$, por tanto se rechaza H_0 si el cociente anterior es mayor que la variable F .

Mientras que si $\beta_i \neq 0$, la suma de cuadrados residual se incrementará significativamente si se elimina la variable x_i del modelo, por tanto:

$$(\Delta SC_{residual} / CM_{residual}) > F_{\alpha, 1, n-1-k}$$

Y el efecto de x_i es significativo.

5.3.4 PRUEBA GENERAL PARA EL CONTRASTE DE HIPÓTESIS EN MODELOS DE REGRESIÓN.

Sea un modelo lineal:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

Se desea contrastar la hipótesis nula que implica "r" restricciones sobre los valores de β (ROMERO y ZUNICA, 1993, cap. 11)

Ejemplos:

1. $H_0: \beta_2 = 0 \quad \beta_3 = 0 \quad \beta_4 = 0$ (3 restricciones)
2. $H_0: \beta_3 = \beta_1$ (1 restricción)

Un procedimiento general para contrastar hipótesis de este tipo consiste en estudiar la significación estadística del $\Delta SC_{\text{residual}}$ que se produce al restringir el modelo obligando a que los estimadores b_i verifiquen la hipótesis. De tal forma que (ROMERO y ZUNICA, 1993, cap. 11):

	Suma de cuadrados residual	Grados de libertad
Modelo restringido	SC_r	$n - 1 - k + r$
Modelo completo	SC_c	$n - 1 - k$
Incremento	$\Delta SC_{\text{RESIDUAL}} = SC_r - SC_c$ (siempre > 0)	r

Se demuestra que si H_0 es cierta (implicando las "r" restricciones) (ROMERO y ZUNICA, 1993, cap. 11)

$$\frac{\Delta SC_{\text{RESIDUAL}} / r}{SC_c / (N - 1 - k)} \sim F_{\alpha, r, n-1-k}$$

Si H_0 es falsa, entonces (ROMERO y ZUNICA, 1993, cap. 11),

$$\frac{\Delta SC_{RESIDUAL}/r}{SC_c/(N-1-k)} > F_{\alpha, r, n-1-k}$$

5.3.5 PREDICCIÓN EN MODELOS DE REGRESIÓN MULTIPLE

Los intervalos de confianza al $100(1-\alpha)\%$ para los β_i pueden calcularse como (JURAN y GRZYNA, 1997, cap. 23):

$$b_i \pm t_{\alpha/2, n-1-k} S \sqrt{c_{ii}} \quad (\text{ec. 5.27})$$

Por otra parte, el intervalo de confianza para el valor medio poblacional de "y" para determinados valores de las x_i , vendrá dado como (ROMERO y ZUNICA, 1993, cap. 11):

$$y_t^* \pm t_{\alpha/2, n-1-k} \sqrt{S^2 \bar{x}_t' (\bar{x}'\bar{x})^{-1} \bar{x}_t'} \quad (\text{ec. 5.28})$$

Donde (ROMERO y ZUNICA, 1993, cap. 11):

$$\bar{x} = \begin{Bmatrix} 1 \\ x_{1t} \\ \vdots \\ x_{kt} \end{Bmatrix} \quad \text{y} \quad S^2 = CM_{RESIDUAL} = \frac{\sum_{j=1}^N \varepsilon_j^2}{n - (k+1)}$$

5.3.6 VALIDACIÓN DEL MODELO: ANÁLISIS DE RESIDUOS

Todo el análisis inferencial expuesto se realiza bajo la hipótesis de que el modelo postulado es correcto. Resulta por tanto esencial utilizar adicionalmente la información contenida en los datos para cuestionarse la adecuación de dicho modelo (ROMERO y ZUNICA, 1993, cap. 11).

Debe recordarse que el modelo se sintetiza en la ecuación:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

Donde los residuos ε son $N(0, \sigma^2)$ e independientes.

Distintas cuestiones pueden plantearse relativas a la adecuación o no de estas hipótesis ante unos datos concretos (ROMERO y ZUNICA, 1993, cap. 11):

- ¿Es admisible el que los residuos, ε , se distribuyen normalmente?
- ¿Hay algún dato claramente anómalo?
- ¿Es admisible que la varianza de los errores no depende de los valores de las variables explicativas?
- ¿Es realmente lineal la relación entre la $E(y)$ y una x_i cuantitativa?

La herramienta más poderosa para analizar estas cuestiones es el *análisis de residuos* (ROMERO y ZUNICA, 1993, cap. 11).

El residuo estimado para cada observación e_t no es más que la diferencia entre el valor realmente observado " y_t " y el valor medio previsto a partir del modelo estimado para esos valores concretos de las variables explicativas ($b_0 + b_1x_{1t} + \dots + b_kx_{kt}$) (ROMERO y ZUNICA, 1993, cap. 11)

$$e_t = y_t - (b_0 + b_1x_{1t} + \dots + b_kx_{kt}) \quad (\text{ec. 5.29})$$

Los residuos, son en realidad las estimaciones de los valores de las perturbaciones aleatorias ε_t en cada observación.

Determinadas representaciones gráficas de los residuos son extremadamente útiles para responder a algunas cuestiones que se plantean en la fase de validación de modelos de regresión (ROMERO y ZUNICA, 1993, cap. 11):

1. Un gráfico de los residuos e_t en papel probabilístico normal permite estudiar si es admisible la hipótesis de normalidad, así como detectar posibles observaciones anómalas. Ver gráfico 5.7.
2. Un gráfico de los cuadrados de los residuos frente a las diferentes x_i puede poner de manifiesto el hecho de que algunas de las variables explicativas afectan a las varianzas de las y_t . En situaciones de este tipo, un análisis aproximado de los

efectos sobre la varianza puede llevarse a cabo realizando un ajuste de regresión múltiple de los valores de los e_t^2 frente a las x_{it} .

3. Un gráfico de los residuos e_t frente a los valores previstos para cada observación \bar{y} puede poner de manifiesto la existencia de relaciones no lineales o bien de dependencia de los residuos, como en los casos presentados en el gráfico 5.8 respectivamente. Para estudiar que variables son responsables de esta no linealidad, son especialmente útiles los gráficos del tipo "componente más residuo" en los que se representan frente a los valores de x_{it} de una variable explicativa los valores y_{it} , una vez eliminados de los mismos los efectos estimados de las restantes variables explicativas. Por ejemplo, si se trata de estudiar la naturaleza del efecto de x_1 :

$$E(Y_t) = \beta_0 + f(x_{1t}) + \beta_2 x_{2t} + \dots + \beta_I x_{It}$$

$$f(x_{1t}) = E(Y_t) - (\beta_0 + \beta_2 x_{2t} + \dots + \beta_I x_{It})$$

Que puede estimarse a partir de:

$$Y_t - (b_0 + b_2 x_{2t} + \dots + b_I x_{It}) = b_1 x_{1t} + e_t \quad (\text{ec. 5.30})$$

(componente + residuo)

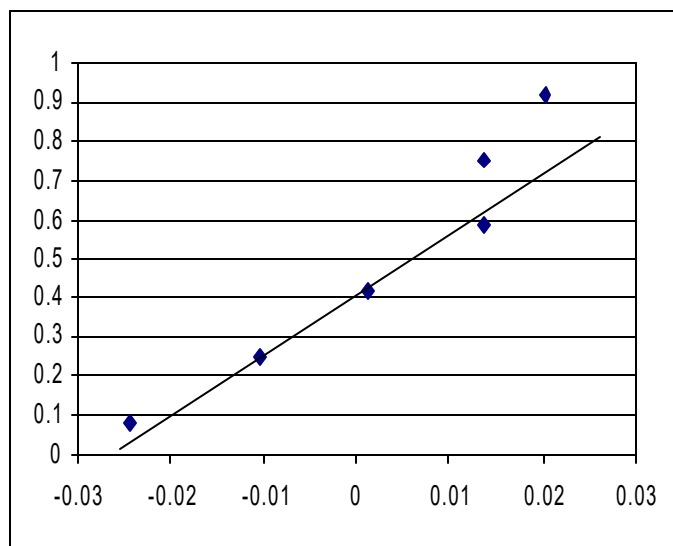


FIGURA 5.7 GRÁFICO DE NORMALIDAD DE LOS RESIDUOS

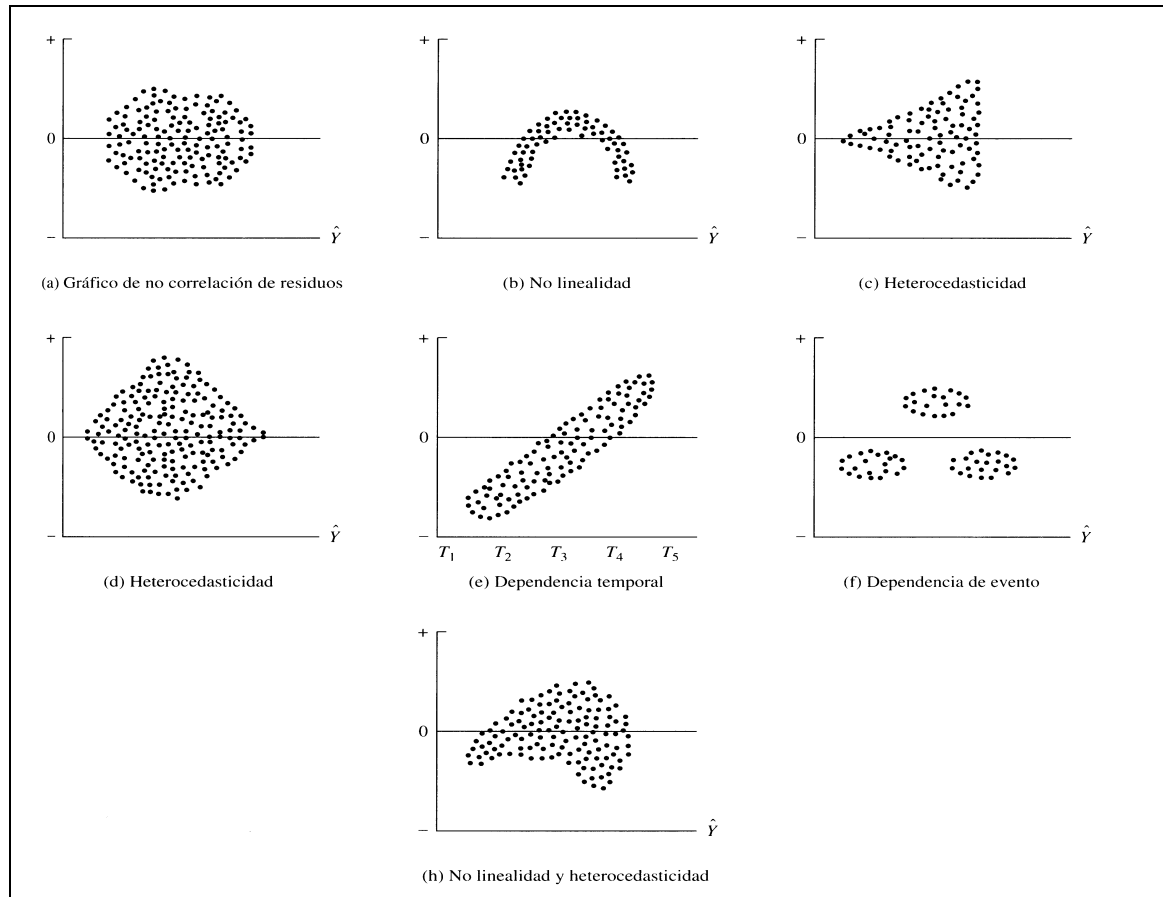


FIGURA 5.8 COMPORTAMIENTO DE LOS RESIDUOS CONTRA X_i Y VALORES PREDICHOS

5.4 REGRESION NO LINEAL

Al modelar una relación como la que se presenta en la figura 5.1, en ocasiones resulta una relación no lineal. Cuando el comportamiento es no lineal, el modelo más apropiado para un conjunto particular de datos viene definido según el modo en que se presentan los coeficientes desconocidos. Si los coeficientes se pueden sumar linealmente, después de aplicar una linealización al modelo, se tiene un modelo lineal. Por el contrario, si los coeficientes no se pueden sumar en forma lineal, el modelo se conoce como no lineal. Ejemplos de estos dos tipos de modelos son como sigue (PERRY, 1997, cap. 3):

Tipo lineal

$$\begin{aligned}
 y &= a + bx + cx^2 \\
 y &= a + bx + c \ln x \\
 y &= a + \frac{b}{x} + \frac{c}{x^2}
 \end{aligned}
 \tag{ec. 5.31}$$

Tipo no lineal

$$\begin{aligned}
 y &= a + be^{cx} \\
 y &= (a + bx_1^c)/(1 + cx_2) \\
 y &= a + \frac{b}{(c + x)}
 \end{aligned}
 \tag{ec. 5.32}$$

Para el primero se puede aplicar en forma directa el método de los mínimos cuadrados para determinar los valores óptimos de los parámetros, usando posiblemente las variables originales transformadas según lo sugiere la linealización del modelo (FOGLER, 2001, cap. 5).

Para la regresión no lineal, el principio de los mínimos cuadrados no se puede aplicar directamente, porque las ecuaciones normales resultantes contienen coeficientes desconocidos en forma no lineal y no es posible resolverlas en forma directa. Por consiguiente, es necesario utilizar un procedimiento iterativo de cálculo, en el que se buscan los valores de los parámetros que minimizan la suma de los cuadrados de las diferencias entre los valores medidos y los valores calculados para todos los puntos de datos (PERRY, 1997, cap. 3) (FOGLER, 2001, cap. 5).

No existen métodos analíticos para determinar el modelo más adecuado para un conjunto particular de datos. Sin embargo, en muchos casos se pueden seleccionar modelos posibles y razonables con base en la naturaleza de la aplicación (MILLER y MILLER, 1993, cap. 5).

Para una regresión simple y cuando no se conoce la forma funcional se pueden utilizar las curvas generales de tendencia como guía inicial de selección. Para desarrollar una relación viable y adecuada, se debe ajustar a los datos una variedad de funciones posibles. La selección más adecuada se logra cuando se dispone de observaciones duplicadas. En estos casos, cabe asegurar directamente la factibilidad de una forma particular de modelo, mediante lo que se conoce como **análisis de falta de ajuste** (PERRY, 1997, cap. 3).

En estas aplicaciones, la práctica común consiste en concentrarse en el desarrollo y cálculo del modelo, en lugar de hacerlo en los aspectos estadísticos. En general, la regresión no lineal sólo debe aplicarse a problemas en los que existe una asociación clara y bien definida entre las variables; en consecuencia, la prueba de hipótesis en la significación del ajuste es en cierto modo risible. Además, la generalización de la teoría para los intervalos asociados de confianza de los coeficientes no lineales no está bien desarrollada (PERRY, 1997, cap. 3).

5.5 EJEMPLOS DE APLICACIÓN DE ANÁLISIS DE REGRESIÓN

5.5.1 REGRESIÓN LINEAL SIMPLE

1. Los siguientes valores son las áreas relativas de los picos de los cromatogramas correspondientes a unas soluciones estándar de metilvinilcetona (MVK) (MILLER y MILLER, 1993, cap.5).

Concentración MVK, moles/ l	Área relativa del pico
0.50	3.76
1.50	9.16
2.50	15.03
3.50	20.42
4.50	25.33
5.50	31.97

- a) Evaluar el ajuste de un modelo de regresión lineal simple a estos datos.
- b) Las áreas relativas de los dos picos correspondientes a dos muestras de MVK fueron 6.3 y 27.5. Estime los límites de confianza para las concentraciones de MVK en cada solución.
- c) Suponer que los resultados del apartado b) representan una única medida así como la media de cuatro mediciones. Calcular en ambos casos las desviaciones estándar absoluta y relativa.

SOLUCIÓN

- a) El modelo a ajustar viene dado por:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

Donde los estimadores mínimo cuadráticos de β_0 y β_1 vienen dados por:

$$b_o = \bar{y} - b_1 \bar{x}$$

$$b_1 = \frac{\sum (x_m - \bar{x})(y_m - \bar{y})}{\sum (x_m - \bar{x})^2}$$

Con

$$\bar{y} = \sum y_m / n$$

$$\bar{x} = \sum x_m / n$$

Para los datos del problema se presentan los cálculos en la tabla E51.1

Tabla E51.1. datos para el problema de determinación de concentraciones de MVK

Concentración MVK, moles/ l	Área relativa del pico	$x_m - \bar{x}$	$y_m - \bar{y}$	$(x_m - \bar{x})(y_m - \bar{y})$	$(x_m - \bar{x})^2$	$(y_m - \bar{y})^2$
0.50	3.76	-2.5	-13.85	34.625	6.25	191.823
1.50	9.16	-1.5	-8.45	12.675	2.25	71.403
2.50	15.03	-0.5	-2.58	1.290	0.25	6.656
3.50	20.42	0.5	2.81	1.405	0.25	7.896
4.50	25.33	1.5	7.72	11.58	2.25	59.598
5.50	31.97	2.5	14.36	35.9	6.25	206.210
$\bar{x} = 18/6 = 3$	$\bar{y} = 17.61$			97.475	17.5	543.585

Las estimaciones de los parámetros del modelo son:

$$b_o = \bar{y} - b_1 \bar{x} = 17.61 - (5.57) * (3) = 0.9$$

$$b_1 = \frac{\sum (x_m - \bar{x})(y_m - \bar{y})}{\sum (x_m - \bar{x})^2} = \frac{97.475}{17.5} = 5.57$$

El coeficiente de determinación para el modelo viene dado por:

$$R^2 = \frac{SC(\text{regresion})}{\sum (y_m - \bar{y})^2} = \frac{b_1 \sum (x_m - \bar{x})(y_m - \bar{y})}{\sum (y_m - \bar{y})^2} = \frac{5.57 * 97.475}{543.585} = 0.9988$$

Así pues la ecuación de la línea es:

$$\text{Área del pico} = 0.9 + 5.57 \text{concentración}$$

La varianza residual, tomando como base el cálculo de residuos de la tabla E51.2 es:

$$S = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n - 2}} = \sqrt{\frac{0.64975}{6 - 2}} = 0.4030$$

TABLA E51.2 CALCULO DE RESIDUOS

Concentración MVK, moles/l (x)	Área relativa del pico (y)	$\bar{y} = 0.9 + 5.57x$	Residuos $y_i - \bar{y}$	Residuos estandarizados	(Residuos) ²
0.50	3.76	3.685	0.075	0.230362	0.005625
1.50	9.16	9.255	-0.095	-0.249992	0.009025
2.50	15.03	14.825	0.205	0.5027	0.042025
3.50	20.42	20.395	0.025	0.0554291	0.000625
4.50	25.33	25.965	-0.635	-4.80987	0.403225
5.50	31.97	31.535	0.435		0.189225
				total....	0.64975

La varianza asociada a cada coeficiente viene dada por:

$$Sb_1 = \frac{S}{\sqrt{\sum (x_m - \bar{x})^2}} = \frac{0.4030}{17.5^{0.5}} = 0.09633$$

$$Sb_0 = S \sqrt{\frac{1}{n - (\sum x_i)^2 / \sum x_i^2}} = 0.4030 * \sqrt{\frac{1}{6 - (18)^2 / (71.5)}} = 0.332555$$

La prueba de significación se hará a partir de la prueba 1, del modelo de regresión lineal simple:

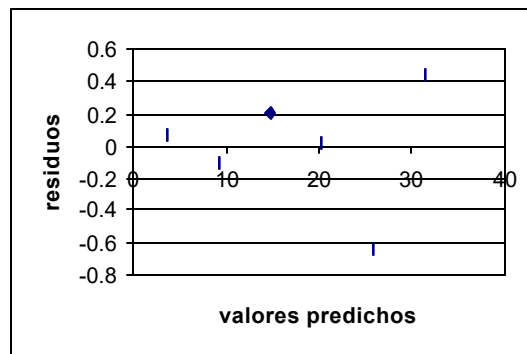
Parámetro	Valor	desv. estándar	t _o
b ₀	0.9	0.09633	9.342884
b ₁	5.57	0.33256	16.7488573

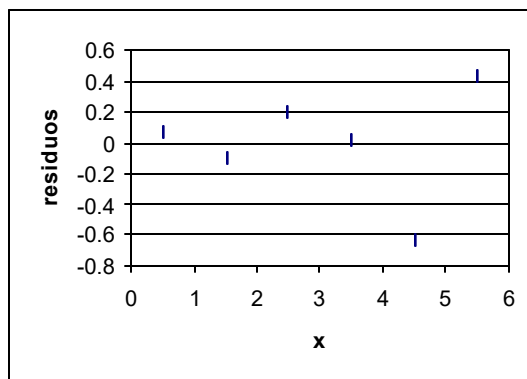
$$t_{0.025,4} = 2.776$$

De acuerdo a esta prueba, ambas estimaciones son significativamente distintas de cero a un nivel de significancia del 5%.

La validación del modelo viene descrita por el análisis de residuos. Al graficar el comportamiento de los residuos frente a los valores predichos y valores de x, grafica E51.1, se observa que esta gráfica no revela la existencia de una relación no lineal:

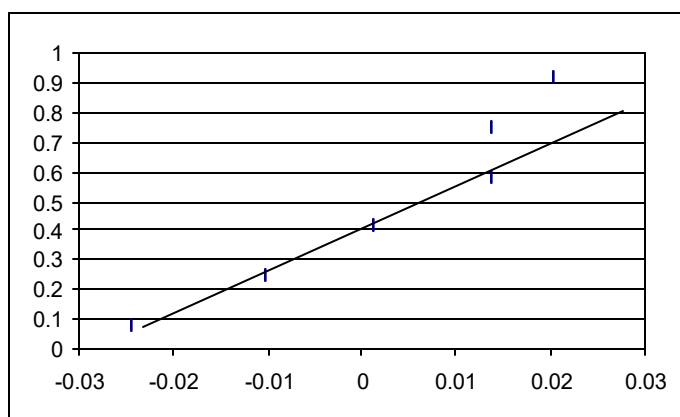
Grafico E51.1 a) residuos versus valores predichos; b) residuos contra valores de x





El gráfico de los residuos en papel probabilístico normal, grafico E51.2, permite observar que los residuos se distribuyen normalmente:

GRAFICA E51.2 GRAFICO DE PROBABILIDAD NORMAL DE LOS RESIDUOS ESTANDARIZADOS



Los intervalos de confianza para cada parametro estimado son:

$$b_1 \pm \frac{t_{\alpha/2, n-2} S}{\sqrt{\sum (x_m - \bar{x})^2}} = 5.57 \pm \frac{2.776 * 0.4030}{\sqrt{17.5}} = 5.57 \pm 0.27$$

b) Las áreas relativas de los dos picos correspondientes a dos muestras de MVK fueron 6.3 y 27.5. Estime los límites de confianza para las concentraciones de MVK en cada solución. Los límites de confianza vienen dados por:

- Para el área de 6.3, se tiene:

$$x_o = \frac{y_o - 0.9}{5.57} = \frac{6.3 - 0.9}{5.57} = 0.9695 \text{ moles / l}$$

- Para el área de 27.5, se tiene:

$$x_o = \frac{y_o - 0.9}{5.57} = \frac{27.5 - 0.9}{5.57} = 4.776 \text{ moles/l}$$

c) Suponer que los resultados para el pico de área 6.3 del apartado b) representan la media de cuatro mediciones. Calcular las desviaciones estándar absoluta y relativa:

- Para 0.9695 moles/l

Desviación absoluta:

$$s_{x_o} = \frac{S}{b_1} \sqrt{\frac{1}{k} + \frac{1}{n} + \frac{(\bar{y}_c - \bar{y})^2}{b_1^2 \sum (x_i - \bar{x})^2}} = \frac{0.4030}{5.57} \sqrt{\frac{1}{4} + \frac{1}{6} + \frac{(6.3 - 17.61)^2}{5.57^2 * (17.5)}} = 0.0584$$

Desviación relativa:

$$\frac{s_{x_o}}{y_o} = \frac{0.0584}{6.3} = 0.00928$$

5.5.2 REGRESION LINEAL MULTIPLE

2. El rendimiento de una reacción química depende de la temperatura de operación y de la concentración inicial del reactivo. Efectué un análisis de regresión a los siguientes datos (MONTGOMERY, 1991, cap. 15).

Rendimiento	concentración	temperatura
81	1	150
89	1	180
83	2	150
91	2	180
79	1	150
87	1	180
84	2	150
90	2	180

SOLUCION

Se ajustara el modelo lineal: $y = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ a los datos, el modelo puede expresarse en términos matriciales mediante:

$$\begin{bmatrix} 81 \\ 89 \\ 83 \\ 91 \\ 79 \\ 87 \\ 84 \\ 90 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 150 \\ 1 & 1 & 180 \\ 1 & 2 & 150 \\ 1 & 2 & 180 \\ 1 & 1 & 150 \\ 1 & 1 & 180 \\ 1 & 2 & 150 \\ 1 & 2 & 180 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \end{bmatrix}$$

$\bar{y} = 85.5$

La estimación de los coeficientes viene dada por

$$\bar{b} = (\bar{x}'\bar{x})^{-1} \bar{x}'\bar{y} = \begin{bmatrix} 39.75 \\ 3.0 \\ 0.25 \end{bmatrix}$$

La varianza residual es entonces:

$$S^2 = \frac{\sum_1^n \epsilon_j^2}{8 - (2 + 1)} = \frac{\sum_1^n (y_j - (b_0 + b_1 x_{1j} + \dots + b_l x_{lj}))^2}{5} = \frac{5.5}{5} = 1.04881$$

Los residuos vienen dados en la tabla E52.1:

TABLA E52.1

y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(y_i - \bar{y})^2$
81	0.75	0.5625	20.25
89	1.25	1.5625	12.25
83	-0.25	0.0625	6.25
91	0.25	0.0625	30.25
79	-1.25	1.5625	42.25
87	-0.75	0.5625	2.25
84	0.75	0.5625	2.25
90	-0.75	0.5625	20.25
Total		5.5	136

El coeficiente de determinación dado a partir de:

$$R^2 = \frac{SC_{EXPLICADA}}{SC_{TOTAL}} = \frac{\sum_{j=1}^{j=n} (y_j - \bar{y})^2 - \sum_{j=1}^n \epsilon_j^2}{\sum_{j=1}^{j=n} (y_j - \bar{y})^2} = \frac{136 - 5.5}{136} = 0.959559$$

Las varianzas asociadas a cada uno de los estimadores de los parámetros que son dadas a partir de $S_{b_i} = S\sqrt{c_{ii}}$ donde c_{ii} es el elemento i, i de la diagonal principal de $(\bar{x}'\bar{x})^{-1}$, se presentan en la tabla E52.2:

TABLA E52.2

Parámetro	estimación	S_{b_i}
β_0	39.75	4.24411
β_1	3.0	0.74162
β_2	0.25	0.02472

La significación global del ajuste se presenta en la tabla E52.3:

TABLA E52.3

fuerce	SC	g.l.	CM	F_0
explicada	130.5	2	65.25	59.32
residuos	5.5	5	1.1	

$$F_{\alpha, I, (n-1-I)} = F_{0.05, 2, 5} = 5.79$$

Al comparar F_0 con el $F_{0.05, 2, 5}$ puede concluirse que el modelo es significativo y que al menos un β_i es distinto de cero.

La significancia del efecto de cada X_i se probara a partir de la prueba t , basada en una prueba "t", dicho análisis se presenta a continuación:

Parámetro	estimación	S_{b_i}	t_0
β_0	39.75	4.24411	9.366
β_1	3.0	0.74162	4.045
β_2	0.25	0.02472	10.11

$$t_{\alpha/2, n-1-I} = t_{0.025, 5} = 2.571$$

Al comparar el t_0 asociado a cada β_i con la $t_{0.025, 5}$ puede observarse que los efectos tanto de la temperatura como de la concentración son significativos a un nivel de confianza del 95%.

El modelo ajustado es por tanto:

$$\bar{y} = 39.75 + 3.0 * \text{concentracion} + 0.25 * \text{temperatura}$$

Los intervalos de confianza al 95% para cada coeficiente del modelo de regresión son:

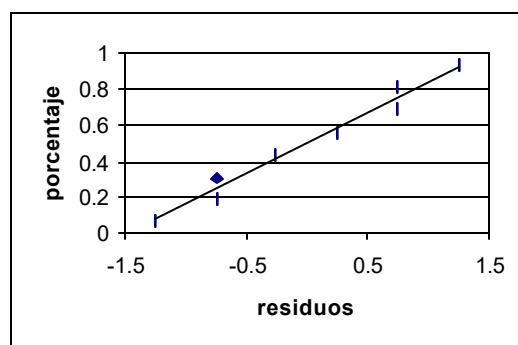
Parámetro	estimación b_i	S_{b_i}	limite inferior	limite superior
β_0	39.75	4.24411	28.8401	50.6599
β_1	3.0	0.74162	1.0936	4.9064
β_2	0.25	0.02472	0.18645	0.31355

El intervalo de confianza para los b_i se calculó como: $b_i \pm t_{0.025, 5} S_{b_i}$

La validación del modelo se hará en base al análisis de los residuos, a través de los siguientes gráficos: a) gráfico de probabilidad normal de los residuos (figura E52.1); b) gráfico de los residuos frente a los valores predichos (figura E52.2 a); c) gráficos de los residuos frente a cada variable (figura E52.2 b):

FIGURA E52.1 GRAFICO DE PROBABILIDAD NORMAL DE LOS RESIDUOS

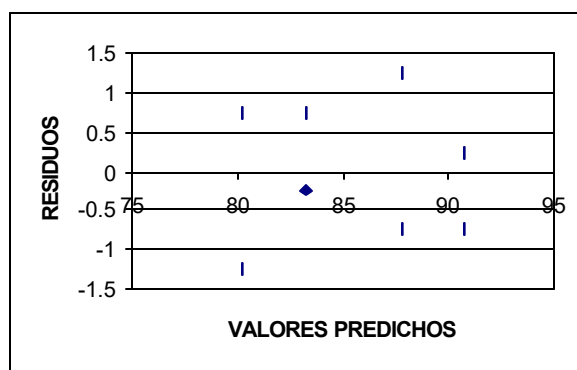
orden	residuo	fracción
1	-1.25	0.0625
2	-0.75	0.1875
3	-0.75	0.3125
4	-0.25	0.4375
5	0.25	0.5625
6	0.75	0.6875
7	0.75	0.8125
8	1.25	0.9375



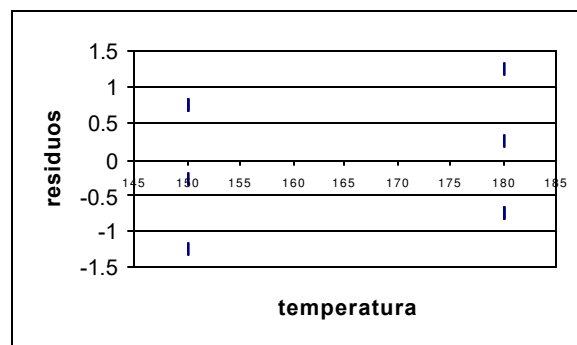
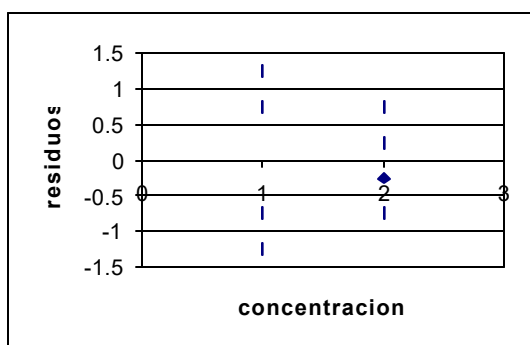
- Según el gráfico anterior puede observarse que la distribución de los residuos es normal.

FIGURA E52.2 GRAFICO DE RESIDUOS VERSUS VALORES PREDICHOS Y VALORES DE X_i

valores predichos \hat{y}	residuos
80.25	0.75
87.75	1.25
83.25	-0.25
90.75	0.25
80.25	-1.25
87.75	-0.75
83.25	0.75
90.75	-0.75



(a)



(b)

- En la grafica de residuos contra valores predichos parece observarse un comportamiento curvilíneo, sin embargo, por venir este ajuste de un diseño 2^2 con 2 replicas no se puede tratar de ajustar un modelo cuadrático, ya que para esto se necesita un diseño con al menos tres niveles para cada factor. Un análisis de los gráficos de los residuos contra las variables concentración y temperatura permite concluir que en cuanto al factor concentración se presenta un efecto muy importante sobre la variabilidad del rendimiento, ya que una mayor concentración reduce la variabilidad en cuanto al rendimiento de la reacción química.

5.5.3 REGRESION NO LINEAL

3. La hidrogenación (H) de etileno para formar etano (EA), $H_2 + C_2H_4 \rightarrow C_2H_6$, se efectúa sobre un catalizador de cobalto y molibdeno [*Collect. Czech. Chem. Común.*, 51,2760 (1988)]. Realice un análisis de mínimos cuadrados no lineal con los datos que se dan en la tabla a continuación:

Número de prueba	Veloc. de reacción (mol/ kg cat. s)	P_E (atm)	P_{EA} (atm)	P_H (atm)
1	1.04	1	1	1
2	3.13	1	1	3
3	5.21	1	1	5
4	3.82	3	1	3
5	4.19	5	1	3
6	2.391	0.5	1	3
7	3.867	0.5	0.5	5
8	2.199	0.5	3	3
9	0.75	0.5	5	1

Determine cuál de las siguientes leyes de velocidad es la que mejor describe los datos:

$$(a) -r_A = \frac{kP_E P_H}{1 + k_E P_E}; \quad (b) -r_A = kP_E^a P_H^b$$

SOLUCION

$$(a) \text{ MODELO: } -r_A = \frac{kP_E P_H}{1 + k_E P_E};$$

El modelo es un modelo no lineal, sin embargo, puede ser linealizable al trabajar con el inverso de $-r_A$, de la siguiente forma:

$$\frac{1}{-r_A} = \frac{1}{kP_E P_H} + \frac{k_E P_E}{kP_E P_H} = \frac{1}{k} \left(\frac{1}{P_E P_H} \right) + \frac{k_E}{k} \left(\frac{1}{P_H} \right)$$

Los parámetros del modelo son k y k_E . Para el análisis de mínimos cuadrados se podrá trabajar con una regresión lineal múltiple sobre las variables transformadas, en la tabla E53.1 se presenta esta transformación de variables:

TABLA E53.1

$1/-r_A$	$1/(P_E^* P_H)$	$1/P_H$
0.96153846	1	1
0.31948882	0.33333333	0.33333333
0.19193858	0.2	0.2
0.2617801	0.11111111	0.33333333
0.23866348	0.06666667	0.33333333
0.41823505	0.66666667	0.33333333
0.2585984	0.4	0.2
0.45475216	0.66666667	0.33333333
1.33333333	2	1

El análisis del modelo de regresión lineal múltiple puede realizarse tal como se presenta en el ejercicio 2, para las variables transformadas:

$$\begin{bmatrix} 0.9615 \\ 0.3195 \\ 0.1919 \\ 0.2618 \\ 0.2387 \\ 0.4182 \\ 0.2586 \\ 0.4548 \\ 1.3333 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0.3333 & 0.3333 \\ 0 & 0.2 & 0.2 \\ 0 & 0.1111 & 0.3333 \\ 0 & 0.0667 & 0.3333 \\ 0 & 0.6667 & 0.3333 \\ 0 & 0.4 & 0.2 \\ 0 & 0.6667 & 0.3333 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1/k \\ k_E/k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_{98} \end{bmatrix}$$

La estimación de los coeficientes viene dada por

$$\bar{b} = (\bar{x}'\bar{x})^{-1} \bar{x}'\bar{y} = \begin{bmatrix} 0.3507 \\ 0.6221 \end{bmatrix}$$

El coeficiente de determinación es dado a partir de:

$$R^2 = \frac{SC_{EXPLICADA}}{SC_{TOTAL}} = \frac{\sum_{j=1}^{j=n} (y_j - \bar{y}_j)^2 - \sum_{j=1}^n \epsilon_j^2}{\sum_{j=1}^{j=n} (y_j - \bar{y}_j)^2} = 0.99894$$

Las varianzas asociadas a cada uno de los estimadores de los parámetros que son dadas a partir de $S_{bi} = S\sqrt{c_{ii}}$ donde c_{ii} es el elemento i, i de la diagonal principal de $(\bar{x}'\bar{x})^{-1}$, se presentan en la tabla E53.2:

TABLA E53.2

Parámetro	estimación	S_{bi}
$1/k$	0.35069	0.01419
k_E/k	0.62207	0.02175

La significación global del ajuste se presenta en la E53.3:

TABLA E53.3

fuelle	SC	g.l.	MC	F_o
explicada	1.2253	2	0.61264	3293.38
residuos	0.00130	7	0.0001857	

$$F_{\alpha, I, (n-1-I)} = F_{0.05, 2, 7} = 4.74$$

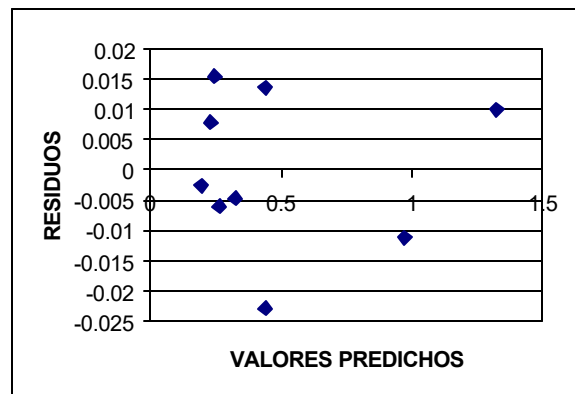
Al comparar F_o con el $F_{0.05, 2, 7}$ puede concluirse que el modelo es significativo, siendo la suma de cuadrados mínima de 0.0013.

En la tabla E53.4 se presentan los residuos calculados y en el grafico E53.1 se presenta el comportamiento de los residuos contra los valores predichos por el modelo.

TABLA E53.4

Observación	Pronóstico para Y	Residuos
0.961546	0.972768	-0.011221
0.319489	0.324260	-0.004767
0.191938	0.194554	-0.002615
0.261780	0.246324	0.015456
0.238663	0.230738	0.007925
0.418235	0.441153	-0.022918
0.258510	0.264692	-0.006094
0.454752	0.441153	0.013599
1.333333	1.323460	0.009873

GRAFICO E53.1



Los valores de k y k_E para el modelo son por tanto:

$$k = 1/0.35069 = 2.852$$

$$k_E = 0.6221k = (0.6221) * (2.852) = 1.774$$

Y por tanto, el modelo queda escrito como: $-r_A = \frac{2.852P_E P_H}{1+1.774P_E}$

Este ejercicio también puede desarrollarse usando la función *regresión* de la opción *análisis de datos* del menú *Herramientas* en la hoja de cálculo de Excel. En el archivo [regresión no lineal m1.xls](#), se presentan las salidas para dicha aplicación:



regresión no lineal
m1.xls

(b) MODELO: $-r_A = kP_E^a P_H^b$

Al igual que en el caso anterior el modelo no es un modelo lineal, sin embargo, es linealizable al aplicar una transformación logarítmica, como se presenta a continuación:

$$-r_A = kP_E^a P_H^b \rightarrow \ln(-r_A) = \ln k + a \ln P_E + b \ln P_H$$

Los parámetros del modelo en este caso son $\ln k$, a y b . Para el análisis de mínimos cuadrados se podrá trabajar con una regresión lineal múltiple sobre las variables transformadas. En la tabla E53.5 se presenta esta transformación de variables:

TABLA E53.5

$\ln(-r_A)$	$\ln(P_E)$	$\ln(P_H)$
0.0392	0.0000	0.0000
1.1410	0.0000	1.0986
1.6506	0.0000	1.6094
1.3403	1.0986	1.0986
1.4327	1.6094	1.0986
0.8717	-0.6931	1.0986
1.3525	-0.6931	1.6094
0.7880	-0.6931	1.0986
-0.2877	-0.6931	0.0000

El análisis del modelo de regresión lineal múltiple puede realizarse tal como se presenta en el ejercicio 2, para las variables transformadas:

$$\begin{bmatrix} 0.0392 \\ 1.1410 \\ 1.6506 \\ 1.3403 \\ 1.4327 \\ 0.8717 \\ 1.3525 \\ 0.7880 \\ -0.2877 \end{bmatrix} = \begin{bmatrix} 1 & 0.0000 & 0.0000 \\ 1 & 0.0000 & 1.0986 \\ 1 & 0.0000 & 1.6094 \\ 1 & 1.0986 & 1.0986 \\ 1 & 1.6094 & 1.0986 \\ 1 & -0.6931 & 1.0986 \\ 1 & -0.6931 & 1.6094 \\ 1 & -0.6931 & 1.0986 \\ 1 & -0.6931 & 0.0000 \end{bmatrix} \begin{bmatrix} \ln k \\ a \\ b \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_{98} \end{bmatrix}$$

La estimación de los coeficientes viene dada por

$$\bar{b} = (\bar{x}'\bar{x})^{-1}\bar{x}'\bar{y} = \begin{bmatrix} -0.10934 \\ 0.26719 \\ 1.05384 \end{bmatrix}$$

El coeficiente de determinación es dado a partir de:

$$R^2 = \frac{SC_{EXPLICADA}}{SC_{TOTAL}} = \frac{\sum_{j=1}^{j=n} (y_j - \bar{y}_j)^2 - \sum_{j=1}^n \varepsilon_j^2}{\sum_{j=1}^{j=n} (y_j - \bar{y}_j)^2} = 0.9898$$

Las varianzas asociadas a cada uno de los estimadores de los parámetros que son dadas a partir de $S_{bi} = S\sqrt{c_{ii}}$ donde c_{ii} es el elemento i , i de la diagonal principal de $(\bar{x}'\bar{x})^{-1}$, se presentan en la tabla E53.6:

TABLA E53.6

Parámetro	estimación	S_{bi}
$\ln k$	-0.042239026	0.061462455
a	0.271021234	0.028559749
b	1.001607145	0.051953273

La significación global del ajuste se presenta en la E53.7:

TABLA E53.7

fuerce	SC	g.l.	MC	F_0
explicada	3.427844373	2	1.713922186	293.540846
residuos	0.035032716	6	0.005838786	

$$F_{\alpha, I, (n-1-I)} = F_{0.05, 2, 6} = 5.14$$

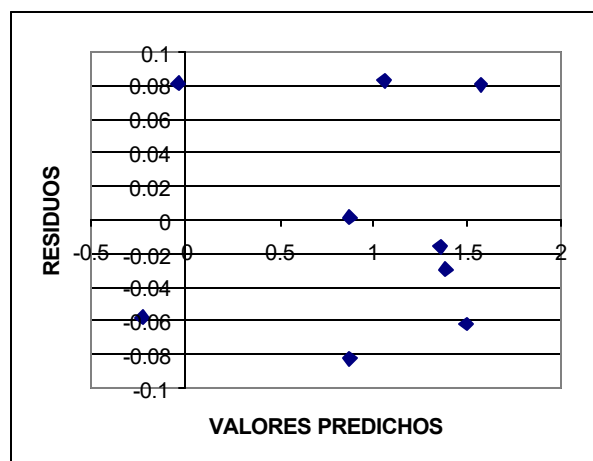
Al comparar F_0 con el $F_{0.05, 2, 6}$ puede concluirse que el modelo es significativo, siendo la suma de cuadrados mínima de 0.03503.

En la tabla E53.8 se presentan los residuos calculados y en el grafico E53.2 se presenta el comportamiento de los residuos contra los valores predichos por el modelo.

TABLA E53.8

Observación	Pronóstico para Y	Residuos
0.0392	-0.042239	0.0814597
1.1410	1.0581389	0.0828941
1.6506	1.5697855	0.0807944
1.3403	1.3558862	-0.0156357
1.4327	1.4943307	-0.0616300
0.8717	0.8702813	0.0014304
1.3525	1.3819279	-0.0294489
0.7880	0.8702813	-0.0822786
-0.2877	-0.230097	-0.0575854

GRAFICO E53.2



Los valores de k , a y b para el modelo son por tanto:

$$k = e(-0.04224) = 0.9586$$

$$a = 0.271$$

$$b = 1.001$$

Y el modelo queda escrito como: $-r_A = 0.9586P_E^{0.27}P_H$

Este ejercicio también puede desarrollarse usando la función *regresión* de la opción *análisis de datos* del menú *Herramientas* en la hoja de cálculo de Excel. En el archivo [regresión no lineal m2.xls](#), se presentan las salidas para dicha aplicación:



Regresión no lineal
m2.xls

En función de la comparación del R^2 y de la MCE (ver tabla E53.7) para ambos modelos puede concluirse que el modelo que mejor representa la ley de velocidad de los datos es el modelo del literal a).

TABLA E53.7

Modelo	R^2	MCE
$-r_A = \frac{2.852P_E P_H}{1 + 1.774P_E}$	0.9989	0.0001857
$-r_A = 0.9586P_E^{0.27}P_H$	0.9898	0.005838

5.5.4 OTROS TEMAS DE APLICACIÓN DE LA REGRESIÓN LINEAL AL ANÁLISIS QUÍMICO

4. Una determinación con un electrodo selectivo de iones (ESI) de sulfato procedente de sulfato reducido por bacterias, se comparó con una determinación gravimétrica. Los resultados obtenidos se expresaron en miligramos de sulfuro, tal como se presenta en la tabla a continuación:

MUESTRA	Método ESI	Gravimetría
1	108	105
2	12	16
3	152	113
4	3	0
5	106	108
6	11	11
7	128	141
8	12	11
9	160	182
10	128	118

Comente sobre la conveniencia del método ESI para esta determinación de sulfuro. (Al-Hitti, I.K., Moody, G.J. y Thomas, J.D.R., *Analyst*, 1983, 108, 43).

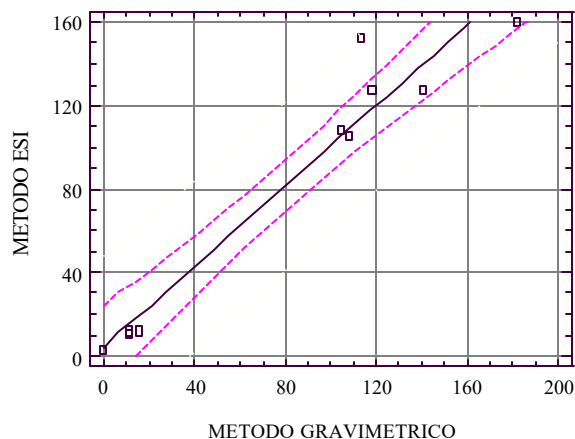
SOLUCION

El objetivo del problema es comparar el "nuevo" método ESI con el método gravimétrico para la determinación de sulfatos. Con los datos proporcionados la comparación de los métodos puede hacerse a través del uso de rectas de regresión. En la práctica de este método, el analista con frecuencia desea realizar una prueba para contrastar una ordenada en el origen significativamente diferente de cero, y una pendiente significativamente diferente de 1. Tales pruebas se realizan determinando los intervalos de confianza para β_0 y β_1 , generalmente al 95% de significancia.

Los resultados de los análisis con los dos métodos se representaran en una línea de regresión, con los datos obtenidos por el método gravimétrico en el eje x y los que resultaron del ESI en el eje y, ver gráfico E54.1.

GRAFICO E54.1

COMPARACION DE METODOS



Usando los métodos de las secciones precedentes las estimaciones de los parámetros β_0 y β_1 , y de R son:

$$b_0 = 4.4337; b_1 = 0.929; R = 0.9697; S = 16.7264; S_{b_1} = 0.0857; S_{b_0} = 8.649$$

Los intervalos de confianza al 95%, para los parámetros son los siguientes ($b_i \pm t_{0.025,9} S_{b_i}$):

$$b_0 = 4.4337 \pm 19.56 \quad \text{y} \quad b_1 = 0.929 \pm 0.1938$$

De estos resultados se desprende que la pendiente y la ordenada en el origen calculadas no difieren significativamente de los valores "ideales" de 1 y 0 respectivamente, y por lo tanto no existen diferencias sistemáticas entre la serie de resultados. Sin embargo, el coeficiente de determinación es inferior al 99%, lo que indica que alguno de los métodos es menos preciso.

5. El contenido de oro de una muestra de mar concentrada se determinó por espectrometría de absorción atómica mediante el método de las adiciones estándar. Los resultados obtenidos fueron los siguientes (MILLER y MILLER, 1993, cap.5):

Oro añadido ng/ ml de muestra concentrada	absorbancia
0	0.257
10	0.314
20	0.364
30	0.413
40	0.468
50	0.528
60	0.574
70	0.635

Estime la concentración de oro en la muestra de agua de mar, y determine los límites de confianza para esta estimación.

SOLUCION

Basándose en las ecuaciones ya descritas en el tema 5.1, y usando el modo de regresión lineal de la calculadora HP-49G, la estimación de los coeficientes de regresión es:

$$b_0 = 0.25692 ; \quad b_1 = 0.00534881; \quad R^2 = 0.99966$$

La razón entre b_0 y b_1 proporciona la concentración de oro en la muestra problema de 48.0331 ng/ml. Los límites de confianza para este resultado se pueden determinar con la ayuda de la ecuación 5.15:

$$s_E = \frac{S}{b_1} \left[\frac{1}{n} + \frac{\bar{y}^2}{b_1^2 \sum (x_i - \bar{x})^2} \right]^{1/2}$$

Aquí $S = 0.0036939$, $\bar{y} = 0.444$, y $\sum (x_i - \bar{x})^2 = 4200$; $S_x = 1.5977$ y los límites de confianza son $48.0331 \pm 2.447 \times 1.5977$, es decir 48.0331 ± 3.9096 ng/ml.

6. La absorbancia de una serie de soluciones de cadmio fue determinada 3 veces, los resultados se exponen a continuación:

Concentración mg/l	1	2	3
0.1	0.028	0.029	0.029
0.3	0.084	0.083	0.081
0.5	0.135	0.131	0.133
0.7	0.180	0.181	0.183
0.9	0.215	0.230	0.216

Determine: (a) La recta de regresión no ponderada; (b) La recta de regresión ponderada; (c) Utilizando ambas rectas calcule los límites de confianza para la concentración de una disolución de cadmio que fue medida dos veces con una absorbancia media de 0.191 (con $s_i = 0.0021$).

SOLUCION

(a) Aplicando las ecuaciones 5.4 a la 5.12, con la concentración representada en el eje x y la absorbancia en el eje y. La estimación de los coeficientes de la recta no ponderada es la siguiente:

$$b_0 = 0.0087 \quad y \quad b_1 = 0.2410 \text{ (de ec. 5.4); } \quad R^2 = 0.9944 \text{ (de ec. 5.7);}$$

$$S = 0.005486 \text{ (de ec. 5.8)}$$

La desviación estándar asociada a cada coeficiente y el calculo del intervalo de confianza (ecs. 5.11 y 5.12) viene dado por:

Parámetro	estimación	S_{b_i}	t_o	$b_i \pm tS_{b_i}$
β_o	0.0087	0.002877	3.0243	0.0087 ± 0.0062
β_1	0.2410	0.005008	48.126	0.2410 ± 0.0108

$$t_{\alpha/2, 13} = 2.160$$

La estimación de la concentración de la solución de cadmio cuando la absorbancia es de 0.191, viene dada a partir de la ecuación 5.14, y es:

$$x_o = \frac{y - b_o}{b_1} = \frac{0.191 - 0.0087}{0.241} = 0.756$$

La incertidumbre asociada al cálculo de la concentración viene dada por:

$$u(c) = \frac{S}{b_1} \sqrt{\frac{1}{k} + \frac{1}{n} + \frac{(x_o - \bar{x})^2}{S_{xx}}}$$

$$\text{con } S_{xx} = \sum_{j=1}^n (x_j - \bar{x})^2 = 1.2, \quad k = 2, \text{ y } t_{\alpha/2, 13} = 2.160$$

$$u(c) = \frac{0.005486}{0.241} \sqrt{\frac{1}{2} + \frac{1}{15} + \frac{(0.756 - 0.5)^2}{1.2}} = 0.018 \text{ mg/l}$$

(b) La recta ponderada viene dada a partir de las ecuaciones:

En donde:

$$w_i = \frac{s_i^{-2}}{\left(\sum s_i^{-2} / n \right)}$$

$$b_1 = \frac{\sum_i w_i x_i y_i - n \bar{x}_w \bar{y}_w}{\sum_i w_i x_i^2 - n \bar{x}_w^2} \quad (\text{ec. 5.17})$$

$$\bar{x}_w = \frac{\sum_i w_i x_i}{n}$$

$$\bar{y}_w = \frac{\sum_i w_i y_i}{n}$$

Las sumatorias y cálculos que involucran las ecuaciones se presentan a continuación:

Concentración mg/l	Absorbancia promedio	desviación (S _i)	1/S _i ²	w _i	w _i x _i	w _i y _i	w _i x _i y _i	w _i x _i ²
0.1	0.028	0.00058	2972652	3.632956	0.363296	0.101723	0.010172	0.03633
0.3	0.084	0.00153	427186.1	0.522075	0.156623	0.043854	0.013156	0.046987
0.5	0.135	0.002	250000	0.305532	0.152766	0.041247	0.020623	0.076383
0.7	0.180	0.00153	427186.1	0.522075	0.365453	0.093974	0.065782	0.255817
0.9	0.215	0.00839	14206.14	0.017362	0.015626	0.003733	0.003359	0.014063
Sumas			4091230	5.0	1.053762	0.28453	0.113093	0.429579

De estos números queda claro que $\bar{y}_w = 0.28453/5 = 0.05691$ y $\bar{x}_w = 1.05376/5 = 0.2108$, Mediante la ecuación se calcula b_1 , como:

$$b_1 = \frac{\sum_i w_i x_i y_i - n \bar{x}_w \bar{y}_w}{\sum_i w_i x_i^2 - n \bar{x}_w^2} = \frac{0.113093 - (5 * 0.05691 * 0.2108)}{0.429579 - (5 * 0.2108^2)} = 0.2561$$

De forma que b_0 esta dado por $\bar{y}_w - b_1 \bar{x}_w = 0.05691 - (0.2561 * 0.2108) = 0.00292$. Estos valores de b_0 y b_1 , se usan para estimar que la concentración de cadmio de la solución para la absorbancia de 0.191 es de:

$$x_o = \frac{y - b_o}{b_1} = \frac{0.191 - 0.00292}{0.256} = 0.7348 \text{ mg/l}$$

Con una incertidumbre en la estimación calculada a partir de la ecuación 5.16:

$$s_{xy,w} = \left[\frac{\sum (w_i y_i^2 - n \bar{y}_w^2) - b_1^2 (\sum w_i x_i^2 - n \bar{x}_w^2)}{n - 2} \right]^{1/2}$$

$$s_{x_o,w} = \frac{s_{xy,w}}{b_1} \left[\frac{1}{w_o} + \frac{1}{n} + \frac{(y_o - \bar{y}_w)^2}{b_1^2 (\sum w_i x_i^2 - n \bar{x}_w^2)} \right]^{1/2}$$

cálculos adicionales:

w_iy_i²
0.002848
0.003684
0.005568
0.016915
0.000803
Σ=0.029818

De esta forma: $S_{xy,w} = 2.7386 \times 10^{-3}$, $w_0 = (1/0.0021^2)/818246 = 0.2771$

$$s_{x_0,w} = \frac{2.7386 \times 10^{-3}}{0.2561} \left[\frac{1}{0.2771} + \frac{1}{5} + \frac{(0.191 - 0.05691)^2}{0.013602} \right]^{1/2} = 0.02422$$

El intervalo de confianza para la estimación es: $0.7348 \pm 3.18 \cdot 0.02422 = 0.7348 \pm 0.0770$ mg/l.

7. Estime el límite de detección para el análisis de Metilvinilcetona del ejemplo 1.

SOLUCION

Una definición que se usa comúnmente en la bibliografía de química analítica es que el límite de detección es la *cantidad de concentración de analito que proporciona una señal igual a la del blanco, y_B , más tres veces la desviación estándar del blanco, s_B* , lo anterior se explica con la ecuación $y - y_B = 3S_B$.

Para el ejemplo, con los valores $y_B (=b_0)$ y $S_B (=S)$ calculados previamente en el ejemplo 1, el valor de y en el límite de detección es $y = y_B + 3S_B = 0.9 + 3 \cdot 0.4030$, es decir, 2.109. El uso de la ecuación de regresión conduce a un límite de detección de:

$$x_0 = \frac{y - b_0}{b_1} = \frac{2.109 - 0.9}{0.57} = 2.121 \text{ moles/l}$$

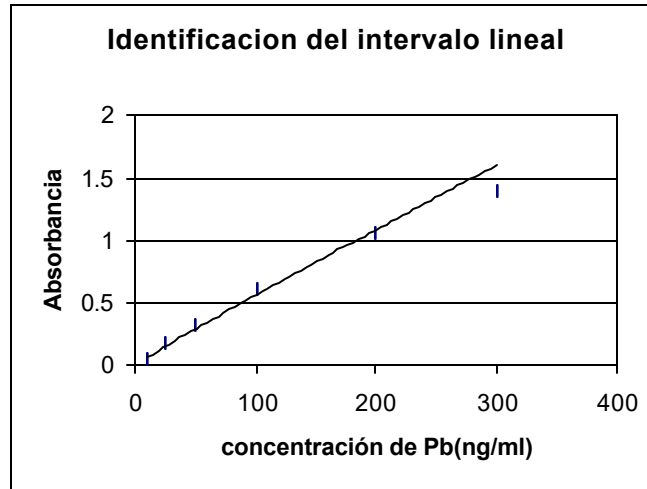
Es importante evitar la confusión entre el límite de detección de una técnica con su sensibilidad. La sensibilidad de una técnica se define correctamente como la pendiente de una línea de calibración, es decir la variación en la señal por cada unidad de aumento de la concentración del analito, y siempre que la representación sea lineal, puede ser medida en cualquier punto de ella. Por el contrario, el límite de detección de un método se calcula con ayuda de la zona de representación cercana al origen, y usando tanto la pendiente como la ordenada en el origen.

8. En la determinación de plomo en solución acuosa por espectrofotometría por absorción atómica con atomización de cámara de grafito, se obtuvieron los siguientes resultados:

Conc. de Pb ng/ml	Absorbancia
10	0.05
25	0.17
50	0.32
100	0.60
200	1.07
300	1.40

Examine el intervalo de calibración lineal de este experimento. (Basado en Giri, S.K., Shields, C.K., Littlejohn, D. y Ottaway, J.M., *Analyst*, 1983, 108,244).
SOLUCION

El punto de partida para examinar el intervalo de calibración lineal es elaborar una grafica en la que la concentración se coloque en el eje de las absisas y la absorbancia en el eje de las ordenadas, tal como se muestra a continuación:



Como puede observarse existe una desviación del comportamiento lineal del último punto, sin embargo, es conveniente, aplicar cálculos de regresión lineal (no ponderada) a todos los datos; de la aplicación de los métodos del tema 5.1, y valiéndose del modo regresión de la calculadora HP-49G, se tiene:

$$R = 0.99355; \quad b_0 = 0.07033; \quad b_1 = 0.00465$$

En este caso los residuos de la absorbancia, y , son:

y	y^2	residuos	residuos ²
0.05	0.11687	-0.06687	0.00447
0.17	0.185583	-0.015583	0.000243
0.32	0.303034	0.016966	0.000288
0.60	0.535735	0.064265	0.00413
1.07	1.00114	0.06886	0.00474
1.40	1.46654	-0.06654	0.00443
Suma de cuadrados		0.0191	0.018301

La tendencia de los valores de los residuos (que no debe esperarse ninguna, si no una distribución aleatoria de los residuos, si el modelo es completamente lineal) sugiere que el último valor de la tabla está probablemente fuera del intervalo lineal. Por otra parte, además, el coeficiente de correlación lineal con un valor de 0.99355, aunque parece marcar una buena correlación lineal entre las variables, es bajo para el caso de

la aplicación de una curva de calibración lineal. En la situación donde una gráfica de calibración sea lineal sobre parte de su intervalo, y curva en otra, es de gran importancia para el químico analítico establecer el intervalo sobre el cual se supone linealidad.

Si se elimina el último punto, los parámetros de la recta de calibración son:

$$R = 0.9972; \quad b_0 = 0.0352; \quad b_1 = 0.00528$$

Como puede observarse el valor de R es mayor que en el primer cálculo. Los residuos de los primeros seis puntos de esta segunda ecuación de regresión son:

y	\bar{y}	residuos	residuos ²
0.05	0.0880195	-0.0380195	0.0014455
0.17	0.167269	0.002731	0.000007
0.32	0.29351	0.02649	0.0007
0.60	0.563516	0.036484	0.001331
1.07	1.09184	-0.02184	0.00048
Suma de cuadrados			0.0039635

El uso de la segunda recta de regresión muestra que la absorbancia esperada para un patrón de 300 ng/ml es de 1.62, es decir el residuo es de -0.22, dado el valor de este último residuo puede decirse que todavía falta ajuste del intervalo lineal. Si se usan los primeros cuatro puntos para el cálculo de una nueva línea recta, se obtienen los siguientes resultados:

$$R = 0.998; \quad b_0 = 0.00738; \quad b_1 = 0.006$$

Como puede observarse hay una mejora en el coeficiente de correlación, los valores de los residuos son ahora:

y	\bar{y}	residuos	residuos ²
0.05	0.0674	-0.0174	0.0003
0.17	0.1574	0.0126	0.0002
0.32	0.3075	0.0125	0.0002
0.60	0.6076	-0.0076	0.0001
Suma de cuadrados			0.0008

Como puede observarse el residuo del punto correspondiente a 200 ng/ml es de -0.14. Puede concluirse, por tanto, que el quinto punto puede excluirse también del intervalo lineal del método. Pruebas similares aplicadas al cuarto punto indican, al compararlos, que probablemente deberían formar parte del intervalo lineal, y un analista podría usar métodos de regresión lineal en el intervalo de absorbancia de 0 – 0.6 con confianza.

REFERENCIAS BIBLIOGRÁFICAS

- **Bartés A. P., Martorell Llabrés. X. T., Cintas. P. G. y Fernandez. L. P. (2000).** Métodos estadísticos. Control y mejora de la calidad. México: Alfaomega Grupo Editor.
- **Box.G.E.P, Hunter.N.G y Hunter. J.S (1989).** Estadística para investigadores. España : Reverté.
- **Capilla Romá C. (1997).** Curso de análisis de regresión, Maestría en estadística. El salvador: Universidad Politécnica de Valencia.
- **Ferré.J. y Rius F.X. (2000).** Introducción al diseño estadístico de experimentos. Tarragona: Departamento de Química Analítica y Química Orgánica. Universitat Rovira i Virgili.
- **Fogler H.S. (2001).** Elementos de Ingeniería de las reacciones químicas. Tercera Edición. México: Pearson Educación.
- **Juran J.M y Gryna F. M. (1998).** Manual de control de calidad. (2 tomos). España: McGRAW-HILL
- **Kenett R. S. y Zacks S. (2000).** Estadística industrial moderna. Diseño y control de la calidad y la confiabilidad. México: International Thomson Editores.
- **Kuehl. R. O. (2001).** Diseño de experimentos. Principios estadísticos de diseño y análisis de investigación. México: International Thomson Editores.
- **Lawson J., Madrigal J.L. y Erjavec J. (1992).** Estrategias experimentales para el mejoramiento de la calidad en la industria. México: Grupo Editorial Iberoamérica.
- **Miller J. C. y Miller J. N. (1993).** Estadística para química analítica. Segunda edición. Wilmington, Delaware, E.U.A: Addison-Wesley Iberoamericana, S.A.
- **Montgomery D. C. (1991).** Diseño y análisis de experimentos. México: Grupo Editorial Iberoamérica.
- **Pérez.C. (2002).** Estadística aplicada a través de Excel. España: Pearson Educación.
- **Perry J. (1998).** Manual del Ingeniero químico, vol. 1. Séptima edición. México: McGRAW HILL.
- **Ramis R.G. (1997).** Curso de Quimiometría. Maestría en Química analítica. España/ Nicaragua: Universidad de Valencia y Universidad Nacional Autónoma de Nicaragua.
- **Romero V. R., Zúñiga L.R.(1993).** Estadística, Proyecto de innovación educativa. España: Universidad Politécnica de Valencia.
- **Skoog.D.A.y West.D.M. (2001).** Química analítica. Séptima edición. México: McGRAW HILL.

CAPITULO VI

ASEGURAMIENTO DE CALIDAD EN LABORATORIOS DE ANALISIS QUIMICO, UNA PERSPECTIVA STATISTICA

CONTENIDO

	Pág.
6.1 Aseguramiento de calidad en laboratorios de análisis químico	6-2
6.2 El muestreo como la llave de la calidad	6-5
6.3 Cálculos en ensayos de colaboración	6-20
6.4 Ejemplos de estimación y separación de varianzas usando ANOVA. Ejemplos de cálculos para ensayos de colaboración	6-22
REFERENCIAS BIBLIOGRAFICAS	6-35
ANEXO A6. Generalidades sobre el muestreo de aire y de agua en el área ambiental	6-36
ANEXO B6. Estimación de los valores de A y s para el calculo de Z-Store.	6-46

6.1 ASEGURAMIENTO DE LA CALIDAD EN LABORATORIOS DE ANÁLISIS QUÍMICO

Cuando se efectúan mediciones y los resultados son utilizados con algún propósito, es obviamente claro que es indispensable tener resultados correctos. Estrictamente hablando, los resultados de toda medición deben poseer dos características básicas: la confiabilidad técnica y que sean legalmente defendibles (PRICHARD et al., 1995, cap. 1).

La confiabilidad técnica de las mediciones implica poseer un profundo entendimiento del proceso de medición, un entendimiento de "qué" necesita ser medido, y un entendimiento de qué se necesita hacer para obtener mediciones confiables. Estos tres puntos relacionan dos cuestiones como son la razón física fundamental del proceso reconocido, así como también los detalles técnicos y físicos de los mecanismos y recursos de medición usados.

Por otra parte, las premisas válidas para que las mediciones sean legalmente defendibles son: Una documentación profunda e imparcial y el uso de procedimientos de reconocimiento en la preparación y conducta de las medidas.

Las buenas prácticas de aseguramiento de la calidad son las bases centrales para alcanzar ambas metas.

El aseguramiento de la calidad (AC) es un conjunto de principios operacionales que, si son estrictamente seguidos durante la recolección de la muestra y el análisis, producirán datos de calidad conocida y defendible. Esto es, la exactitud del resultado analítico puede ser presentada con un alto nivel de confianza.

El aseguramiento de la calidad va acompañado de dos conceptos básicos que son: El control de calidad y La valoración de la calidad (PRICHARD et al., 1995, cap. 1).

El control de calidad es un conjunto de medidas dentro de la metodología de análisis de una muestra que permiten asegurar que el proceso está en estado de control, es decir, es el mecanismo establecido para controlar errores. Mientras que la valoración de la calidad esta relacionada con el mecanismo para mostrar que el procedimiento de medición (analítica) está trabajando dentro de límites bien definidos y aceptables. En otras palabras, la valoración de la calidad es el control de calidad interno y externo para los datos medidos (PRICHARD et al., 1995, cap. 1).

6.1.1 PLANEACION PARA EL ASEGURAMIENTO DE LA CALIDAD

La planeación para el aseguramiento de la calidad consiste en establecer un conjunto de principios operacionales que constituirán el *programa de aseguramiento de la calidad*. La preparación de un programa para el AC incluye lo siguiente (GARFIELD, 1984, cap. 1):

- a) Hoja explicativa con las firmas de aprobación del plan.
- b) Determinación de la organización y responsabilidades del personal.
- c) Control de las muestras y la documentación de procedimientos.
- d) Procedimientos estándar de operación para cada método analítico (PEO).
- e) Requerimientos de capacitación de los analistas.
- f) Procedimientos de mantenimiento preventivo de los equipos.
- g) Procedimientos de calibración.
- h) Acciones correctivas.
- i) Actividades de control de calidad interno.
- j) Desarrollo de auditorías.
- k) Procedimientos para el aseguramiento y precisión de los datos.
- l) Reducción, validación y reporte de datos.

La hoja explicativa con las firmas de aprobación del plan indica que el plan ha sido revisado y juzgado convenientemente, así como también las responsabilidades y organización de la sección fuera de línea en la cadena de mando, y las asignaciones de funciones específicas para cada persona involucrada.

El control de muestras y la documentación de procedimientos permiten dar seguimiento a una muestra y sus derivados a través de todos los pasos desde su colección hasta el análisis y reporte de resultados.

La documentación siempre es importante pero lo es especialmente cuando es impuesta una cadena de requerimientos custodiada.

Los procedimientos estándar de operación (PEO) describen el método con detalles tales, que aún un analista no familiarizado con el mismo puede llegar a obtener resultados aceptables. Los requerimientos de entrenamiento para los analistas deben ser

especificados. El número de análisis requerido y la incertidumbre de los resultados variará para los distintos tipos de análisis, características de muestras y la experiencia del analista.

Un programa estricto de mantenimiento de equipo reducirá sensiblemente el mal funcionamiento de los mismos, el mantenimiento de la calibración y el tiempo perdido.

Los procedimientos de calibración, acciones correctivas, actividades de control de calidad interno, desarrollo de auditorías y la valoración del sesgo y precisión de los datos son descritos en los temas de control y valoración de la calidad.

La reducción, validación y reporte de datos son las características finales de un programa de AC. Las lecturas obtenidas para un instrumento analítico deben ser ajustadas para factores tales como eficiencia del instrumento, eficiencia de extracción, tamaño de muestra y el valor de fondo, antes de llegar a dar un resultado útil. El plan de AC especifica los factores de corrección que deben ser aplicados así como los pasos a seguir para validar el resultado. El reporte de resultados debe hacerse en unidades estándar de masa, volumen o concentración. El uso de un método prescrito debe ir seguido del límite de detección del método. Cada resultado o conjunto de resultados debe acompañarse de la declaración de su incertidumbre.

En la figura 6.1, se presenta, el esquema de la relación de los elementos en un plan de AC.

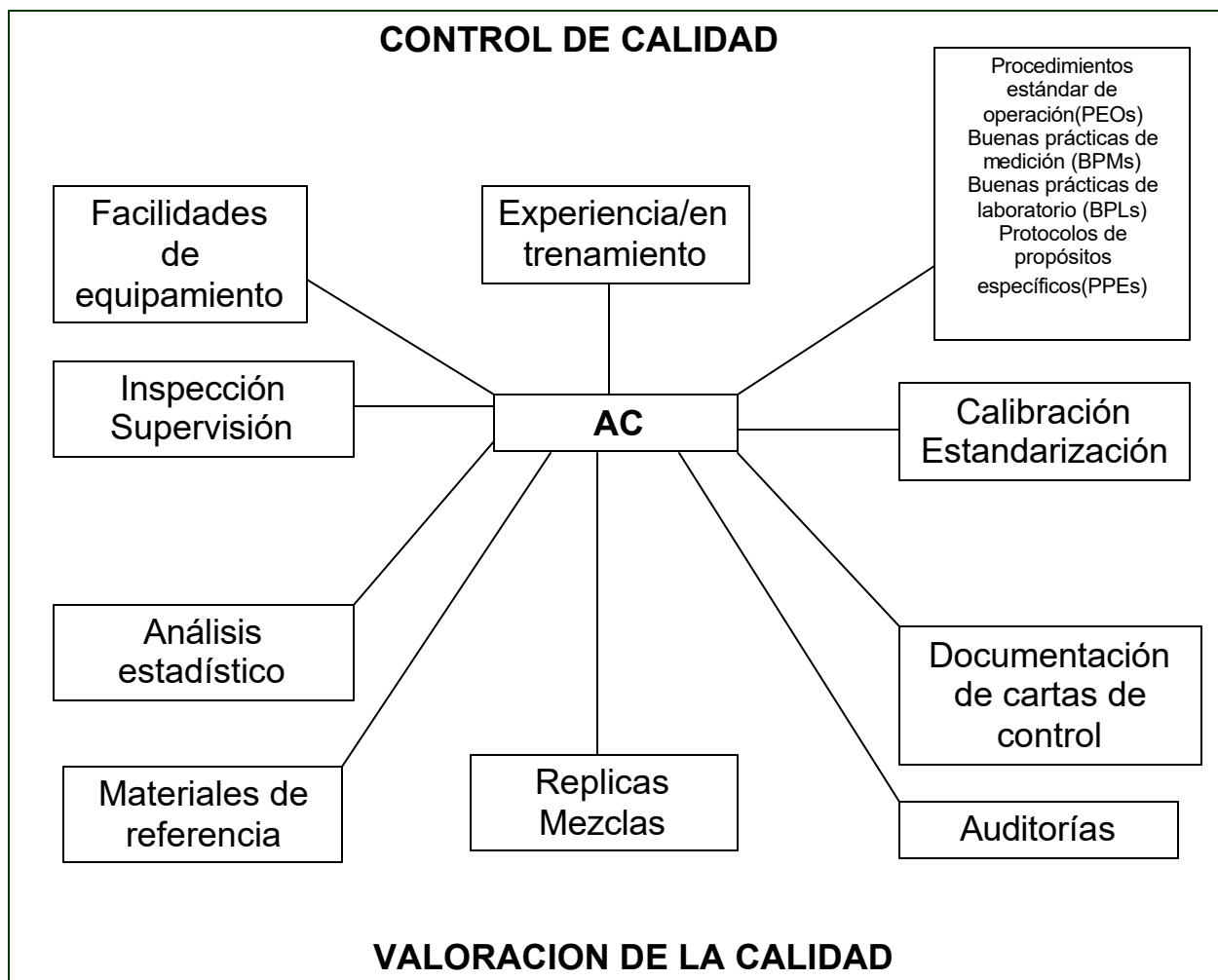


FIGURA 6.1 ELEMENTOS DE UN PLAN DE ASEGURAMIENTO DE CALIDAD
(PRICHARD et al., 1995, cap. 1)

6.2 EL MUESTREO COMO LLAVE DE LA CALIDAD

Existen tres actividades básicas para solucionar un problema analítico (GARFIELD, 1984, cap. 6):

- La recolección de una muestra representativa.
- La preparación de la muestra para el análisis.
- El análisis usando métodos apropiados.

Estas actividades son independientes una de la otra. Sin embargo, una puede tener una influencia decisiva en la otra. Puesto que en cada una de estas actividades existe la probabilidad de cometer errores, se deben tomar medidas para identificarlos y evitar que ocurran. Los laboratorios deben desarrollar un plan para el desempeño adecuado de cada actividad, y luego establecer los estándares de calidad y procedimientos escritos que satisfagan estos estándares. La elaboración de planes apropiados dependerá del entendimiento de los problemas involucrados en cada actividad, seguido de la aplicación de criterios razonables en la búsqueda de soluciones (GARFIELD, 1984, cap. 6).

La calidad de un análisis empieza en la toma de muestra. Siempre que la toma de muestra no se realice correctamente, todos los métodos estadísticos, así como los resultados del análisis posterior, no tendrán utilidad alguna.

Cualquier analista debe trabajar con muestras, ya que en la mayor parte de los casos es imposible analizar todo el objeto que se está considerando. Por ejemplo, es imposible analizar las grasas de un tanque lleno de leche, o todo el cuerpo de un río, para determinar el contenido de contaminantes (MILLER y MILLER, 1993, cap. 4).

Si el muestreo presenta riesgos, el análisis es incapaz de detectarlos, y mucho menos de eliminarlos. El analizador más sofisticado no podrá conducirlos a buenos resultados. Así, el análisis exacto de una muestra sesgada proporcionará, inevitablemente, un resultado sesgado (GARCIA, 1998, cap. 5).

En función de la constitución de los objetos que se van a muestrear, el muestreo será una operación más o menos difícil. La toma de muestra más difícil corresponde a los sólidos heterogéneos, especialmente si se pretende analizar un componente que se encuentra en pequeña cantidad y repartido desigualmente, por ejemplo, minerales de metales preciosos (GARCIA, 1998, cap. 5).

Si la materia es homogénea, se puede decir que todas sus partes son idénticas entre sí, y cualquier fracción del lote, por pequeña que sea, será exactamente representativa de todas sus propiedades. Los materiales jamás serán rigurosamente homogéneos, por lo que la operación de muestreo nunca será exacta, y llevará siempre asociado un error determinado (GARCIA, 1998, cap. 5).

En las normas sobre las buenas prácticas de laboratorio (BPLs) de la internacional Organization for Standardization (ISO 9000) o de la European Norms (EN 45000), se dice que los químicos deberán analizar muestras representativas. Sin embargo, en los

documentos estandarizados no se define que es una muestra representativa, ni se indica como obtenerla (GARCIA, 1998, cap. 5).

6.2.1 DEFINICIONES

La guía UNE 007 define la terminología utilizada en el proceso de toma de muestras (GARCIA, 1998, cap. 5):

- **Partida:** Cantidad de un producto expedida o recibida de una sola vez, en el marco de un contrato particular o de un momento de transporte.
- **Lote:** cantidad determinada de un producto, presumiblemente de las mismas características, identificada en el seno de una partida y que puede permitir conocer la idoneidad de esta, en un periodo estimado de tiempo.
- **Muestra:** Parte o porción de características, tomadas de un conjunto por métodos que permiten considerarla como representativa del mismo. Una muestra representativa, es por lo tanto, aquella parte o porción de características que se supone están asociadas a un material que tiene la misma composición y/o características del producto sometido a muestreo, cuando se le considera como un todo homogéneo.
- **Muestreo:** Acción material de toma de muestras, siguiendo el plan de muestreo previamente establecido.
- **Plan de muestreo:** Procedimiento a seguir, planificado, para la selección, la toma y el tratamiento de una o varias muestras de un lote, con el fin de obtener a partir de la muestra final, la información que interesa sobre una o varias características, para adoptar una decisión sobre el lote.
- **Unidad de muestreo:** Cantidad determinada de material obtenida del conjunto, y que se encuentra dentro de unos límites físicos (masa, envase, probeta, etc.), o

bien, dentro de límites de naturaleza convencional (por ejemplo, un intervalo de tiempo).

- **Muestra global:** Cantidad de un producto obtenida, reuniendo si es posible y mezclando, todas las tomas de muestra elementales obtenidas en un lote determinado.
- **Muestra elemental:** Cantidad de un producto o mercancía tomada de un punto del lote de una sola vez, en una unidad de muestreo.
- **Muestra para laboratorio:** El conjunto de unidades que constituyen la muestra, reglamentaria o concertada, sobre la que van a realizarse los análisis o ensayos previos, para conocer la idoneidad, aptitud de uso o calidad de un producto. Esta muestra en casos reglamentarios debe estar compuesta, como mínimo, por tres partes: una de ellas para verificar el análisis inicial, otra el contradictorio y la tercera, el dirimente, si procede.
- **Muestra para análisis:** Cada una de las partes de la muestra para laboratorio, sobre la que se van a realizar las pruebas analíticas, o los ensayos previamente seleccionados.
- **Toma para ensayo:** Fragmento o fracción representativo de una muestra para análisis, debidamente homogeneizada en su caso mensurable, sobre el que se lleva a cabo la determinación de la especificación técnica a valorar, mediante el método analítico o de ensayo normalizado. En determinados casos, y fundamentalmente para sectores específicos y ensayos de análisis físicos, esta toma para ensayo se denomina también probeta.

6.2.2 EL OBJETIVO DE LA TOMA DE MUESTRA

Una muestra es una parte de un conjunto que representa, con más o menos exactitud, las propiedades que se quieren medir de todo el conjunto. Este conjunto puede ser muy

variado, desde un material en el que se pretende determinar un contaminante a nivel de nanogramos, hasta una selección de candidatos para una encuesta.

La toma de muestra o muestreo consiste en la obtención de una muestra representativa de todo el conjunto, con respecto a la propiedad que se quiere medir. La sencillez de la definición no debe confundir, respecto a la dificultad de su realización (GARCIA, 1998, cap. 5).

El primer paso y fundamental, en una toma de muestra, es establecer claramente cual es el objetivo de la misma. Los objetivos más comunes son (GARCIA, 1998, cap. 5):

- i) Conocimiento cualitativo del material.
- ii) Obtener información sobre problemas específicos. En determinados casos, pueden interesar algunos constituyentes particulares. Se deben entonces diseñar el régimen de toma de muestras, con atención particular a dichos constituyentes, cuya medida deberá tener una fiabilidad y precisión específicas.
- iii) Satisfacer un contrato de compraventa. El régimen de muestreo deberá ser diseñado específicamente para este contrato.

Una vez establecido el objetivo de la toma de muestras, se podrá confeccionar ya el plan para obtener la máxima exactitud en la información, al mínimo coste posible. El plan de muestreo es un procedimiento que define la cantidad de muestra, forma de extracción, número de extracciones, estado físico del material, operaciones de homogenización y tratamientos posteriores, que garanticen la representatividad de la muestra final (GARCIA, 1998, cap. 5).

6.2.3 TIPOS DE MUESTREO

Existen dos grandes tipos de muestreo:

- a) *muestreo de aceptación (laboratorio de control)*
- b) *muestreo de caracterización (laboratorio analítico)*

a) Muestreo de aceptación

Es el muestreo típicamente utilizado, cuando se tienen unidades discretas. Se trata de obtener una muestra de varias unidades de un conjunto, representativas del mismo respecto a la característica de ensayo o análisis (JURAN y GRAYNA, 1997, cap. 25). En función de esta característica, se habla de (GARFIELD, 1985, cap. 6):

a.1) Muestreo de atributos

Un atributo es una característica de la muestra, por ejemplo, color olor, etc., sobre la cual se establece el control de calidad del objeto estudiado (correcto o incorrecto).

a.2) Muestreo por variables

La evaluación se hace midiendo una propiedad del material a la cual se le puede asignar un valor numérico, por ejemplo, densidad y pH. El plan de muestreo exige la extracción de un número de elementos, que es función del número de unidades del conjunto, de su homogeneidad y del nivel de confianza requerido.

Existen diversas normas oficiales aplicables a este tipo de muestreo: normas UNE 66020 73, 63020 88, 66021 74 (referentes al muestreo por atributos), 66030 84 (muestreo por variables). En función del tamaño de la muestra, la norma indica el número de unidades que deben tomarse, estableciendo los niveles de tolerancia en relación al número de defectos (GARCIA, 1998, cap. 5).

b) Muestreo de caracterización

Se pretende determinar la calidad media del lote, respecto a las variables a estudiar. Este es el muestreo analítico. Se supone que se tiene una población que sigue una distribución normal o fácilmente normalizable, y se obtiene un valor medio de la variable. La característica básica a considerar en el muestreo de caracterización es la heterogeneidad del material, y de acuerdo con ella, deberá establecerse el tamaño de la muestra (GARCIA, 1998, cap. 5).

Los parámetros a tener en cuenta en el plan de muestreo son (GARCIA, 1998, cap. 5):

- Cantidad de muestra.
- Sistema de homogeneización.
- Numero y localización de las extracciones.
- Tamaño de partícula, en el caso de sólidos pulverulentos.

6.2.4 EL PLAN DE MUESTREO

Antes de efectuar un muestreo, se deben definir las características principales del plan de muestreo mediante la realización de ensayos previos, o bien, haciendo uso de la experiencia anterior.

Las principales características de cualquier esquema de muestreo son (GARCIA, 1998, cap. 5):

- i) Exactitud requerida: siempre hay errores de muestreo y análisis. Debe efectuarse un balance entre ambos tipos de errores, respecto a los costes relativos y el valor intrínseco del material.
- ii) El método de muestreo: la experiencia suele determinar la técnica a utilizar, perforaciones, palas, disolución, etc.
- iii) El tamaño de la muestra global: se determina estableciendo, primero, el tamaño de la muestra elemental y el número de éstas que la componen.
- iv) Tratamiento de la muestra: una vez tomada la muestra, se suele mezclar bien y se trata como una masa relativamente homogénea. Se han desarrollado varias teorías para calcular relaciones entre tamaño de partícula, peso de la muestra y errores relativos. Este parámetro vendrá también condicionado por el método de análisis.
- v) Evaluación: en cualquier procedimiento de muestreo, se deben incluir pruebas que garanticen la calidad de la muestra obtenida.
- vi) Exactitud requerida: En general, cuanto mayor sea la exactitud, mayor será el coste del muestreo. Debe evaluarse considerando el valor del material y las consecuencias del error.

6.2.4.1 MODELOS DE EXTRACCIÓN DE LA MUESTRA

La forma de extracción de la muestra elemental depende del tipo de material, de la manera de transportarlo, así como del objetivo del muestreo. Existen distintos tipos de muestreo, que deben seleccionarse antes de elegir el peso de la muestra primaria. Estos son (PRICHARD et al., 1995, cap. 2):

a) Muestreo aleatorio

La muestra elemental se selecciona, de manera que todas las partes del material tengan la misma probabilidad de ser elegidas. Esto se hace, normalmente, con la ayuda de tablas de números aleatorios.

Esta técnica es de utilidad cuando hay poca información acerca del material estudiado. Es el muestreo típicamente analítico.

b) Muestreo sistemático

Las muestras elementales se recogen a intervalos previamente determinados y que han sido previstos en el plan de muestreo. Los intervalos no tienen que ser necesariamente iguales. Por ejemplo, un muestreador mecánico recoge las muestras a intervalos regulares, por lo tanto, hace un muestreo sistemático. Cuando se da un patrón de segregación conocido, el muestreo sistemático puede utilizarse para minimizar sus efectos.

El muestreo sistemático reduce el esfuerzo, por lo que es muy utilizado, si bien puede resultar peligroso cuando se dan variaciones periódicas en el material, de manera que el plan sistemático puede coincidir con ellas y no detectarlas.

c) Muestreo estratificado

Es una extensión del anterior, que involucra la división del lote en grupos. Los subgrupos se muestrean en proporciones a sus pesos estadísticos en el lote.

Este tipo de muestreo es útil cuando la partida está constituida por materiales diferentes, que no se pueden mezclar de manera homogénea. Un ejemplo típico sería el muestreo de un líquido que contiene sedimentos, o en una cantera, el muestreo de un mineral que se presenta en capas.

d) Muestreo de dos pasos

Es útil cuando se tienen lotes muy grandes de material, que no merecen un muestreo estratificado profundo. Por ejemplo, si se tiene una carga de mineral de 20 vagones, se pueden seleccionar primero cinco de ellos, en los cuales se realiza a continuación el muestreo.

e) Muestreo secuencial

Los esquemas secuenciales se utilizan cuando se muestrea siguiendo una prueba de especificación, con objeto de reducir esfuerzos.

El muestreo secuencial consiste en utilizar pocas muestras elementales para determinar la calidad del lote. Si el material está claramente fuera o dentro de la especificación, el lote se rechaza o acepta; si el resultado es próximo al nivel de especificación, se procederá a un muestreo más exhaustivo.

f) Híbridos

En muchos casos, se tienen situaciones intermedias. Por ejemplo, un esquema estratificado puede ser aleatorio o sistemático, dependiendo de las bases empleadas para seleccionar las muestras elementales, dentro de cada estrato.

6.2.5 TIPOS DE UNIDADES DE MUESTREO EN ANÁLISIS QUÍMICO

En análisis químico puede tratarse sobre todo, con dos tipos de materiales, que son (JURAN y GRAYNA, 1997, cap. 25):

- a. Materiales a granel claramente fraccionado (o "envasado" en unidades discretas).
- b. Materiales a granel no fraccionados.

Considere el caso en el que se tiene un lote de botes con jugo para bebé y se desea establecer el volumen de llenado medio de los mismos. En lugar de medir el volumen del jugo contenido en todos los botes del lote, se tomarán unos cuantos, cuatro por ejemplo, y se mide el volumen de jugo. En este ejemplo, los botes de jugo constituyen unidades discretas obvias, las cuales nominalmente son las mismas, o sea, los botes de jugo, y en donde los volúmenes de llenado de todos los botes de jugo del lote forman una población y los cuatro volúmenes medidos constituyen una muestra de dicha población.

Si la muestra se va a utilizar para deducir propiedades de la población, debe ser lo que se considera una muestra aleatoria, esto es, una muestra tomada de tal manera que todos los miembros de la población tengan una probabilidad conocida de ser incluidos en la misma (ver capítulo 2).

Debe tenerse en cuenta que el término aleatorio tiene, en sentido estadístico, un significado diferente de casual. Aunque en la práctica, un analista puede repartir todos los

botes de jugo en una mesa, e intentar cuantificar el volumen de jugo de cuatro de ellos de una manera casual, tal método podría conducir a un sesgo. La mejor manera de obtener una muestra aleatoria es mediante el uso de la tabla de números aleatorios (como la presentada en el anexo A4). A cada miembro de la población se le asigna un número, y los componentes de la muestra se extraen a partir de una tabla de números aleatorios, empezando en un punto arbitrario.

Una alternativa más simple consiste en seleccionar los miembros de la población a intervalos regulares, por ejemplo, tomar un bote de cada cincuenta en k unidades, en la línea de producción. Este método no es del todo satisfactorio, ya que podría, por ejemplo, darse una coincidencia periódica en el volumen de llenado de los mismos (MILLER y MILLER, 1993, cap. 4).

El muestreo de materiales como rocas, polvo, gases y líquidos, para los que esto no se cumple, se denomina muestreo **a granel**. Se llamará **incremento de muestra** a una pequeña parte tomada del todo continuo y a un grupo de tales incrementos se le llamará **muestra bruta** (MILLER y MILLER, 1993, cap. 4).

Si un material a granel fuera perfectamente homogéneo, sólo sería necesario un incremento de muestra para determinar las propiedades de la totalidad. En la práctica los materiales a granel no son homogéneos por diversas razones; por ejemplo, las aleaciones y los sedimentos están formados por partículas microscópicas con diferente composición y puede ser que no estén uniformemente distribuidas en el conjunto, o puede ser que los fluidos no sean homogéneos en una escala molecular, debido a los gradientes de concentración. Tal falta de homogeneidad se detecta sólo tomando incrementos de muestra de diferentes partes del conjunto. Para obtener una muestra bruta aleatoria deberían tomarse incrementos en puntos del conjunto elegidos al azar. Esto puede hacerse, por ejemplo, si se considera el grueso del conjunto como una colección de celdillas y si se seleccionan las celdillas que van a ser muestreadas, utilizando números aleatorios como se describió anteriormente. Sin embargo, dada la enorme variedad de posibilidades en cuanto a objetivos y procedimientos analíticos, en el muestreo de materiales a granel, no existen normas estandarizadas generales que permitan seguir un patrón de calidad, a la hora de establecer un plan de muestreo. De hecho, el muestreo de estos materiales, es un campo de continuo desarrollo, y que sirve actualmente como base a un número importante de proyectos de investigación (GARCIA, 1998, cap. 5). En el anexo A6, se indican aspectos relativos tanto al muestreo de aire, como al de aguas en el

área ambiental, que de manera general podrían considerarse en cualquier plan de muestreo.

Los aspectos especiales que diferencian al muestreo a granel del muestreo de unidades discretas son (JURAN y GRANA, 1997, cap. 25):

1. La posibilidad de hacer una mezcla física y la subsiguiente reducción (o submuestreo) que suele ser necesaria.
2. La necesidad de utilizar, en muchos casos, un dispositivo mecánico de muestreo para obtener las distintas fracciones que formarán la muestra. En este caso es probable que esas fracciones se "creen" mediante el uso del dispositivo de muestreo y no puedan considerarse como preexistentes.

La figura 6.2 muestra la relación entre las diferentes operaciones involucradas en el esquema de muestreo y análisis. Este también ayuda a identificar algunos de los términos expuestos en la sección 6.2.1.

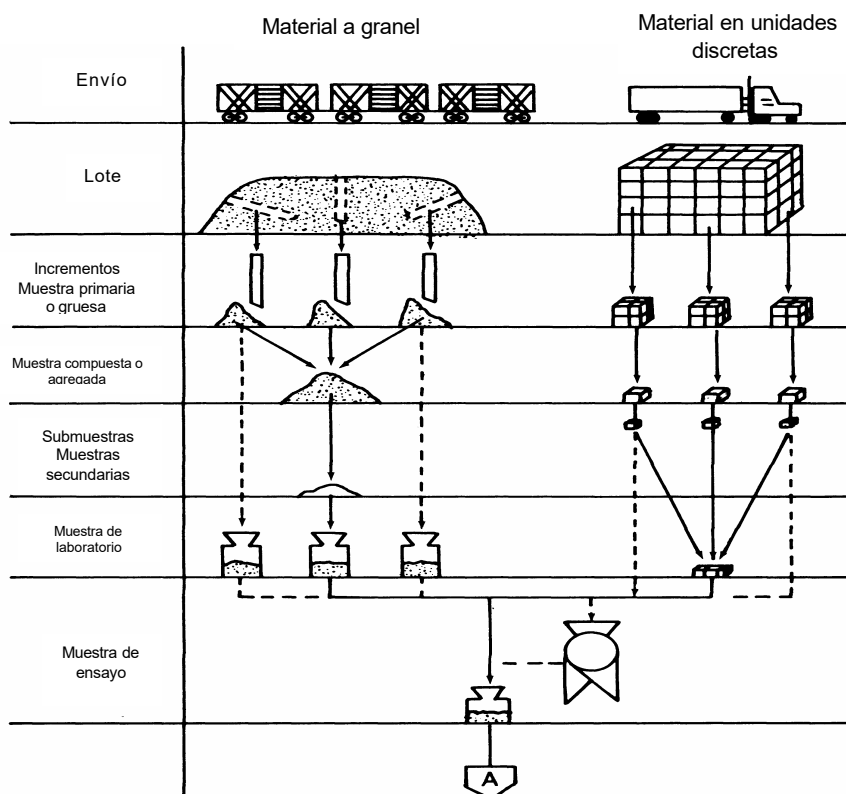


FIGURA 6.2 ESQUEMA DE LAS OPERACIONES DE MUESTREO
(PRICHARD et al., 1995, cap. 1)

6.2.6 TEORÍA DEL MUESTREO

Todos los esfuerzos de la normativa existente acerca del muestreo, se centran en asegurar un conocimiento de los parámetros estadísticos, tales como precisión y exactitud de las medidas que se realicen, a partir de la muestra tomada. En otras palabras, los límites de confianza de los valores que se obtengan para cada parámetro en relación al conjunto total muestreado. Esto exige una toma de muestras "representativas", es decir, una muestra recogida de manera que todas las partículas del conjunto tengan la misma probabilidad de ser seleccionadas (GARCIA, 1998, cap. 5).

El muestreo es una selección que puede ser o no probabilística. Se dice probabilística, cuando se someten todos los elementos al proceso de selección, con una probabilidad distinta de cero de ser seleccionados, y no probabilística, cuando no se cumple esta condición (ver capítulo 2). Por ejemplo, cuando a una fracción del lote se le da una probabilidad cero de ser seleccionada.

No puede establecerse una teoría de muestreo no probabilístico. La razón y la experiencia demuestran que el muestreo no probabilístico es peligroso, ya que siempre genera errores incontrolados. Esto debería evitarse a toda costa, pero por desgracia, es una situación omnipresente en todas las ramas de la investigación, industria y comercio (ver capítulo 2). La selección probabilística puede ser o no correcta. Es *correcta* cuando al mismo tiempo todos los elementos del lote se someten al proceso de selección, con una probabilidad conocida de ser elegidos, y la muestra y sus elementos permanecen inalterados en el proceso. Esta última condición cubre pérdidas, contaminaciones, alteraciones físicas o químicas, tanto de manera accidental como intencionada. Es *incorrecta* cuando no se cumple cualquiera de estas condiciones (GARFIELD, 1984, cap. 6).

En la práctica, los resultados de esta selección deben jugarse en términos de errores de muestreo. Bajo el punto de vista estadístico, cada error aparece como una variable aleatoria que se suele definir con la ayuda de sus momentos sucesivos, y en particular, de los dos primeros, es decir (GARCIA, 1998, cap. 5):

- i) Primer momento o media: que se traduce como la tendencia central de la distribución del error. Cuando esta media no es nula, el error presenta un carácter sistemático y se le denomina error sistemático o sesgo. Como se ha dicho anteriormente, afecta a la exactitud del muestreo.

- ii) Momento centrado de orden 2 o varianza: que traduce la dispersión del error alrededor de su media. La varianza caracteriza la parte puramente aleatoria del error, y se designa por error aleatorio. Afecta a la precisión del muestreo. La raíz cuadrada de la varianza es la desviación típica de la distribución.

El valor exacto de un error de muestreo no se conoce nunca. Se puede definir sumariamente con la ayuda de (PRICHARD et al., 1995, cap. 2):

- i) Un cierto intervalo de confianza.
 ii) La probabilidad de que un error observado caiga en ese intervalo.

6.2.7 ESTIMACION Y SEPARACION DE VARIANZAS

Cuando se determina un analito en materiales a granel, el trabajo se hace generalmente en dos etapas, la primera de las cuales es tomar una muestra de un "lote", y la segunda es el análisis químico de la muestra obtenida. Debido a estos procedimientos, existen dos fuentes básicas de error en la estimación del analito. Uno es **el error debido al muestreo**, y ocurre cuando la muestra se toma del lote, y el otro es **el error debido al análisis o al proceso de medición**, el cual ocurre cuando se "mide" el analito (HITOSHI, 1997, cap. 8). Cuando el error de muestreo se expresa como e_s y el error de medición como e_M , los datos medidos x se pueden expresar como (HITOSHI, 1997, cap. 8):

$$x = \mu + e_s + e_M \quad (\text{ec. 6.1})$$

donde μ es el valor verdadero del analito si la media de e_s es μ_{es} y la de e_M es μ_{eM} , la media de x esta dada por (HITOSHI, 1997, cap. 8):

$$\mu_x = \mu + \mu_{es} + \mu_{eM} \quad (\text{ec. 6.2})$$

donde μ_{es} y μ_{eM} son conocidos como el sesgo de muestreo y de medición respectivamente. La varianza de x esta dada por (HITOSHI, 1997, cap. 8):

$$\sigma_x^2 = \sigma_{es}^2 + \sigma_{eM}^2 \quad (\text{ec. 6.3})$$

donde σ_{es}^2 es la varianza del muestreo y σ_{eM}^2 es la varianza debida a la medición.

Las varianzas son aditivas, por lo tanto, estas dos varianzas pueden separarse y estimarse usando ANOVA.

Cuando un analito se mide "n" veces y se toma la media, la varianza de la media viene dada por (MILLER y MILLER, 1993, cap. 4):

$$\sigma_{\bar{x}}^2 = \sigma_{es}^2 + \frac{\sigma_{eM}^2}{n} \quad (\text{ec. 6.4})$$

La varianza del error de medición se divide "n" veces. Si se toman "m" incrementos de muestra y se hacen "n" mediciones de cada una de ellos, la varianza estará dada por (MILLER y MILLER, 1993, cap. 4):

$$\sigma_{\bar{x}}^2 = \frac{\sigma_{es}^2}{m} + \frac{\sigma_{eM}^2}{nm} \quad (\text{ec. 6.5})$$

Es obvio que para alcanzar la precisión máxima se requiere que la varianza de la media sea tan pequeña como fuera posible. El término debido a la varianza de la medición puede ser reducido ya sea utilizando un método de análisis más preciso o incrementando "n", el número de mediciones repetidas, pero no hay que esforzarse por hacer que la varianza de la medición sea menor que un décimo de la varianza muestral, ya que cualquier reducción posterior no mejorará en absoluto la varianza total. Es preferible, en cambio, tomar un número mayor de incrementos de muestra ya que el intervalo de confianza disminuye cuando aumenta "m". Si se ha utilizado una muestra preliminar para estimar σ , entonces el tamaño de la muestra requerido para lograr un tamaño dado del intervalo de confianza se puede calcular de manera aproximada .

En general, los errores pueden clasificarse en dos categorías sesgo y variación. Mientras que la variación se reduce cuando se repiten las mediciones, el sesgo permanece igual sin importar cuantas observaciones se repiten, y la obtención de la media no lleva a su reducción (MILLER y MILLER, 1993, cap. 4).

Una estrategia de muestreo con materiales a granel es tomar "m" incrementos de muestra y mezclarlos antes de hacer "n" mediciones repetidas. La varianza de la media de estas mediciones es (MILLER y MILLER, 1993, cap. 4):

$$\sigma_{\bar{x}}^2 = \frac{\sigma_{es}^2}{m} + \frac{\sigma_{eM}^2}{n} \quad (\text{ec. 6.6})$$

Esta varianza total debería compararse con la resultante de analizar "n" veces cada incremento de muestra y promediar las medias de los incrementos, siendo entonces la varianza la presentada en la ecuación 6.5.

El conocimiento de los valores de σ_{es}^2 y σ_{eM}^2 de experiencias anteriores, y los costos de muestreo y análisis, pueden utilizarse para calcular el costo de las estrategias de muestreo relativas. Por lo general, se usará el esquema más económico que proporcione el grado de precisión requerido. En el caso de materiales a granel la varianza muestral depende del tamaño del incremento de muestra relativo a la escala de no homogeneidades, y disminuye con el aumento del tamaño del incremento de muestra. En algunas experiencias puede ser necesario seleccionar un límite superior de la varianza muestral, de manera que puedan detectarse cambios en la media. Pueden tomarse mediciones preliminares para decidir el tamaño del incremento de muestra mínimo para proporcionar un nivel aceptable de varianza muestral (GARCIA, 1998, cap. 5).

6.2.8 NUMERO DE MUESTRAS ELEMENTALES

Si se realiza un análisis de cada uno de los h incrementos de muestra, entonces los límites de confianza de la media están dados por la ecuación (MILLER y MILLER, 1993, cap. 4):

$$\mu = \bar{x} \pm t(s/\sqrt{n}) \quad (\text{ec. 6.7})$$

donde \bar{x} es la media de las mediciones y s la desviación estándar de ellas que se estima de acuerdo a lo planteado en el tema 6.2.7; t es la variable t-Student definida por el nivel de confianza admitido para (n - 1) grados de libertad.

La predicción que se hace de la variabilidad del material, y de los límites de confianza, dependen de que la población tenga una distribución normal. En la práctica, no se da esta situación. Por ejemplo, en el caso de un lote de material que ha sido inspeccionado, siguiendo un control de calidad previo, o una mezcla de distintos materiales no bien homogenizados, pueden encontrarse distribuciones binomiales. En estos casos, se utilizan aproximaciones estadísticas (desigualdades de Tchebycheff y de Camp- Meidell), para realizar los cálculos (MILLER y MILLER, 1993, cap. 4).

Como los límites de confianza de la media vienen dados por la ecuación 6.7, se requiere que la desviación típica, s , sea lo menor posible.

6.3 CALCULOS EN ENSAYOS DE COLABORACION

Una de las herramientas más potentes para detectar y corregir fuentes de error debidas a una técnica particular de análisis, o a un fallo en el control de calidad dentro de un laboratorio, la constituye la participación en ejercicios intercomparativos o interlaboratorios (JURAN y GRZYNA, 1997, cap. 26).

Los programas de prueba interlaboratorio se clasifican en pruebas de competencia, exámenes interlaboratorios, muestras de verificación y programas "round-robin". En estos programas, organizaciones externas a los laboratorios (organismos de arbitraje) distribuyen las muestras para su análisis. Por lo general, son programas de participación voluntaria, usados como parte de un sistema de acreditación o certificación de los laboratorios, para evaluar su competencia técnica, o para elevar la calidad total del desempeño del laboratorio (GARFIELD, 1985, cap. 7).

Los exámenes colaborativos son una forma especial de pruebas para evaluar el desempeño del método en condiciones reales de trabajo en varios laboratorios, por medio de análisis de muestras homogéneas cuidadosamente preparadas. La participación en un programa de pruebas interlaboratorios puede reportar muchas ventajas: proporciona un medio para comparar el trabajo del laboratorio con el de otros laboratorios; puede estimular la autoevaluación y reducir los errores del laboratorio; proporciona evidencia externa de la calidad del desempeño analítico del laboratorio y la competencia individual de los analistas; y puede ser usada para reducir la frecuencia de pruebas interlaboratorios cuando se alcanzan resultados consistentemente favorables en pruebas de muestras interlaboratorios. Aún más, combinando con otras partes del esfuerzo de garantía de

calidad, las pruebas íterlaboratorio pueden ayudar a identificar necesidades de capacitación y cambio de procedimientos del laboratorio (GARFIELD, 1985, cap. 7).

Generalmente, los programas voluntarios de pruebas de competencia interlaboratorios tienen una organización o laboratorio coordinador que proporciona las muestras que serán examinadas por los laboratorios participantes. Este tipo de programas tiene las siguientes características (GARFIELD, 1985, cap. 7):

- El laboratorio coordinador prepara las muestras homogéneas y distribuye porciones a los laboratorios participantes para su análisis (de 10 a 20 laboratorios).
- La muestra puede o no estar acompañada por un método estándar. Si no se utiliza el método estándar, normalmente se exige al laboratorio analista que describa el método usado.
- El laboratorio analista debe examinar las muestras dentro de un tiempo específico y enviar los resultados a la organización patrocinadora a través del director del laboratorio.

Cuando se concluye el trabajo experimental de un ensayo de colaboración y se envían los datos al organismo de arbitraje, debe efectuarse la evaluación estadística de los resultados. Estos cálculos deberían revelar hasta que punto la variación entre los laboratorios participantes se debe a errores sistemáticos. En caso de que los resultados sean deficientes, el hecho se comunica al laboratorio, al que se le pueden hacer sugerencias para mejorar su desempeño (GARFIELD, 1985, cap. 7).

Existen diferentes métodos para la evaluación de los resultados dados por los participantes, dependiendo de la finalidad del estudio de intercomparación. Generalmente, se basan en el diagnóstico gráfico y en procedimientos de cálculo de estadísticos como la media, la mediana, la desviación estándar de los datos enviados, etc. Sin embargo, la complejidad de los métodos empleados dependerá muchas veces del número de muestras enviadas a los laboratorios (JURAN y GRAYNA, 1997, cap. 26), ver tabla 6.1.

TABLA 6.1 TECNICAS ESTADISTICAS EMPLEADAS EN ENSAYOS DE COLABORACION

Número de muestras enviadas	Análisis estadístico
1. Una sola muestra	Aplicación simple de ANOVA de un factor. Comparación de tratamientos con un control. Gráficos de control
2. Dos muestras	<i>Técnicas paramétricas:</i> ANOVA de dos factores Análisis de sumas y diferencias <i>Técnicas no paramétricas</i> Técnica gráfica de Youden
3. Más de dos muestras	Técnicas de clasificación, procedimiento debido a Youden (1967). Gráficos de control para cada muestra.

En los ejemplos E6.2 y E6.3 se presenta el análisis estadístico de datos, asociado a dos diferentes ensayos de colaboración.

6.4 EJEMPLOS DE ESTIMACION Y SEPARACION DE VARIANZAS USANDO ANOVA Y DE ENSAYOS DE COLABORACION

6.4.1 ESTIMACION Y SEPARACION DE VARIANZAS

1. La aplicación de pesticida a menudo es parte de programas de manejo de insectos usados en Agronomía u Horticultura. Una de las preocupaciones que sigue a la aplicación de pesticidas es la concentración de sus residuos que permanecen en las plantas después de cierto período. Los residuos de pesticidas se evalúan con análisis químicos de laboratorio en muestras de plantas de los cultivos tratados con

pesticidas. Se obtuvieron muestras en campo de tres lotes de plantas y se prepararon para el análisis de residuos, haciéndose dos mediciones para cada lote. Estime los componentes de varianza. En la tabla E61.1 se presentan los residuos de pesticidas determinados para cada muestra (MILLER y MILLER, 1993, cap. 4).

TABLA E61.1 RESIDUOS DE PESTICIDAS EN μg ENCONTRADOS EN MUESTRAS DE PLANTAS DE ALGODÓN

MUESTRA	RESIDUOS DE PESTICIDA(μg)
1	120, 110
2	120, 100
3	140,130

SOLUCION:

Como se explicó en el tema 6.2.7, existen dos posibles fuentes de variación: aquella debida al error aleatorio en la medición de los residuos de pesticida, dada por la varianza en la medición, σ_{eM}^2 , y la que se debe a las variaciones reales en el contenido de residuos de pesticida en las diferentes muestras, dada por la varianza muestral, σ_{es}^2 . Puesto que el cuadrado medio dentro de las muestras no depende de la media muestral puede usarse para estimar σ_{eM}^2 . El cuadrado medio entre las muestras no se puede utilizar para estimar σ_{es}^2 en forma directa, ya que la variación entre medias muestrales se debe al error aleatorio en la medición y a las posibles variaciones en el contenido de residuos de pesticida. Se demuestra que el cuadrado medio entre las muestras proporciona una estimación de $\sigma_{eM}^2 + n\sigma_{es}^2$ (donde n representa el número de mediciones repetidas. A continuación se desarrolla el cálculo de los cuadrados medios utilizando las formulas planteadas en el capítulo de diseño de experimentos.

MUESTRA	RESIDUOS DE PESTICIDA(μg)	Totales	Promedios
1	120, 110	230	115
2	120, 100	220	110
3	140,130	270	135

En la tabla E61.2 se presenta el análisis de la varianza para los datos de residuos de pesticidas:

TABLA E61.2

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	F _o
Entre muestras	700	2	350	3.50
Dentro de muestras	300	3	100	
Total	1000	5		

Las estimaciones de las componentes de varianza son: $\sigma_{eM}^2 = 100$

$$\sigma_{es}^2 = \frac{350 - 100}{2} = 125$$

Por tanto, la estimación de la varianza de cualquier observación de la muestra es $\sigma_{eM}^2 + \sigma_{es}^2 = 100 + 125 = 225\mu\text{g}^2$, donde el 55.56% de la variabilidad se atribuye a las diferencias entre las muestras.

6.4.2 EJEMPLOS DE ANÁLISIS ESTADÍSTICO EN ENSAYOS DE COLABORACION

- En un ensayo de colaboración, se enviaron dos muestras muy similares de aceite (A y B) a 15 laboratorios, cada uno de los cuales realiza una determinación del nivel de cadmio en cada muestra utilizando la técnica de emisión de plasma por acoplamiento inductivo. Se obtuvieron los siguientes resultados (MILLER y MILLER, 1993, cap. 4):

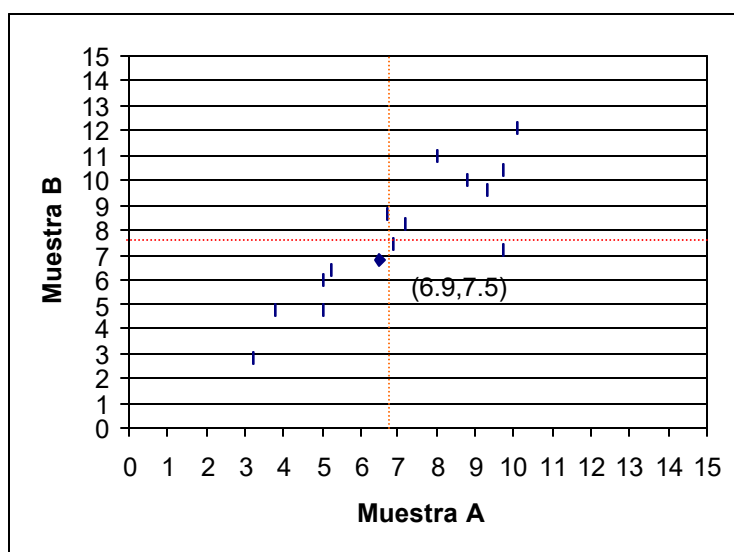
Laboratorio	Niveles de Cd (ppm) aceite A	Niveles de Cd (ppm) aceite B
1	8.8	10.0
2	3.8	4.7
3	10.1	12.1
4	8.0	11.0
5	5.0	4.7
6	5.2	6.4
7	6.7	8.7
8	9.3	9.6
9	6.9	7.5
10	3.2	2.8

Laboratorio	Niveles de Cd (ppm) aceite A	Niveles de Cd (ppm) aceite B
11	9.7	10.4
12	7.2	8.3
13	6.5	6.8
14	9.7	7.2
15	5.0	6.0

Dibuje el diagrama de Youden estas dos muestras para estos datos y comente la principal fuente de error en el ensayo de colaboración. Estime la varianza global, la varianza de medición, y el componente de error sistemático de la varianza de los resultados.

SOLUCION

a) Grafica de Youden: Se trazará un diagrama de dispersión de puntos en los que los datos de la muestra A serán la coordenada "x" y los resultados para la muestra B serán la coordenada en el eje "y", es decir cada punto es un par de resultados de un único laboratorio. Para la diagnosis de la gráfica se debe trazar un eje coordinado usando como origen el punto de las medianas para las dos muestras, es decir el punto (Mediana muestra A = 6.9; Mediana muestra B = 7.5). De esta forma, se tiene:



Los puntos que están muy alejados del grueso de los resultados indican laboratorios que probablemente, deberían ser cribados del análisis. Las dos líneas

medias dividen el espacio en cuatro cuadrantes y el primer (y, con frecuencia revelador) paso del análisis es observar la distribución de los puntos en los cuadrantes. Si sólo se han producido errores aleatorios de medición, habrá una distribución circular con, aproximadamente, el mismo número de puntos en cada cuadrante. Pero, en la mayor parte de los casos reales, los datos se concentran en el superior derecho y en el inferior izquierdo. Si un laboratorio da alto en ambas muestras, su punto caerá en el cuadrante superior derecho; si da bajo en las dos, caerá en el inferior izquierdo. Si da alto (o bajo) en ambas muestras, es una indicación de que el laboratorio, por alguna razón, pone su propio sello en el procedimiento, es decir, acusará siempre diferencias sistemáticas. Cuando existen esas, diferencias, los puntos tienden a caer a lo largo de una estrecha y larga elipse, tal y como se distribuyen los datos del ejemplo.

Asumiendo que los dos materiales son similares en tamaño y valor de la propiedad medida, como es requisito en esta prueba, y que la distribución de resultados de la muestra A ha de ser aproximadamente la misma que para la muestra B, se puede calcular una desviación tipo de cada resultado de la forma que se presenta en la tabla E62.1:

TABLA E62.1

Laboratorio	Niveles de Cd (ppm) aceite A	Niveles de Cd (ppm) aceite B	(1) $d = A - B$	(2) $d'_i = d_i - \bar{d} $
1	8.8	10.0	-1.2	0.46
2	3.8	4.7	-0.9	0.16
3	10.1	12.1	-2	1.26
4	8.0	11.0	-3	2.26
5	5.0	4.7	0.3	1.04
6	5.2	6.4	-1.2	0.46
7	6.7	8.7	-2	1.26
8	9.3	9.6	-0.3	0.44
9	6.9	7.5	-0.6	0.14
10	3.2	2.8	0.4	1.14
11	9.7	10.4	-0.7	0.04
12	7.2	8.3	-1.1	0.36
13	6.5	6.8	-0.3	0.44
14	9.7	7.2	2.5	3.24
15	5.0	6.0	-1	0.26
			$\bar{d} = -0.74$	$\bar{d}' = 0.864$

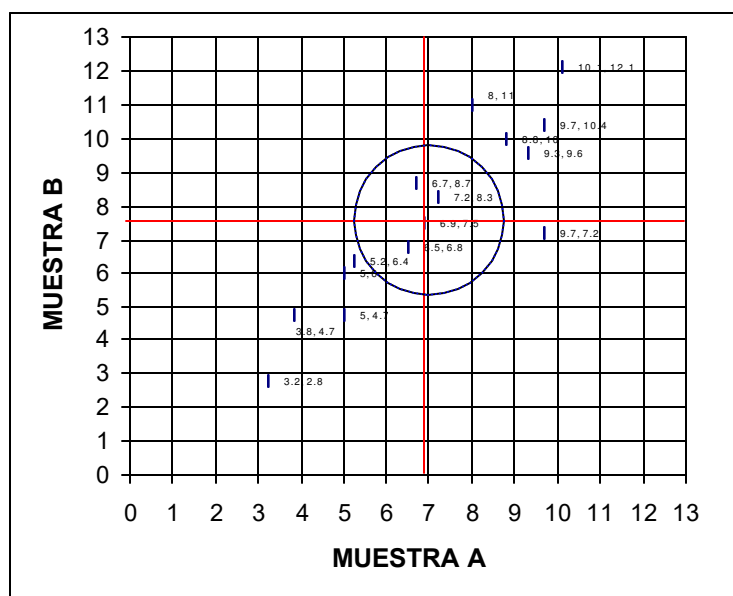
La estimación de la desviación estándar de las desviaciones se obtiene por:

$$s = \frac{\bar{d}'}{d_2} = 0.7660$$

donde d_2 , se obtiene de la tabla del anexo A7, para $n = 2$, es decir $d_2=1.128$. Ahora debe dibujarse un círculo que se espera contenga un porcentaje de puntos. El círculo se centrará en el punto mediano y su radio, para el porcentaje dado que ha de contener, se obtiene multiplicando la estimación de s por el factor de la tabla E62.2:

Porcentaje de puntos dentro del círculo	múltiplo de la desviación tipo
90	2.146
95	2.488
99	3.035

Así para un porcentaje de puntos dentro del círculo del 95% el radio es:
 $r = 2.488 \times 0.7660 = 1.9057$



Como puede observarse los laboratorios 7, 9, 12 y 13, generan resultados que se encuentran dentro del círculo, es decir resultados que parecen no tener un sesgo significativo. Mientras que los laboratorios 3 y 10 poseen para ambas muestras grandes sesgos positivos y negativos respectivamente. Las distancias en perpendicular desde los puntos experimentales a la línea de 45° que pasa por (6.9,

7.5) mide los errores aleatorios de los laboratorios participantes, y queda claro que los laboratorios 4 y 14 pueden tener grandes errores aleatorios en sus resultados. Los componentes de varianza pueden determinarse a partir de las ecuaciones siguientes:

$$\sigma_{EM}^2 = \frac{\sum (d_i - \bar{d})^2}{2(n-1)} \quad \sigma^2 = \frac{\sum (T_i - \bar{T})^2}{2(n-1)} \quad \sigma^2 = 2\sigma_1^2 + \sigma_{EM}^2$$

En donde: $d = A-B$, $T = A+B$, y \bar{d}, \bar{T} son los promedios de d y T respectivamente

σ_{EM} es la varianza de la medición

σ^2 es la varianza global, y

σ_1^2 es la componente de varianza debido a los errores sistemáticos

Los resultados de este ensayo se presentan en la tabla E62.3:

TABLA E62.1

Laboratorio	Niveles de Cd (ppm) aceite A	Niveles de Cd (ppm) aceite B	$(d_i - \bar{d})^2$	$(T_i - \bar{T})^2$
1	8.8	10.0	0.2116	16.3758
2	3.8	4.7	0.0256	39.1038
3	10.1	12.1	1.5876	55.4533
4	8.0	11.0	5.1076	18.03446
5	5.0	4.7	1.0816	25.5358
6	5.2	6.4	0.2116	9.9433
7	6.7	8.7	1.5876	0.4182
8	9.3	9.6	0.1936	17.1951
9	6.9	7.5	0.0196	0.1248
10	3.2	2.8	1.2996	76.62
11	9.7	10.4	0.0016	28.5872
12	7.2	8.3	0.1296	0.5576
13	6.5	6.8	0.1936	2.1121
14	9.7	7.2	10.4976	4.6083
15	5.0	6.0	0.0676	14.0873
			$\Sigma = 22.216$	$\Sigma = 308.75$

Los valores \bar{d}, \bar{T} son respectivamente -0.74 y 14.753 y de aquí que las estimaciones de σ_{EM} y σ_1^2 son las siguientes:

$$\sigma_{EM} = \frac{\sum (d_i - \bar{d})^2}{2(n-1)} = \frac{22.216}{2(15-1)} = 0.7934$$

Mientras que la estimación de σ^2 viene dada por:

$$\sigma^2 = \frac{\sum (T_i - \bar{T})^2}{2(n-1)} = \frac{308.75}{2(15-1)} = 11.027$$

Por lo que σ_1^2 es: $\sigma_1^2 = \frac{\sigma^2 - \sigma_o^2}{2} = \frac{11.027 - 0.7934}{2} = 5.117$

3. Se realiza un ejercicio de intercomparación para 14 laboratorios que analizan residuos de pesticidas, para lo cual se envió una muestra de Lindano en isoetano en tres niveles de concentración, esta muestra no ofrece dificultades de identificación y de matriz, por lo que el resultado de su análisis permite verificar el funcionamiento de los equipos de medición y la calidad de los materiales de referencia utilizados para la calibración. Los valores de concentración de lindano en las muestras se presentan en la tabla siguiente:

Muestra	identificación	Valores nominales(mg/l)
1	L ₁	0.015±0.001
2	L ₂	0.029±0.001
3	L ₃	0.060±0.002

Los datos enviados por los laboratorios participantes pueden verse en la tabla E63.1. (INTI, 1998).

SOLUCION

El análisis de los datos se hará en base a los valores de los estadísticos promedio y % de desviación ((valor promedio del laboratorio - valor nominal)/ valor nominal) de la serie de datos que se presentan en la tabla E63.1. Con el recurso a gráficos de control (gráficos 1 a 3) se pone de manifiesto el comportamiento de los laboratorios.

TABLA E63.1
Datos remitidos por los participantes
Análisis de Lindano

Participante	Lab. N ^o	L1				L2				L3			
		dato 1	dato 2	dato 3	promedio	dato 1	dato 2	dato 3	promedio	dato 1	dato 2	dato 3	promedio
1	501	0,014	0,015	0,015	0,015	0,029	0,028	0,029	0,029	0,058	0,058	0,058	0,058
2	502	0,011	0,011	0,011	0,011	0,028	0,029	0,028	0,028	0,061	0,067	0,062	0,063
3	503	0,016	0,016	0,016	0,016	0,032	0,032	0,033	0,032	0,053	0,063	0,062	0,063
4	504	0,013	0,013	0,013	0,013	0,027	0,026	0,026	0,026	0,055	0,053	0,055	0,054
4	505	0,013	0,013	0,013	0,013	0,027	0,028	0,027	0,028	0,057	0,057	0,057	0,057
5	506	0,013	0,013	0,013	0,013	0,027	0,027	0,027	0,027	0,053	0,053	0,053	0,053
6	507	0,018	0,016	0,017	0,017	0,028	0,027	0,028	0,028	0,041	0,043	0,042	0,042
7	508	0,015	0,015	0,014	0,015	0,029	0,028	0,028	0,028	0,060	0,062	0,059	0,060
8	509	0,016	0,016	0,016	0,016	0,033	0,034	0,035	0,034	0,073	0,070	0,070	0,071
9	510	0,013	0,012	0,012	0,012	0,026	0,028	0,027	0,027	0,040	0,044	0,043	0,042
9	511	0,014	0,013	0,013	0,013	0,028	0,027	0,028	0,028	0,043	0,042	0,041	0,042
10	512	0,014	0,014	0,014	0,014	0,028	0,028	0,028	0,028	0,059	0,060	0,059	0,059
11	513	0,015	0,015	0,015	0,015	0,032	0,032	0,032	0,032	0,060	0,059	0,062	0,060
12	514	0,017	0,017	0,016	0,017	0,039	0,038	0,042	0,040	0,080	0,077	0,085	0,080
13	515	0,014	0,014	0,014	0,014	0,030	0,031	0,030	0,030	0,059	0,059	0,058	0,059
13	516	0,012	0,012	0,015	0,013	0,030	0,032	0,037	0,033	0,059	0,062	0,059	0,060
14	517	0,015	0,015	0,015	0,015	0,027	0,027	0,029	0,028	0,056	0,051	0,057	0,055
14	518	0,014	0,013	0,012	0,013	0,028	0,027	0,028	0,028	0,055	0,054	0,053	0,054
14	519	0,013	0,014	0,013	0,013	0,027	0,027	0,027	0,027	0,053	0,053	0,053	0,053
14	520	0,011	0,010	0,011	0,011	0,024	0,025	0,024	0,024	0,051	0,048	0,050	0,050

Nota 1: Se aclara que algunos participantes enviaron más de una serie de datos.

Nota 2: Los datos que se informaron con un número mayor a tres decimales fueron redondeados en ese número de cifras.

TABLA E63.2

Promedios y porcentajes de desviación con respecto a los valores especificados para las muestras de lindano analizadas

Participante	Lab.No.	L1		L2		L3	
		promedio	%desv.	promedio	%desv.	promedio	%desv.
1	501	0,015	0	0,029	0	0,058	-3.33
2	502	0,011	-26.67	0,028	-3.45	0,063	5.0
3	503	0,016	6.67	0,032	10.34	0,063	5.0
4	504	0,013	-13.33	0,026	-10.34	0,054	-10.0
4	505	0,013	-13.33	0,028	-3.45	0,057	-5.0
5	506	0,013	-13.33	0,027	-6.90	0,053	-11.7
6	507	0,017	13.33	0,028	-3.45	0,042	-30.0
7	508	0,015	0.00	0,028	-3.45	0,06	0.0
8	509	0,016	6.67	0,034	17.24	0,071	18.3
9	510	0,012	-20.00	0,027	-6.90	0,042	-30.0
9	511	0,013	-13.33	0,028	-3.45	0,042	-30.0
10	512	0,014	-6.67	0,028	-3.45	0,059	-1.7
11	513	0,015	0.00	0,032	10.34	0,06	0.0
12	514	0,017	13.33	0,04	37.93	0,08	33.3
13	515	0,014	-6.67	0,03	3.45	0,059	-1.7
13	516	0,014	-6.67	0,028	-3.45	0,06	0.0
14	517	0,015	0.00	0,028	-3.45	0,055	-8.3
14	518	0,013	-13.33	0,028	-3.45	0,054	-10.0
14	519	0,013	-13.33	0,027	-6.90	0,053	-11.7
14	520	0,011	-26.67	0,024	-17.24	0,05	-16.7

Gráfico 1
 Datos enviados por los participantes- L1

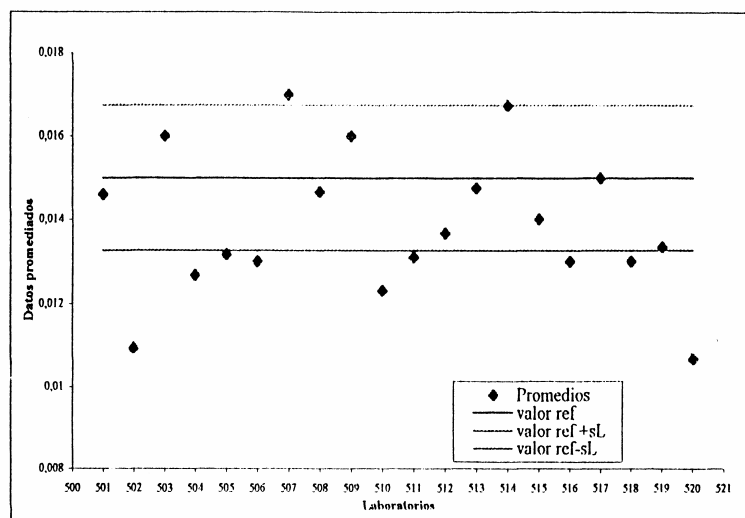


Gráfico 2
 Datos enviados por los participantes- L2

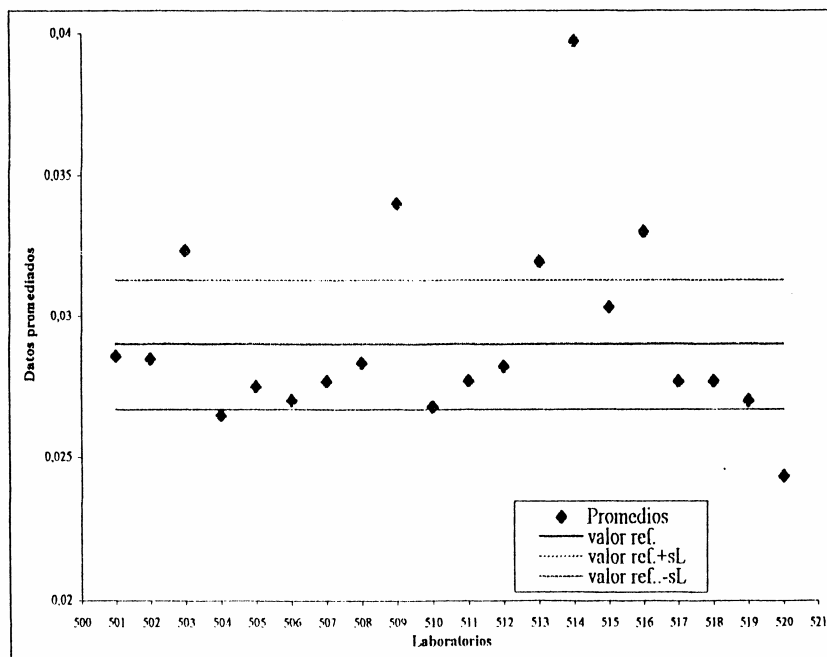
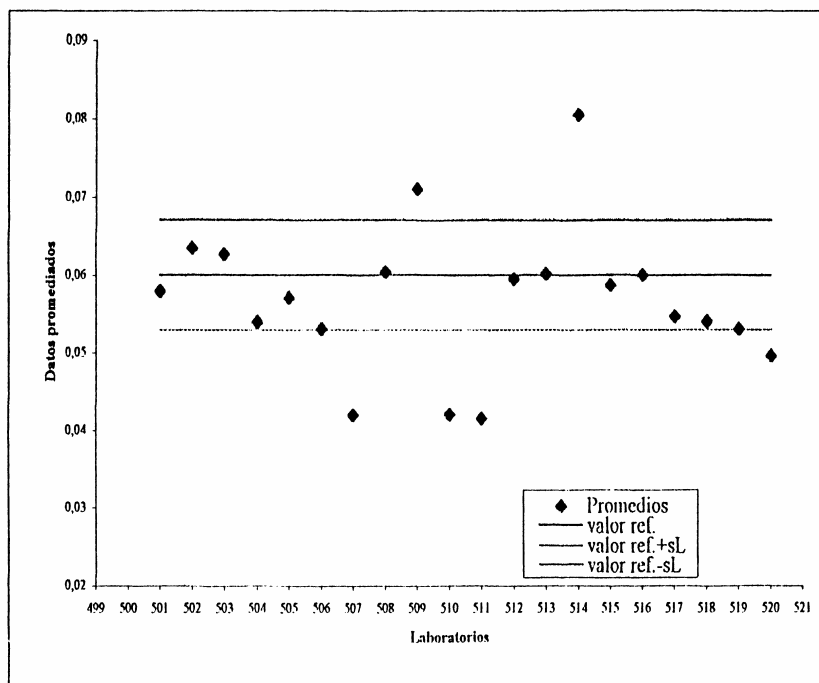


Gráfico 3
Datos enviados por los participantes- L3



Para cada juego de datos se calcula la media y la desviación estándar de la muestra.

A fin de evaluar el desempeño del conjunto de laboratorios, se usó como criterio el cálculo del parámetro "z" definido de la siguiente manera:

$$z = \left| \frac{(x_i - A)}{s} \right|$$

Donde: A es el valor "verdadero" de la concentración de analito; s es la desviación estándar seleccionada, y x_i es el valor medido de la concentración de analito, representado por su promedio (o mediana).

Los organizadores del ensayo decidirán sobre los parámetros A y s de la ecuación anterior, (ver anexo B6).

Usando el criterio z , es posible clasificar el desempeño de los laboratorios de la siguiente forma:

$$|z| \leq 2 \text{ satisfactorio} \quad 2 < |z| < 3 \text{ cuestionable} \quad |z| \geq 3 \text{ no satisfactorio}$$

Sin embargo, antes es conveniente seleccionar los "outliers", lo cual puede hacerse de diversas formas. Primero, en este caso se estima que, por las características del método de análisis, puede preverse hasta un 20% de discrepancia en los resultados. Para poder aplicar las pruebas estadísticas se descartaron los resultados que discrepan en más de un 30% del valor de referencia. Los datos eliminados de esta forma se denominan de aquí en adelante "inconsistentes", de acuerdo a la tabla E6.3 se consideran inconsistentes los datos para la muestra L_2 del laboratorio 514 y los datos para la muestra L_3 de los laboratorios con códigos 507, 510, 511 y 514. Posteriormente se hace otra selección de datos (con los datos inconsistentes eliminados) calculando el valor medio y la desviación estándar para cada concentración, los resultados que caen fuera del intervalo $\bar{X} \pm 3s$ deben ser ignorados a la hora de calcular el promedio y la desviación estándar que se servirán en el cálculo del z-score. De acuerdo a lo anterior y en función de los datos de la tabla E6.2 y al cálculo de los intervalos $\bar{X} \pm 3s$ para las muestras, ya no hay más datos anómalos.

En la tabla E6.3, se presenta el cálculo del parámetro z obtenido para cada muestra. Como puede observarse para el caso de la muestra L_1 , los laboratorios con código 502 y 520 resultan cuestionables. Mientras que para la muestra L_2 el laboratorio con código 514 resulta no satisfactorio. Para la muestra L_3 , el laboratorio con código 509 resulta cuestionable, mientras que los laboratorios con códigos 507, 510, 511 y 514 son no satisfactorios.

En el archivo, [z-score, cap6.xls](#) se encuentra la hoja de cálculo para el cálculo de los estadísticos y los z-scores para las diferentes muestras.



z-score, cap6.xls

TABLA E6.3

Laboratorio	z-scores L_1	z-scores L_2	z-scores L_3
501	0	0	0.392
502	2.27	0.401	0.588
503	0.57	1.202	0.588
504	1.14	1.202	1.176
505	1.14	0.401	0.588
506	1.14	0.801	1.373
507	1.14	0.401	3.529

TABLA E63.3 (continuación)

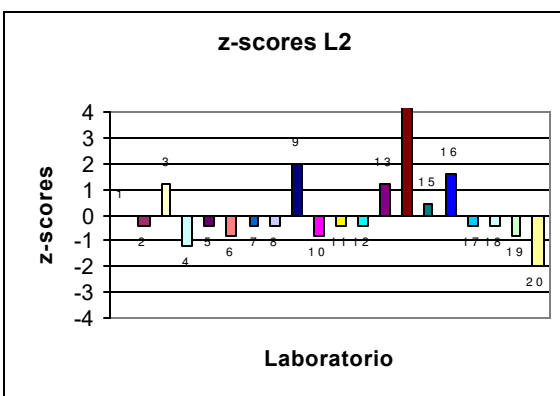
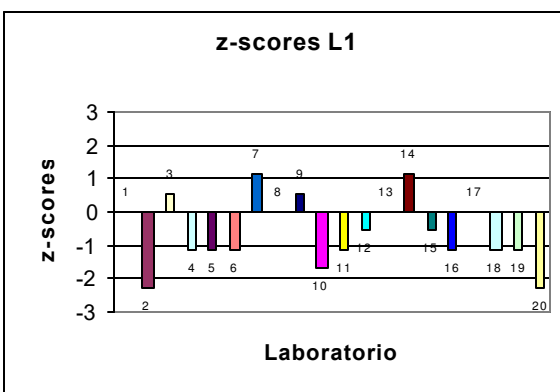
Laboratorio	z-scores L_1	z-scores L_2	z-scores L_3
508	0.00	0.401	0.000
509	0.57	2.003	2.157
510	1.70	0.801	3.529
511	1.14	0.401	3.529
512	0.57	0.401	0.196
513	0.00	1.202	0.000
514	1.14	4.407	3.922
515	0.57	0.401	0.196
516	1.14	1.603	0.000
517	0.00	0.401	0.980
518	1.14	0.401	1.176
519	1.14	0.801	1.373
520	2.27	2.003	1.961

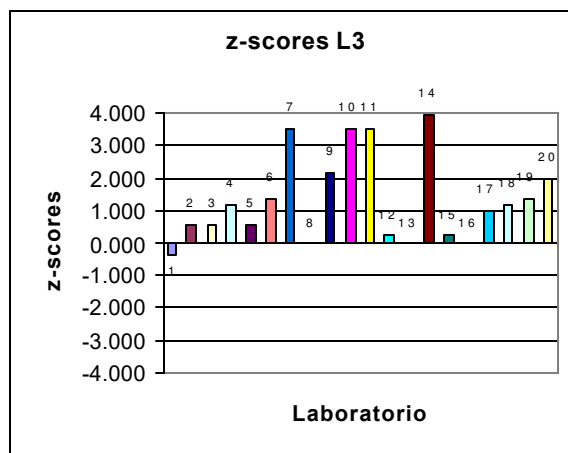
Los valores de z fueron calculados para cada muestra, usando los siguientes valores de A y s

$$A_{L_1} = 0.015 \quad A_{L_2} = 0.029 \quad A_{L_3} = 0.060$$

$$S_{L_1} = 0.0017 \quad S_{L_2} = 0.0025 \quad S_{L_3} = 0.0051$$

En las figuras siguientes se presentan los diagramas de barras para los z-score de los diferentes laboratorios.





REFERENCIAS BIBLIOGRÁFICAS

- **Garfield, F.M. (1985).** Principios de garantía de calidad para laboratorios analíticos. Segunda edición. España: AOAC.
- **García Alvarez, M.C. (1998).** Análisis, Homologación y Control de Calidad, Maestría en química analítica. Nicaragua: Universidad Nacional Autónoma de Nicaragua, Universidad de Valencia.
- **Instituto Nacional de Tecnología Industrial (1998).** Resultados del ensayo interlaboratorio de residuos de pesticidas. Buenos Aires, Argentina.
- **Juran, J.M y Gryna, F. M. (1997).** Manual de control de calidad. (2 tomos). España: Mcgraw-Hill.
- **Kenett, R. S. y Zacks, S. (2000).** Estadística industrial moderna. Diseño y control de la calidad y la confiabilidad. México: International Thomson Editores.
- **Miller, J. C. y Miller, J. N. (1993).** Estadística para química analítica. Segunda edición. Wilmington, Delaware, E.U.A: Addison-Wesley Iberoamericana, S.A.
- **Prichard, E. et al. (1995).** Quality in the Analytical Chemistry Laboratory. New York, U.S.A.: John Wiley & Sons Ltd.

ANEXO A6. GENERALIDADES SOBRE EL MUESTREO DE AIRE Y DE AGUA EN EL AREA AMBIENTAL (GARCIA, 1998, cap. 5).

a) *Muestreo de aire ambiental*

El muestreo de aire ambiental se define como la recolección de muestras de aire, en cualquier localización expuesta a la atmósfera. En esta amplia clasificación, se incluye un grupo infinito de posibilidades, desde la toma de muestra en un parking a una altura de 0.50 m del suelo, hasta la que recoge un globo sonda libre en un océano, por lo que no se pueden definir una reglas de calidad en el muestreo que cubran todas las áreas, pero sí se pueden indicar aquellos aspectos a considerar de manera general, previamente a un muestreo ambiental.

El primer paso es seleccionar, claramente, los objetivos del estudio. Estos pueden ser:

- i) Determinar la calidad del aire, en relación con sus efectos sociales y ambientales.
- ii) Establecer cuáles son las fuentes de emisión que inciden en la calidad del aire.
- iii) Serenar información para ayudar a planificar el control global de la contaminación en un área, y las estrategias para localizar las zonas municipales e industriales.
- iv) Realizar programas de investigación de los mecanismos de reacción de los contaminantes en el aire y de la dispersión.

La meteorología juega un papel fundamental en la determinación de la calidad del aire. Las variaciones diurnas y estacionales en las fuentes de emisión se reflejan en los niveles de contaminación ambiental, y este efecto se incrementa con las variaciones en las condiciones meteorológicas. De todas ellas, las más significativas son la estabilidad atmosférica, velocidad y dirección del viento, precipitación, temperatura, humedad y radiación solar.

En cuanto a la elección de los puntos de muestreo, no existen reglas generales a modo de procedimiento estándar, debido a la gran diversidad que plantea el muestreo ambiental. Conviene señalar algunas reglas de carácter general:

- i) Normalmente, la toma de muestra debe hacerse al menos a 1.5 a 2.0 metros del suelo. La máxima altura viene determinada por el objetivo concreto del estudio, pero debe ser consistente entre los distintos puntos, en el caso de una red de muestreo diseñada para un objetivo común.
- ii) El punto más alto de los grandes obstáculos, tales como edificios, montañas, etc., debe formar un ángulo con el muestreador de al menos 30°.
- iii) Debe evitarse exponer el muestreador a fuentes localizadas específicas, por ejemplo, una chimenea si el muestreo se hace en un tejado. Para establecer la calidad del aire en un área determinada, el punto de muestreo no debe situarse directamente en las principales fuentes de emisión (autopistas, parkings, chimeneas industriales, etc.).
- iv) El punto debe ser accesible y seguro.
- v) Dispondrá de toma eléctrica, siempre que los equipos lo requieran.

En el caso de redes de muestreo, la elección de los puntos se hace utilizando métodos estadísticos basados en modelos de difusión, así como aprovechando los datos meteorológicos históricos. Por ejemplo, se han utilizado en la práctica los siguientes criterios:

- i) Situar los puntos en círculos concéntricos, respecto al área de interés.
- ii) Seguir la trayectoria típica de los vientos.
- iii) Situar de manera aleatoria un número grande de puntos de muestreo, en las proximidades de la zona más interesante, y un número más reducido de puntos en el exterior, elegidos también de manera aleatoria.
- iv) Dividir la zona de manera regular (por ejemplo, mediante una cuadrícula), situando los puntos de manera equidistante.

La frecuencia y duración de los muestreos dependerá del objetivo, contaminantes a estudiar y método de análisis. Por ejemplo, un estudio del plomo es significativo haciendo un promedio semi-anual, ya que su efecto es acumulativo, mientras que la determinación de dióxido de azufre debe efectuarse con bastante frecuencia, ya

que se pueden dar episodios puntuales de contaminación elevada, que es necesario controlar.

Se pueden distinguir dos grupos diferentes de muestreo ambiental, según se recojan partículas o gases, ya que por su diferencia naturaleza, requerirán el control de parámetros bien diferenciados.

b) Muestreo de partículas

La recolección de partículas, de manera que la muestra no sea muy diferente de la población de interés, es un tema muy difícil y que requiere un tratamiento especial. El problema se plantea por el momento de inercia de las partículas, que es el producto de su masa por su velocidad.

Cuando cambia la dirección de una corriente en la que hay partículas y gases, éstas al ser sus moléculas mucho más grandes que las de los gases, tienden a continuar en su línea original y desplazarse de la misma; es decir, se producen deposiciones de las partículas más grandes. Además, las partículas pueden depositarse en los tubos o en otros componentes del muestreador, o perderse por atracción electrostática hacia las superficies del captador, especialmente en aquéllas no conductoras, como vidrio o plástico. También se pueden perder partículas por adhesión en superficies húmedas o con grasa, en diferentes proporciones, según la humedad relativa y la rugosidad de las superficies.

Las partículas pueden cambiar de tamaño en el muestreo, modificándose con ello la eficiencia de recolección. Por ejemplo, las moléculas de ácido sulfúrico adsorben agua, cuando la humedad relativa es alta (por encima del 50%). Las partículas más pequeñas tienden a adherirse a las mayores, variando el tamaño de las mismas.

En un muestreo de partículas, se deben considerar las siguientes posibilidades:

- i) Reacciones entre contaminantes posteriores a la recolección
- ii) Reacciones con los filtros, disolventes, etc.
- iii) Algunos gases pueden adsorberse en los aerosoles sólidos o líquidos. Los aerosoles pueden sublimar o evaporarse, después de la toma de muestra.

Sin embargo, si se sigue un control estricto del muestreo, estos últimos aspectos no suelen causar problemas serios, y en general, se puede decir que las partículas tienden a mantener su integridad, durante grandes periodos de tiempo.

Las fuentes puntuales no suelen distribuir las partículas en todas las direcciones. Estas se mueven con el viento, característica muy variable. Un muestreador situado en el camino de un penacho puede captar entre 10 y 100 veces más partículas que uno situado a sólo unos pocos metros del mismo. En las proximidades de edificios, árboles y otros obstáculos, el aire permanece inactivo y se obtiene menor concentración de partículas.

También se dan gradientes de concentración importantes en sentido vertical. Ya que, normalmente las fuentes están cerca del suelo, la cantidad de partículas disminuye con la altura.

La hora de muestreo es también una variable importante. En muchas ciudades, la altura de la capa de mezcla varía a lo largo del día. A menudo se produce una inversión nocturna, que sitúa los contaminantes cerca del suelo, aumentando relativamente las concentraciones. Cuando amanece, la inversión sube gradualmente y se rompe en la mañana o a primeras horas de la tarde, momento en el que se produce buena dispersión de los contaminantes.

La concentración de partículas puede variar, fácilmente, en un factor de 10 o más en un periodo de 24 horas. Así, una muestra tomada durante la tarde puede dar muy poca información sobre la contaminación real de una zona. Por otra parte, una muestra de 24 horas puede promediar un episodio de contaminación severa ocurrido durante unas pocas horas. En cualquier caso, los muestreos de partículas de 24 horas son una práctica estándar.

En resumen, la muestra debe tomarse en el punto de mayor interés, siempre que esto sea posible, pero adoptando todas las precauciones debidas a la hora de extrapolar datos a un área más amplia.

Si se desea obtener una medida del aerosol, en cuanto a su influencia sobre la salud de los seres vivos en un área, debe tenerse en cuenta:

- i) No estar directamente a favor del viento de la fuente emisora.
- ii) Localizar el muestreador aproximadamente a 1.5 m del suelo.
- iii) Situarse lejos de los principales obstáculos, a una distancia de al menos 10 veces su altura.
- iv) Tomar varias muestras de la zona de interés en diferentes puntos.
- v) Considerar las horas del día más significativas.

c) Muestreo y almacenamiento de gases

Los gases siguen las leyes normales de la difusión y se mezclan perfectamente con la atmósfera. No están afectados por las fuerzas de inercia o electrostáticas que presentan las partículas, por lo que son más sencillos de muestrear que estas. Sin embargo, es mucho más difícil mantener la integridad de la muestra.

Las precauciones más importantes a considerar, para mantener la calidad de un muestreo de gases, se indican a continuación:

- i) La integridad de la muestra. Las degradaciones y reacciones de los gases, dentro del sistema recolector, son fáciles.
- ii) El procedimiento de muestreo debe ser el óptimo para el procedimiento analítico posterior.
- iii) Debe considerarse la posibilidad de reacción con las paredes del contenedor.
- iv) Previsión de fugas y contaminaciones de la muestra. Esto se controla previamente al muestreo, siguiendo los programas adecuados de control de calidad. En general se debe reducir al máximo el tiempo transcurrido entre muestreo y análisis, proteger las muestras de la radiación solar y de las temperaturas extremas, durante el transporte y almacenamiento.
- v) Existe una gran variedad de sistemas de bombeo de aire, con diferentes tamaños y tipos. Cuando la muestra debe atravesarlos, es necesario que los mismos estén contruidos en materiales inertes, como teflón o acero inoxidable. Este tipo de bombas son esenciales, cuando se muestrean gases como ozono, o tazas de hidrocarburos o halocarburos.
- vi) Mención especial merecen los tubos utilizados para el muestreo. Es esencial que la muestra de aire llegue inalterada al muestreador. Para asegurar que no hay pérdidas de gases o vapores, el tubo debe estar construido de un material inerte, como teflón o vidrio. No son recomendables materiales como la silicona o el nylon, incluso en el muestro de gases inertes. Los tubos deben tener la menor longitud posible, y nunca dejar que acumulen polvo o humedad.
- vii) Cuando se utilizan absorbentes, se deben sellar y guardar de manera adecuada, inmediatamente después del muestreo.

c) Muestreo de aguas

El objetivo del muestreo de aguas es recoger una pequeña porción de material, es decir de agua, de manera que tenga un volumen adecuado para el transporte y posterior

manipulación en el laboratorio de análisis, y que siga siendo una representación exacta del cuerpo de agua que se pretende analizar. Esto significa que las proporciones o concentraciones relativas de todos los componentes de interés deben permanecer inalteradas, antes de efectuar los ensayos.

Algunas precauciones generales en el muestreo de agua son las siguientes:

- i) debe obtenerse una muestra que se ajuste a los requisitos del programa de muestro y manipularla, de manera que no se deteriore o contamine, antes de llegar al laboratorio. El envase o botella utilizados deben enjuagarse dos o tres veces con el agua a muestrear, a no ser que contengan algún conservante.
- ii) Dependiendo del tipo de análisis, el envase debe llenarse totalmente (mayor parte de contaminantes orgánicos), o dejar espacio para la aeración, mezcla, etc. (análisis microbiológicos). Para muestras que van a ser almacenadas, es preferible dejar un espacio de aire de aproximadamente el 1% de la capacidad del recipiente, para permitir la expansión térmica.
- iii) En determinados casos, las muestras representativas solo pueden obtenerse preparando mezcla compuestas de muestras elementales, recogidas en periodos de tiempo o en diferentes puntos de muestreo. Los detalles del muestreo variarán enormemente con las condiciones locales, de manera que no se pueden establecer recomendaciones específicas generales. A veces, proporciona más información el análisis de varias muestras elementales, que el de una muestra global.
- iv) Factores importantes que afectan a los resultados son la presencia de materia en suspensión o turbidez, el método elegido para eliminarlos, así como los cambios fisicoquímicos, debidos al almacenamiento o aireación de las muestras.
- v) Se deben tener precauciones especiales al procesar (mezclas, filtrar, etc) muestras en las que se van a determinar trazas de componentes, especialmente metales. Por ejemplo, la determinación de plomo puede invalidarse, con facilidad, por contaminación durante estos procesos.
- vi) Se debe tratar cada muestra de acuerdo con las sustancias que se van a determinar, la cantidad y la naturaleza de la turbidez, y otras condiciones que pueden influir en los resultados.

- vii) En general, se deben separar la materia en suspensión por decantación, centrifugación, o mediante un procedimiento de filtración adecuado. A veces, puede tolerarse una turbidez ligera, ya que puede corregirse en una prueba calorimétrica, que es donde potencialmente tiene el mayor efecto interferente.
- viii) Debe anotarse cada muestra e identificarla posteriormente, el nombre del muestreador, fecha, hora, localización exacta, temperatura del agua y otros datos necesarios, como las condiciones climáticas, nivel del agua, caudal, etc. Debe dejarse un espacio para las iniciales de la persona que asuma la muestra, y la hora y fecha de transporte. Se deben fijar los puntos de muestreo mediante mapas, para permitir su identificación por otras personas.
- ix) Antes de recoger muestras procedentes de sistemas de distribución, deben dejarse pasar una cantidad de agua suficiente para asegurar que se obtiene una muestra representativa, teniendo en cuenta el diámetro y la longitud de la tubería, así como el caudal.
- x) Al recoger muestras de pozos, debe asegurarse que se ha bombeado suficientemente para que la muestra represente al agua subterránea de interés.
- xi) Cuando las muestras son de ríos o corrientes, los valores analíticos pueden variar con la profundidad, el caudal, la distancia a la orilla y de una orilla a la otra. Si se dispone del equipo adecuado, es conveniente tomar una muestra integrada desde la superficie al fondo, en mitad de la corriente. Si solo se puede tomar una muestra, debe hacerse en el centro de la corriente y a una profundidad media.
- xii) Los lagos están sujetos a grandes variaciones por causas naturales, tales como la estratificación estacional, las lluvias, la deposición y los vientos. Debe elegirse el punto, profundidad y frecuencia de muestreo, de acuerdo con las condiciones locales y el objetivo de la investigación. Se debe evitar recoger a espuma superficial.
- xiii) En general, debe tenerse siempre en cuenta antes de diseñar el plan de muestreo, los análisis que se van a efectuar, así como el objetivo del muestreo.

Tipos de muestras de aguas:

- i) Muestras instantáneas: Estrictamente hablando, una muestra que se recoge en un instante y en un punto determinado, sólo puede representar a la fuente en

ese momento y en ese punto. Sin embargo, cuando se sabe que una fuente es bastante constante en cuanto a su composición, durante largos periodos de tiempo, o en un amplio margen de distancias y direcciones, se puede considerar como una muestra representativa. Este es el caso de algunos sistemas de distribución de agua, determinadas aguas superficiales, y ocasionalmente también, de corrientes de aguas residuales. Cuando se sabe que la composición de la fuente varía con el tiempo, estas muestras, recogidas a intervalos adecuados y analizadas por separado, pueden documentar la extensión, frecuencia, y duración de las variaciones. Se deben elegir los intervalos, de acuerdo con la frecuencia con la que se esperan los cambios. Si la composición varía en el espacio, se deben tomar muestras en los puntos adecuados, siguiendo diversos planes estadísticos, o basándose en la experiencia anterior.

- ii) Muestras compuestas: En la mayor parte de los casos, este término se refiere a una mezcla de muestras instantáneas recogidas en un mismo punto, a diferentes tiempos. Este tipo de muestra es útil para observar promedios de concentraciones, que se utilizan por ejemplo, para calcular la eficiencia de una planta de tratamiento de aguas residuales. Como alternativa al análisis separado de una gran número de muestras, seguido por tratamiento de datos de promedios y resultados totales, estas muestras suponen un ahorro tanto económico, como de esfuerzo. Por ello, se considera de manera estándar, para la mayor parte de las determinaciones, una muestra compuesta representando un periodo de 24 horas, a no ser que se elijan otros ciclos, debido a problemas particulares.

Cuando se determinen componentes o características variables con el almacenamiento, no se puede utilizar este tipo de muestreo. Tales determinaciones deben realizarse rápidamente sobre las muestras instantáneas, preferiblemente en el mismo punto de muestreo. Ejemplos de este tipo de análisis son los gases disueltos, la temperatura y el pH. Cuando se producen cambios en algunos de estos componentes, por ejemplo el pH o la temperatura, pueden a su vez ocasionar variaciones en otros constituyentes, como hierro, manganeso, alcalinidad o dureza.

Se deben tomar porciones individuales, recogidas a intervalos regulares, utilizando una botella con un diámetro de entrada de al menos 35 mm y una capacidad mínima de 120 ml. La mezcla puede efectuarse, tanto al final como durante el muestreo. Si se utilizan conservantes, deben añadirse al inicio del muestreo, de manera que todas las porciones estén en contacto con ellos desde el momento en que son muestreadas. El volumen debe ser proporcional al caudal. El volumen final de 2 ó 3 litros es suficiente, en el caso de afluentes y residuos. Las botellas se deben limpiar a diario, así como todos los objetos de muestreo, para eliminar el crecimiento biológico, así como otros depósitos.

iii) Muestras integradas: Para determinados objetivos, puede ser necesario analizar mezclas de muestras instantáneas recogidas simultáneamente de diferentes puntos. Un ejemplo es el caso de un río, en el que la composición varía a través de su profundidad y anchura. Para establecer una composición promedio, deben tomarse varias muestras en puntos de la sección transversal, en proporción a sus flujos relativos. La preparación de muestras integradas normalmente requiere equipos especiales para recoger una muestra a una profundidad conocida, sin contaminación del agua de las capas superiores. Se debe conocer el volumen, movimientos, y composición de las capas del agua que se va a muestrear.

Cantidad de muestra.

De manera general, se establece la cantidad de 2 litros, para la mayor parte de las determinaciones físicas y químicas. No obstante, para determinados análisis, se pueden necesitar mayores cantidades. Para la mayor parte de los análisis físicos y químicos es suficiente una muestra de 2 litros, pero para ciertas determinaciones especiales, se puede necesitar un mayor volumen.

No debe utilizarse la misma muestra para los estudios químicos, bacteriológicos y exámenes microscópicos, ya que los métodos de muestreo y manipulación de la muestra son diferentes.

Conservación de la muestra

La conservación completa e inequívoca de una muestra de agua es prácticamente imposible. Las técnicas de conservación, en el mejor de los casos, retardan los cambios biológicos y químicos, que inevitablemente tienen lugar después del muestreo.

Algunos constituyentes son más sensibles a modificaciones. Por ejemplo, puede haber pérdidas de algunos cationes por absorción o intercambio iónico, con las paredes del recipiente que los contiene (aluminio, cadmio, cromo, cobre, hierro, plomo, magnesio, plata, zinc). Se debe recoger, separadamente, una muestra en una botella limpia y acidificarla con ácido nítrico, a un pH inferior a 2, para minimizar precipitaciones y adsorciones.

La temperatura cambia rápidamente, y el pH puede variar de manera significativa en cuestión de minutos. Pueden darse pérdidas de gases disueltos (e.g., oxígeno, dióxido de carbono), por lo que estos parámetros deben determinarse en campo. Con cambios en el balance Ph- alcalinidad- dióxido de carbono, pueden precipitar el carbonato cálcico disminuyendo la dureza total.

Hierro y magnesio son solubles en sus estados de oxidación más bajos, pero relativamente insolubles en los superiores. Así, pueden precipitar o disolverse a partir de un sedimento, dependiendo del potencial redox de la muestra. Se pueden producir variaciones en el contenido en nitratos, nitritos y amoníaco, disminución de la concentración de fenol y en la demanda biológica de oxígeno; todo ello por la actividad microbiológica, etc. Las características organolépticas pueden aumentar, disminuir o modificarse sus cualidades. No se pueden utilizar recipientes de vidrio para muestrear sílice, boro o sodio. Algunas transformaciones biológicas pueden cambiar los estados de oxidación de determinación de metales.

Es imposible precisar terminantemente el tiempo que puede transcurrir entre la recolección de una muestra y su análisis, pues depende de la índole de la muestra, del análisis particular por verificar y de las condiciones de almacenamiento. Se pueden reducir mucho los cambios producidos por la proliferación de organismos si la muestra se mantiene en la oscuridad y a baja temperatura hasta que se analice. Se sugieren, como razonables, los siguientes límites máximos para muestras destinadas a análisis químicos y físicos:

- Aguas no contaminadas..... 72 horas
- Aguas ligeramente contaminadas..... 48 horas
- Aguas contaminadas 12 horas

En general, algunos de estos aspectos se recogen en los métodos de análisis individuales, pero será fundamental el criterio del analista, en relación con la naturaleza del agua en estudio.

ANEXO B6. ESTIMACION DE LOS VALORES DE A Y s PARA EL CALCULO DEL z-score (PRICHARD et al., 1995, cap. 3).

Los organizadores del ensayo de colaboración son los encargados de decidir sobre los parámetros A y s para el cálculo del z-score.

Estimación del verdadero valor, A

Hay tres formas de obtener una estimación de A .

(i) La adición de una cantidad conocida de analito a una matriz que no contiene analito. Este método es completamente satisfactorio en muchos casos, especialmente si lo que se requiere es que la cantidad total de analito sea mayor que la concentración. Si la concentración requerida es del orden de los ppm entonces la cantidad exacta de la matriz a la cual es añadido el analito debe conocerse tan bien como la cantidad añadida. Si sólo una porción de esta será usada para su análisis, es importante conocer la homogeneidad de la distribución del mismo en la muestra.

(ii) El uso de un valor consensuado producido por un grupo de laboratorios expertos usando los métodos mejores posibles. Esta es probablemente, la aproximación más cercana al verdadero valor para el material de prueba, pero "producirlo" es bastante caro.

(iii) El uso de un valor consensuado, producido en cada corrida de un ensayo de colaboración y basándose sobre los resultados obtenidos por los participantes. El dato consensuado es estimado como la media de los resultados después de que se han eliminado los datos anómalos. Esta aproximación es claramente la forma más barata de obtener una estimación del verdadero valor, pero podrían haber problemas sino existe un consenso real entre los participantes o bien el consenso está sesgado debido al uso de una metodología general defectuosa.

Estimación de la desviación estándar, s

Hay cuatro formas de estimar los valores de la desviación estándar, estas son:

(i) Uso de la desviación estándar de un blanco, el cual es representativo de la concentración de analito. Usar esta aproximación tiene la ventaja de que los resultados de diferentes ensayos pueden ser comparados.

(ii) Calcular la desviación estándar de los resultados obtenidos por todos los laboratorios participantes en el ensayo.

(iii) Excluir los datos anómalos de los resultados obtenidos y calcular la desviación estándar.

(iv) Si es posible, use la desviación estándar obtenida en ensayos de colaboración donde todos los participantes en el ensayo usaron el mismo método.

CAPITULO VII

CONTROL DE CALIDAD EN LABORATORIOS DE ANALISIS QUIMICO

CONTENIDO

	Pág.
7.1 Control de la calidad	7-2
7.2 División del control de calidad	7-6
7.3 Control estadístico de la calidad	7-8
7.4 Estudio de la variación	7-11
7.5 Diagramas de control	7-12
7.6 Cartas de control de SHEWHART	7-15
7.7 Diagramas de suma acumulativa (ACUSUMA)	7-20
7.8 Diagramas de medias móviles	7-22
7.9 Diagramas de control para observaciones dependientes	7-22
7.10 Ejercicios de construcción y análisis de gráficas de control	7-23
REFERENCIAS BIBLIOGRAFICAS	7-41
ANEXO A7. Factores para construir diagramas de control para variables	7-43
ANEXO B7. Plantilla para la elaboración de gráficas de control	7-44
ANEXO C7. Ilustración de prueba de curvas asignables	7-45
ANEXO D7. Procedimiento de construcción de diagramas ACUSUMA	7-48
ANEXO E7. Procedimiento de construcción de diagramas de control bivariantes	7-50

7.1 CONTROL PARA LA CALIDAD

“Control” se refiere al proceso que se emplea con el fin de cumplir con los estándares. Esto consiste en observar el desempeño real, compararlo con algún estándar y después tomar medidas si el desempeño observado es significativamente diferente del estándar. El proceso de control tiene la naturaleza de un ciclo de retroalimentación. (ver figura 7.1). El control incluye la siguiente secuencia universal de pasos (JURAN, 1990, cap. 9):

1. Seleccionar al sujeto de control: esto es lo que se quiere regular.
2. Elegir la unidad de medida
3. Establecer la meta para el sujeto de control
4. Crear un censor que pueda medir el sujeto de control en términos de la unidad de medida
5. Medir el desempeño real
6. Interpretar la diferencia entre el desempeño real y la meta
7. Tomar medidas (si es necesario) sobre la diferencia

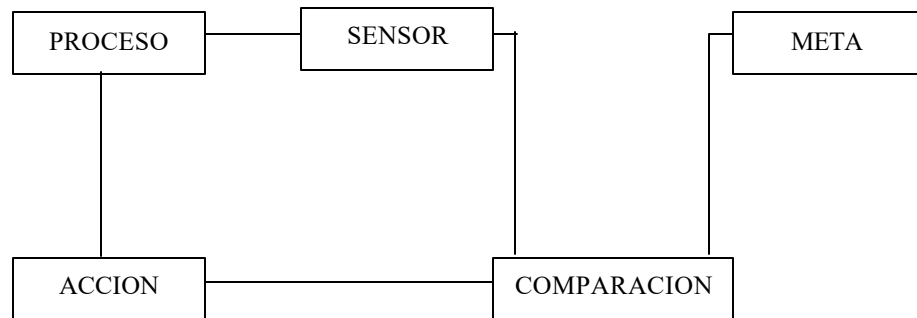


FIGURA 7.2 PROCESO DE CONTROL

Esencial en el proceso de control de calidad es el hecho de “medir la calidad”. Para el control de la calidad la medición proporciona retroalimentación y advertencias a tiempo sobre los problemas.

7.1.1 VARIACION Y NECESIDAD DE CONTROL

Al igual que no hay dos hojas en un árbol exactamente iguales, no hay procesos que sean exactamente iguales.

La meta de la "producción", es producir "productos" de calidad consistente, que cumplan con las especificaciones técnicas (EVANS y LINDSAY, 1995, cap. 8).

Un proceso cualquiera tiene muchas *fuentes de variación*, algunas de ellas son (EVANS y LINDSAY, 1995, cap. 8):

- La variación de lo que entra
- El proceso de conversión
- La medición de lo que sale

Por ejemplo, en la realización de un ensayo químico la distribución de la cantidad y otras características del analito en el material que lo contiene no es del todo uniforme. También el proceso de muestreo (obtención de la muestra) en sí tiene su variabilidad inherente. Durante el proceso de conversión, que en el análisis químico podría asociarse al método de análisis seguido para la determinación del analito, y todo lo que su seguimiento involucra como preparación del material que contiene el analito, preparación de reactivos, etc. Las fuentes de variabilidad podrían asociarse a factores ambientales, humanos, de instrumentación y de equipo de medición (EVANS y LINDSAY, 1995, cap. 8).

Debido a estas variaciones, se pueden expresar las especificaciones de las medidas como valores únicos, sino más bien con las incertidumbres aceptables.

En general, todo proceso tiene cierta variabilidad que no puede atribuirse a una causa única, siendo el resultado de los efectos combinados de muchas. Estas causas responsables de esta variabilidad se llaman causas no asignables o comunes, entre estas causas pueden citarse: variabilidad en material entrando, precisión y exactitud propias del método de análisis, destreza de los analistas, etc. Estas causas no asignables hacen que el repetir el proceso en condiciones aparentemente análogas, se obtengan resultados distintos. Aunque no es posible medir las causas asignables en forma individual, los efectos combinados pueden ser estables y ser descritos con bastante exactitud mediante distribuciones de probabilidad (EVANS y LINDSAY, 1995, cap. 8).

Existen otras causas de variabilidad, que cuando actúan, producen ciertos efectos previsibles y definidos; por ejemplo, la descalibración de un instrumento de medición produce mediciones defectuosas, sin embargo al calibrarlo se elimina la causa de variabilidad y los efectos desaparecen. Estas se conocen como causas asignables o especiales de variación (EVANS y LINDSAY, 1995, cap. 8).

Todo proceso de funcionamiento regular tiene variabilidad debida a ambos tipos de causas. Las causas no asignables están presentes siempre produciendo una variabilidad homogénea y estable que es predecible al ser constante. Las asignables solo intervienen en determinados momentos, y producen entonces una variabilidad muy grande. Los defectos debidos a causas no asignables aparecen aleatoriamente y la aparición de un defecto no hace más probable la aparición del siguiente. Por el contrario, los defectos debido a causas asignables se mantienen hasta que se elimina la causa que los produce (EVANS y LINDSAY, 1995, cap. 8). En el análisis químico la variabilidad debido a causas no asignables se conoce como error aleatorio y la variabilidad debida causas asignables se conoce como sesgo.

Estudiando un proceso es posible eliminar sucesivamente las causas asignables de manera que la variabilidad restante sea debida únicamente a causas no asignables. *Se dirá entonces que el proceso se encuentra en estado de control estadístico.*

Ningún proceso se encuentra espontáneamente en estado de control; llevarlo a dicho estado y mantenerlo en él es un logro. Este es el objetivo del control de procesos. Si se tiene presencia de variabilidad debida a causas asignables se dice que el proceso esta fuera de control (EVANS y LINDSAY, 1995, cap. 8).

La responsabilidad de eliminar las causas asignables corresponde al supervisor del proceso: la descalibración de equipo, el error del analista, etc. Son causas directamente detectables y resolubles dentro del marco del proceso. Sin embargo, la responsabilidad de reducir la variabilidad producida por las causas no asignables, que son la mayoría, corresponde a la dirección de la empresa: mejorando la tecnología, cambiando a los proveedores, y, en general, mejorando el proceso (PEÑA, 1986).

En resumen, cuando un proceso esta en estado de control, la variabilidad es constante a lo largo del tiempo y, por tanto, predecible. Por el contrario, cuando el proceso esta fuera de control la variabilidad no es constante, siendo sus valores futuros impredecibles (PEÑA, 1986).

En la figura 7.2, se presenta la representación de procesos bajo y fuera de control.

Causas no asignables	Causas asignables
-Existen muchas, cada una de pequeña importancia.	-Existen un número pequeño pero que producen fuertes efectos
-Producen una variabilidad estable	-Producen una variabilidad imprevisible
-Es difícil reducir sus efectos	-Sus efectos desaparecen al eliminar la causa
Ejemplos: Variaciones debidas al material entrante al proceso, a diferencias de habilidad de los analistas, a factores ambientales.	Variabilidad debida a desajuste, errores humanos, reactivos defectuosos, fallos de controles.

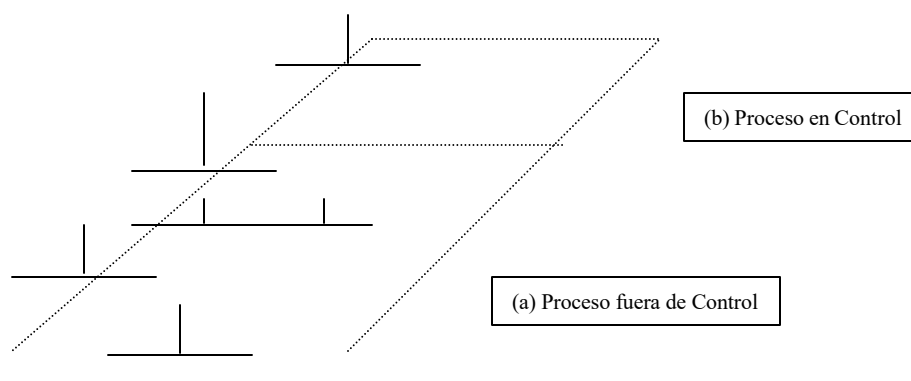


FIGURA 7.3 (b) PROCESOS BAJO CONTROL Y (a) FUERA DE CONTROL
(PEÑA, 1986)

La salida de un proceso de medición llega a ser creíble cuando este se encuentra en Estado de Control Estadístico (EVANS y LINDSAY, 1995, cap. 8).

Lo anterior, significa que hay evidencia estadística que el proceso de medición genera resultados válidos, es decir (PRICHARD et al., 1995, cap.5):

- El valor medio de un gran número de valores individuales tiende a un valor límite.

- Los valores individuales estarán distribuidos alrededor del valor medio limitante en una distribución estable descrita por una desviación estándar también estable.

Para demostrar la estabilidad de un sistema se requiere el continuo monitoreo del mismo.

7.2 DIVISION DEL CONTROL DE CALIDAD

Con frecuencia el Control de Calidad, se divide en dos categorías (PEREZ, 1999, cap. 7):

- a) Control de calidad en línea o control de procesos.
- b) Control de calidad fuera de línea (de recepción y producto acabado)

El en línea consiste en todas las actividades de control que se llevan a cabo durante el proceso. Estas actividades usualmente se llaman *control del proceso*. Los métodos en línea tienen un ciclo de realimentación para tener acción correctiva inmediata. Los métodos fuera de línea son los que se llevan a cabo externamente al proceso. En el laboratorio de ensayo estas dos categorías de control se entenderían mejor como control de calidad interno y control de calidad externo (o valoración de la calidad).

7.2.1 CONTROL DE CALIDAD INTERNO EN EL LABORATORIO DE ENSAYO

Los elementos básicos del control de calidad interno en un laboratorio de ensayo son (GARFIELD, 1984, cap. 2):

- Competencia técnica del personal
- Facilidades convenientes y equipo
- Buenas prácticas de laboratorio (BPLs)
- Buenas practicas de medición (BPMs)
- Procedimientos de operaciones estándar (POEs)
- Protocolos para procesos específicos (PPEs)
- Inspección
- Documentación

- Entrenamiento

En general, los controles de calidad internos en un programa de Aseguramiento de la calidad deben incluir (PRICHARD et al., 1995, cap. 5):

- a. Inspección y mantenimiento preventivo de equipo e instrumentación
- b. Inspección y pruebas periódicas y ajuste para funcionamiento apropiado
- c. Mantenimiento de la destreza del personal
- d. Validación de métodos
- e. Etiquetado de todos los reactivos y soluciones
- f. Uso de materiales de vidrio y reactivos de alta calidad
- g. Disponibilidad de un manual de procedimientos del laboratorio; uso exclusivo de métodos documentados
- h. Localización de los manuales en un área definida
- i. Manuales actualizados y diseñados para reflejar revisión de supervisión anual (escrito para la aprobación de todos los cambios)
- j. Calibración apropiada de pipetas y buretas
- k. Procedimientos de estandarización para reactivos convenientes
- l. Uso de muestras representativas, muestras replicadas, control de muestras, blancos, muestras ciegas y materiales de referencia
- m. Observación directa de ciertos ensayos de evaluación críticos para controlar errores
- n. Monitoreo continuo de los resultados
- o. Mantenimiento de cartas de control y registros apropiados
- p. Comparación de medias de los resultados replicados del laboratorio con otros laboratorios
- q. Medidas correctivas para mejorar la calidad de los datos
- r. Respuesta a las quejas de los usuarios
- s. Actividades específicas para cada técnica analítica en el laboratorio

7.2.2 CONTROL DE CALIDAD EXTERNO EN EL LABORATORIO DE ENSAYO (VALORACION DE LA CALIDAD).

La valoración de la calidad tiene dos componentes que son (APHA, 1996, cap.1):

1. La valoración interna de la calidad: que permite día a día en un laboratorio particular precisión en los resultados.
2. La valoración externa de la calidad: que evalúa la exactitud en los resultados.

Los elementos básicos de la valoración de la calidad son los siguientes (APHA, 1996, cap.1):

Para la valoración interna:

- Mediciones repetidas
- Pruebas internas de muestras
- Mediciones de blancos analíticos
- Cartas de control
- Análisis de precisión
- Intercambio de operadores
- Auditorias

Para la valoración externa:

- Pruebas colaborativas
- Intercambio de muestras
- Materiales Estándar de referencia
- Auditorías

7.3 CONTROL ESTADÍSTICO DE CALIDAD

El Control Estadístico de Calidad, es la aplicación de técnicas estadísticas para medir y mejorar la calidad de los procesos. El Control Estadístico de Calidad incluye (EVANS y LINDSAY, 1995, cap. 8):

- El control Estadístico del proceso
- Herramientas de diagnosis
- Planes de muestreo

Las técnicas de control Estadístico de procesos, comenzaron a ser desarrolladas en 1920 en EEUU por W.A. Shewart, cobrando especial importancia su utilización durante la segunda guerra mundial en las empresas de armamento. Hasta entonces las pruebas de calidad que se adoptaban en las empresas estaban basadas en la inspección de productos acabados, eliminando los defectuosos. El procedimiento se reveló ineficaz, por lo que el control de calidad se desplazó al proceso. A partir de entonces el Control de Procesos ha ido evolucionando respondiendo a las necesidades de las empresas, dando lugar a dos corrientes (PRAT et al., 2000, cap. 11):

1. La primera sigue denominándose *Control Estadístico de Procesos (CEP)* (Statistical Process Control, SPC), y ha estado más relacionada con las industrias de producción en serie y de servicios. Este se desarrolló principalmente a partir de la crisis de los años 70 en empresas relacionadas con el sector de automoción.
2. A la segunda se le denomina *Control Adaptativo o Automático de Procesos (CAP)* (Automatic Process Control, APC), y ha estado mas ligado a empresas de producción continua, como pueden ser las empresas químicas.

Hoy en día la utilización de unas técnicas u otras es compartida cada vez más por ambos sectores. Sin embargo, en el área de control de procesos de medición en laboratorios de ensayo el CEP es el más aplicable.

7.3.1 CONTROL ESTADÍSTICO DE PROCESOS (CEP)

El Control Estadístico de Procesos (CEP), es la aplicación de técnicas estadísticas para la medición y análisis de las variaciones de los procesos.

Los objetivos principales del CEP son los siguientes (PRAT et al., 2000, cap. 11):

- I. Minimizar la "producción" defectuosa
- II. Mantener una actitud de mejora continua del proceso
- III. Comparar respecto a las especificaciones

Bill Hunter, dice: "Todo proceso genera un "producto", pero además genera información", por lo que las técnicas del CEP han de ser aplicadas lo más próximas posible al proceso

que genere información para poder disminuir el tiempo de reacción ante el proceso. Por eso han de ser sencillas de utilizar e interpretar (PRAT et al., 2000, cap. 11).

Hay dos razones básicas para aplicar el CEP (EVANS y LINDSAY, 1995, cap. 8):

- La primera, es que el CEP permite determinar *cuando emprender acciones para ajustar el proceso que se ha salido de control, (presencia de causas asignables de variación)*.
- La segunda, es que el CEP señala cuando *dejar solo un proceso*.

El CEP se puede considerar como si se efectuara la prueba de una hipótesis (EVANS y LINDSAY, 1995, cap. 8):

H₀: el proceso está bajo control

Comparándola con:

H₁: el proceso está fuera de control

Las decisiones correctas son ajustar el proceso cuando esta fuera de control y dejarlo solo cuando esta bajo control. Naturalmente, se pueden tomar decisiones incorrectas por errores de muestreo o errores en la interpretación de los datos. El riesgo de ajustar el proceso bajo control equivale al error **tipo I**; si no se corrige el proceso que esta fuera de control, equivale al error **tipo II**. La aplicación correcta del CEP reduce al mínimo estos riesgos (EVANS y LINDSAY, 1995, cap. 8). En la tabla 7.1, se resume lo anteriormente expuesto.

TABLA 7.1 DECISIONES Y RIESGOS EN EL PROCESO DE CONTROL

<i>Decisión</i>	<i>Bajo Control</i>	<i>Fuera de control</i>
Ajustar el proceso	Error tipo I	Decisión correcta
Dejarlo solo	Decisión correcta	Error tipo II

7.4 ESTUDIO DE LA VARIACION

Como se mencionó anteriormente el CEP es la aplicación de técnicas estadísticas para la medición y análisis de las fuentes de variación de los procesos. Existen en general, tres

técnicas estadísticas dirigidas a estos fines: (1) Gráficos de control; (2) Estudios de transmisión del error; (3) Estimación de los componentes de la varianza. Los primeros sirven para descubrir causas de variación asignables y pueden provocar vigilancia y realimentación en los sistemas operativos de control. Los segundos pueden mostrar como los errores introducidos en diversas partes de un sistema se *sintetizan* en sus efectos sobre el resultado final. La tercera permite disgregar la variación final en componentes asociados con causas individuales (BOX et al., 1989, cap. 17).

7.4.1 LA VARIACION. REPRESENTACION GRAFICA

De todos los medios para analizar datos puede que los más valiosos sean los gráficos. La comprensión surge como resultado de la información *correctamente transmitida*; la existencia de información no es suficiente. Lo mismo si el objetivo es controlar una planta de fabricación, la gestión de un laboratorio de pruebas o el cuidado de un paciente, es importante no sólo recoger la información apropiada, sino también que esta información sea transmitida de forma fácilmente comprensible a los responsables de tomar decisiones (BOX et al., 1989, cap. 17).

Desgraciadamente es bastante corriente que se realice el costoso esfuerzo de recoger los datos y no, en cambio el comparativamente menor adicional de transmisión. La información puede permanecer oculta en libretas de notas, tablas, hojas de registro de procesos, o en memorias de ordenadores o en complicadas ecuaciones matemáticas incomprensibles. Para que los datos sean informativos deben presentarse de forma que las experiencias presentes y pasadas puedan compararse fácilmente y la variación concomitante de dos o más respuestas relacionadas pueda considerarse simultáneamente. Sea lo que sea lo que se haga posteriormente en un análisis más formal, la representación gráfica apropiada de los datos no es nunca una pérdida de tiempo y frecuentemente revela características de los datos que de otra forma hubiesen pasado desapercibidas (BOX et al., 1989, cap. 17).

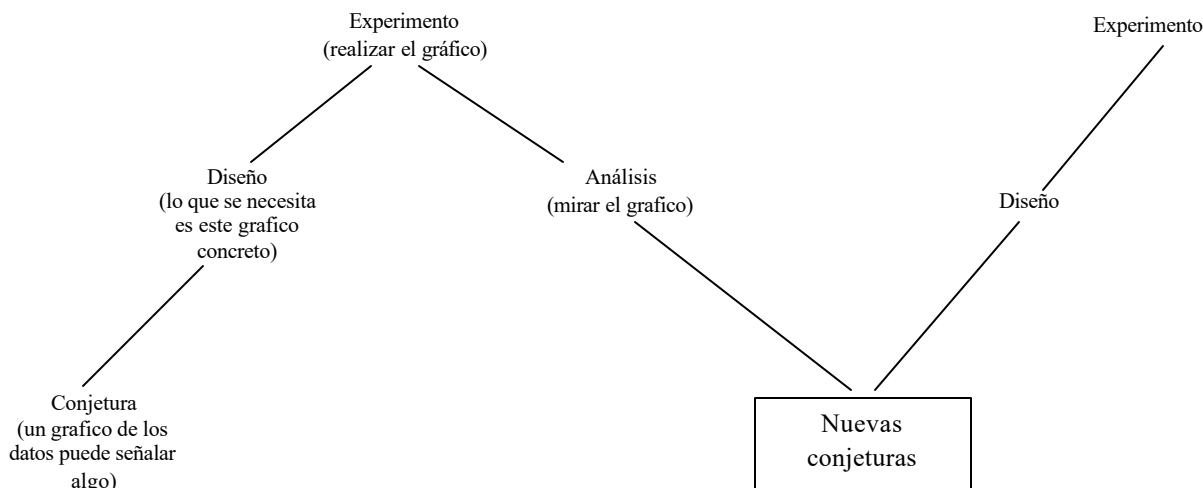


FIGURA 7.4 REPRESENTACION GRAFICA COMO EXPERIMENTACION ITERATIVA

En la figura 7.3 se ilustra que realizar gráficos es un caso especial de experimentación iterativa. Quizás un gráfico revele errores de transcripción, tendencias insospechadas o ciclos. Si es así, hay que examinar con más detalle estos fenómenos. Es posible que esto se pueda conseguir con nuevos gráficos o con algún análisis más formal de los mismos datos; también es posible que sea necesario recoger más datos. En cualquier caso, los gráficos pueden en ocasiones cambiar por completo el curso del estudio. Es un crimen antinatural, pero frecuentemente cometido, el querer forzar a la pieza cuadrada de los datos inapropiados a encajar en el agujero redondo de los métodos estadísticos computarizados (BOX et al., 1989, cap. 17).

Una aplicación muy importante de los gráficos de datos son los *gráficos de control* que se utilizan para la vigilancia de operaciones repetitivas en plantas industriales y laboratorios de ensayo (BOX et al., 1989, cap. 17).

7.5 DIAGRAMAS DE CONTROL

En el programa de garantía de calidad los *diagramas de control* son herramientas útiles, cuyo uso se da en los análisis de rutina más propensos a error (habitualmente sesgo aumentado, pero algunas veces variabilidad aumentada).

Un diagrama de control es un trazo gráfico de los resultados de las pruebas con relación al tiempo o secuencia de mediciones, con límites dentro de los cuales se espera caigan los resultados cuando el esquema analítico esté en un estado de "control estadístico". Un

procedimiento está bajo control estadístico cuando los resultados caen siempre dentro de los límites de control establecidos. Un proceso se encuentra en esta condición cuando se han corregido o eliminado las causas asignables a la variación significativa, permitiendo que se use un grupo de mediciones para predecir límites de variación y para asignar un nivel de confianza en el sentido de que las mediciones futuras estarán dentro de estos límites. Además de identificar los resultados que están fuera de control un diagrama revelará las tendencias y ciclos, por lo cual es una herramienta para proporcionar el tiempo real de análisis de los datos y la información en la que se pueden basar las medidas correctivas apropiadas, que se aplicarán después mediante un investigación subsiguiente del proceso.

7.5.1 CONSTRUCCION DE UN DIAGRAMA DE CONTROL

Para hacer un diagrama de control, las mediciones deben tomarse cuando el proceso esta bajo control. Esto significa que el analista debe estar familiarizado con el método, haber explotado las fuentes de error, por ejemplo, procedimientos de muestreo, técnicas instrumentales y separación de los analitos) y haber realizado suficientes validaciones del método para estar seguros de que los resultados son aceptables. Los siguientes son pasos razonables en la construcción de un diagrama de control (GARFIELD, 1984, cap. 2):

- Seleccionar el tipo de diagrama de control a construir
- Identificar los elementos esenciales del diagrama
- Decidir las metas
- Seleccionar las características que serán medidas
- Decidir cómo, dónde y cuando deberán ser medidas las características
- Decidir que tamaño de subgrupo es apropiado
- Definir el procedimiento para obtener muestras aleatoriamente
- Obtener estimaciones confiables de la media del proceso (sino es conocida de otra manera) y de la desviación estándar a largo plazo
- Colocar los elementos del diagrama (la línea central y los límites preventivos y de acción)
- Colocar los datos

Existen diferentes tipos de diagramas de control, tal como se presenta en la sección 7.5.2. Sin embargo, los diagramas de control de Shewhart son los más fáciles de construir, usar e interpretar. Probablemente su característica más útil sea la identificación de resultados individuales remotos (BOX et al., 1989, cap. 17). La figura 7.4 es un modelo estándar para un diagrama de control de Shewhart y de todos sus elementos.

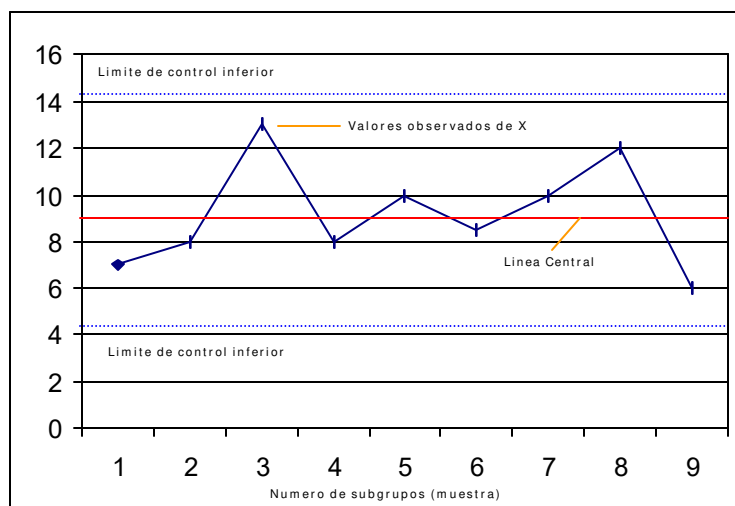


FIGURA 7.5 ELEMENTOS DE UNA GRAFICA DE CONTROL

7.5.2 TIPOS DE CARTAS DE CONTROL

Existen en general, dos clasificaciones para las graficas de control en base al tipo de variables que se manejen, y estas son (MONTGOMERY, 1991, cap. 4):

- a) **Graficas de control para variables**
- b) **Graficas de control para atributos**

Las de mayor aplicabilidad en el área de análisis químico son las *graficas de control para variables*. De las cuales hay dos tipos principales (GARFIELD, 1984, cap. 2):

1. Cartas de control para propiedades
2. Cartas de control para precisión

Habitualmente, estos dos tipos de cartas se usan en asociación.

Existen diferentes tipos de cartas de control para propiedades entre las que se encuentran: las graficas de medias y rangos, graficas de mediciones individuales, gráficos de sumas acumulativa o ACUSUMA y las gráficas de medias móviles y cartas de control de mediciones dependientes.

En la tabla 7.2, se presenta una guía comparativa de los distintos tipos de cartas de control.

TABLA 7.2 GUIA COMPARATIVA DE TIPOS DE FICHAS DE CONTROL PARA VARIABLES

GRAFICO	Parámetros representados	Uso Característico	Ventajas	Desventajas	Comentarios
\bar{X} y R/s	Media y recorrido (o desviación tipo) del subgrupo	Proceso con predominio de máquina	Un buen observatorio de las variaciones estadísticas de un proceso	Cálculos complejos; respuesta lenta; relación indirecta entre límites de control y tolerancia	Selección del tamaño de los subgrupos, frecuencia y cantidad de subgrupos utilizados para fijar y modificar cuidadosamente los límites de control.
X y R	Unidades y recorrido del subgrupo	Cuando solo es posible una sola observación por lote	Muy rápido y fácil de complementar y explicar. Directamente comparable con las tolerancias.	No tan sensible como la ficha de \bar{X} y R	
ACUSUMA	Sumas acumuladas de desviaciones de la media del subgrupo a partir de un valor de referencia	Productos o ensayos de costo elevado en los que es corriente una variación de 0.5σ a 2σ	Más rápida respuesta a un cambio brusco en la media que la ficha de \bar{X} y R	Complejo, difícil de explicar	El Pre-Control es más fácil

7.6 CARTAS DE CONTROL DE SHEWHART

7.6.1 CARTAS DE CONTROL DE SHEWHART PARA PROPIEDADES: DE MEDIAS Y MEDICIONES INDIVIDUALES

Estas cartas pueden utilizar: *Mediciones individuales* o *El promedio de varias mediciones* de una propiedad seleccionada de *una muestra de control*.

Una muestra de control (MC) es una cantidad de material que ha sido seleccionada para hacer mediciones de un parámetro de calidad dado de un material, con el propósito de controlar el proceso de medición. Una muestra seleccionada para el control estadístico de calidad debe cumplir con los siguientes criterios, para asegurar la propia evaluación del método de ensayo (BOVER, 1990):

- La MC generalmente es similar a los materiales típicamente analizados por el método de ensayo.
- Algunas veces la MC puede ser un material de referencia o estándar certificado.
- La MC debe ser estable con respecto a la propiedad medida sobre el periodo supuesto de uso y bajo condiciones adecuadas de almacenaje.
- La MC debe ser homogénea, o el procedimiento debe ser establecido para procurar homogeneidad en una muestra de material a granel para análisis.
- El valor más probable de la propiedad medida, debe ser conocido o determinado por análisis repetidos.

Los elementos básicos de las cartas de control de medias y mediciones individuales, son los siguientes (BOVER, 1990):

1. La línea central, que puede ubicarse en la mejor estimación del valor medio de la variable o en el valor certificado de un material de referencia.
2. Los límites de control de sesgo cada uno localizado inmediatamente abajo y arriba de la línea central. Estos son dibujados a una desviación estándar de la línea central ($\pm 1\sigma$).
3. Límites de control de alarma cada uno localizado abajo y arriba de los límites de control de sesgo y están colocados a dos desviaciones estándar de la línea central ($\pm 2\sigma$).
4. Límites de control de acción cada uno localizado después de los límites de alarma (arriba y abajo). Ellos están a tres desviaciones estándar del valor medio ($\pm 3\sigma$).

5. Zona de alarma es el área de la carta localizada entre los límites de alarma y los límites de acción.

Para la construcción de una carta de control deben considerarse los siguientes aspectos:

- I. Asumir una distribución normal cuando se usen los límites de control
- II. Cuando se construye un gráfico para la media utilizar al menos 25 valores de \bar{X} en la etapa inicial de construcción del mismo.
- III. Si la desviación estándar es estimada experimentalmente debe estar basada sobre al menos 25 mediciones independientes.
- IV. La desviación estándar es usada para determinar los límites de control.
- V. Las mediciones deben ser hechas preferiblemente en diferentes días.

Los límites de control para los diagramas de características se calculan como se presenta en la tabla 7.3.

Cuando se dispone un número relativamente grande de datos y el sistema está bajo control (APHA, 1996, cap. 1):

- ± 1 desviaciones estándar deben contener 2/3 de los puntos
- ± 2 desviaciones estándar deben contener 19/20 de los puntos
- ± 3 desviaciones estándar deben contener todos los puntos

Los indicadores de fuera de control son los siguientes (PEREZ, 1999, cap. 7):

- Un punto fuera de los límites de control de tres sigma
- 2 de 3 puntos sucesivos fuera de los límites 2 sigma
- 4 de 5 puntos sucesivos fuera de los límites 1 sigma
- 8 puntos consecutivos se hallan al mismo lado de la línea central
- Cualquier patrón anormal o no aleatorio en los datos (tendencia sistemática) (ver anexo C7)

TABLA 7.3 ELEMENTOS DE GRAFICAS DE CONTROL PARA PROPIEDADES

<p>Mediciones individuales (carta de X)</p> <p>Línea central Limites de alarma LA Limites de control LC</p>	<p>\bar{x} o valor conocido* $\pm 2s$ (ó σ conocido)* $\pm 3s$ (ó σ conocido)*</p>
<p>Promedios de subgrupos (carta de \bar{X})</p> <p>Línea central Limites de alarma LA Limites de control LC</p>	<p>$\bar{\bar{X}}$ o valor conocido* $\pm 2s/\sqrt{n}$ (ó σ conocido)* $\pm 3s/\sqrt{n}$ (ó σ conocido)*</p>

(*) Cuando los valores de \bar{x} y σ son dados como valores especificados, las graficas de control se nombran como gráficas con patrón dado. Se puede demostrar que $s \approx \bar{R} / d_2$

7.6.2 CARTAS DE CONTROL DE LA PRECISI ON

Estas utilizan las estimaciones de la desviación estándar de las mediciones del proceso. Las cartas típicas de precisión son las siguientes: **a) Carta de control de rangos; b) Carta de control de desviaciones estándar (o varianzas).**

Estas cartas utilizan las mediciones replicadas de muestras de ensayo comunes. En las cartas de control de rango, este es utilizado como una estimación de la desviación estándar.

Estas cartas permiten monitorear la precisión de las mediciones pero no permiten evaluar los rumbos sistemáticos que podrían afectar a cada medición en la misma forma.

Los límites de control para los diagramas de dispersión se calculan como se presenta en la tabla 7.4.

La interpretación de las cartas de precisión es algo similar a la interpretación de las cartas de control para propiedades. Cinco o más puntos consecutivos arriba del límite de confianza de 50% indican una tendencia del proceso a salirse de control. Más de un punto arriba del límite preventivo superior en veinte determinaciones duplicadas es causa de preocupación. Cualquier punto arriba del límite de control superior debe ser causa de alarma (PRICHARD, 1995, cap. 5).

TABLA 7.4 ELEMENTOS PARA LA CONSTRUCCION DE GRAFICOS DE DISPERSION

<p>Cartas de control de rangos (carta de R)</p> <p>Línea central</p> <p>Limites de alarma LA</p> <p>Limites de control LC</p>	<p>\bar{R} (promedio de rangos basado en 10 a 15 análisis por duplicado en diferentes tiempos)</p> <p>$\frac{2}{3}(D_4\bar{R} - \bar{R}) + \bar{R}$</p> <p>$LCS = D_4\bar{R}$</p> <p>$LCL = D_3\bar{R}$</p> <p>(para análisis por duplicado $n=2$ y $D_4 = 3.27$ y $D_3 = 0$, para otro número de réplicas estos valores pueden encontrarse en la tabla del anexo A7)</p>
<p>Carta de control de desviaciones estándar (o varianzas (carta de s o s^2))</p> <p>Línea central</p> <p>Limites de alarma LA</p> <p>Limites de control LC</p>	<p>\bar{S} o valor conocido</p> <p>$B_4\bar{S}$</p> <p>$B_3\bar{S}$</p> <p>(Los valores de B_3 y B_4 pueden encontrarse en la tabla del anexo A7)</p>

La elaboración general de gráficos de control para propiedades y precisión puede hacerse manualmente o mediante el apoyo de software como hojas de cálculo, programas estadísticos, o bien módulos informáticos propios para el área de control de calidad. En el anexo B7 se presenta una plantilla para la recogida de datos y para la elaboración de gráficos de control.

Las cartas de control para mediciones individuales son las de más frecuente aplicación en el análisis químico. Para la construcción de gráficos de mediciones individuales en el laboratorio, se recomienda, para obtener mejores resultados, realizar al menos 15 determinaciones de la MC antes de calcular la desviación estándar y construir la gráfica de control. No necesariamente todos los resultados deben obtenerse el mismo día.

Óptimamente la acumulación de los mismos debe formar parte de la operación diaria del laboratorio (APHA, 1996, cap. 1).

Las cartas de control de \bar{X} - R(ó s) (realmente dos cartas pero usualmente tratadas como una sola) son probablemente las más utilizadas en aplicaciones industriales. Para construir cartas \bar{X} - R(ó s), en el laboratorio deben acumularse al menos 25 pares de determinaciones duplicadas de la MC, puede utilizarse también, mayor número de réplicas, sin embargo, el tiempo y el costo involucrados aumentan al aumentar el número de réplicas, por lo que la carta \bar{X} - R(ó s) en el control del laboratorio se limita muchas veces a los análisis por duplicado (APHA, 1996, cap. 1).

Es deseable hacer el proceso tan sensible como sea posible para detectar situaciones fuera de control teniendo cuidado de no rechazar los datos que están realmente bajo control. Cuando se toma la decisión de que un proceso está fuera de control, se marca el diagrama, los análisis se detienen y se investigan la o las causas. Cuando se hace la corrección y el sistema está otra vez bajo control se reanuda la preparación del diagrama (EVANS y LINDSAY, 1995, cap. 16).

Los límites de control de los gráficos deben ser recalculados periódicamente, según lo especifiquen los procedimientos de operación del laboratorio (APHA, 1996, cap. 1).

Existen muchas causas para un cambio o tendencia en el diagrama usado en análisis químico: por ejemplo, la preparación incorrecta del estándar o reactivo, contaminación de la muestra, calibración incorrecta de los instrumentos, técnicas analíticas deficientes y desviaciones del método (APHA, 1996, cap. 1). En el anexo C7, se presenta el tema general de cómo interpretar comportamientos o tendencias en gráficas de control.

7.7 DIAGRAMAS DE SUMA ACUMULATIVA (ACUSUMA)

Otro tipo de diagramas es el diagrama de control de sumas acumulativas (ACUSUMA). El diagrama de acusuma tiene la ventaja de que identifica cambios persistentes pequeños en el esquema analítico más rápido que en los diagramas antes descritos. Los gráficos de control ACUSUMA surgieron como una alternativa a los gráficos de Shewhart. Las diferencias principales entre los dos gráficos se deben a los objetivos que se persiguen (MILLER y MILLER, 1993, cap. 4).

Mientras que el objetivo de los gráficos de Shewhart es detectar la aparición de causas asignables de variabilidad, el objetivo del ACUSUMA es algo diferente. En el sistema de

diagramas de ACUSUMA, cada resultado se compara con una referencia valor nominal o Target T y se pretende detectar cualquier evidencia de alejamiento por parte del proceso de T en una magnitud superior al valor preestablecido. Este valor T puede ser, dependiendo de la característica en estudio (PRAT et al., 2000, cap. 11):

- Un valor constante: el valor nominal de una variable continua, la varianza del proceso, etc.
- Un valor no constante: los valores que predice un modelo teórico.

En el caso particular en el que se pretende controlar la media de cierta característica suele tomarse como valor de control el nominal u objetivo $T = \mu$. Las diferencias con la referencia son acumuladas, respetando su signo, para obtener una suma acumulativa de las diferencias con el estándar. En otras palabras, se calcula la diferencia entre cada resultado y la media, y la segunda diferencia se suma a la primera, la tercera diferencia se suma a la suma de la primera y la segunda, y así sucesivamente. La ACUSUMA se traza frente al tiempo. Su virtud más importante es que elimina las variaciones a corto plazo, pero resalta las diferencias persistentes o recurrentes. Por lo tanto, en un gráfico ACUSUMA la magnitud del valor presentado no tiene tanto interés como en los gráficos de Shewhart, pues aquí la importancia la tiene la *pendiente* que forma la trama de puntos. Por lo tanto, los límites de control en lugar de estar formados por líneas paralelas estarán formados por dos "*pendientes*", que representarán las máximas pendientes permitidas antes de concluir que hay pruebas de que hay causas asignables actuando en el proceso provocando un cambio en la media superior al admitido. La pendiente de los límites de control dependerá de cuatro factores (PRAT et al., 2000, cap. 11):

- La escala del gráfico
- La variabilidad innata del proceso σ
- El cambio en el parámetro del proceso que se pretende detectar
- El riesgo que se admite tomar en las decisiones

La ACUSUMA proporciona advertencias más rápidas sobre la desviación del proceso del control y permite "ver y mirar" la estimación gráfica de las tendencias. Algunos usuarios de

diagramas prefieren los diagramas de propiedades y de dispersión porque son más fáciles de interpretar (PRAT et al., 2000, cap. 11).

En el anexo D7, se proporcionan las instrucciones para la construcción de diagramas de acusuma.

7.8 DIAGRAMAS DE MEDIA MÓVIL

Se puede comprobar que las graficas de \bar{X} -R (ó s) son más eficientes para detectar pequeños cambios en el proceso a medida que el tamaño del subgrupo se incrementa (MONTGOMERY, 1991, cap. 6).

Un arreglo razonable entre el uso de mediciones individuales y uso de subgrupos grandes es el usar diagramas de media móvil. La media móvil para n muestras es el promedio de los resultados para tales muestras, y el rango móvil es la diferencia entre los resultados extremos en el grupo de n muestras. El promedio móvil evidencia variaciones entre los resultados. El rango móvil también puede graficarse y servir como una medida de la dispersión. De esta forma es posible obtener cercanamente la misma información utilizando mediciones individuales sobre una MC, y evitándose de este modo invertir costos y tiempo adicional en hacer determinaciones por duplicado (APHA, 1996, cap. 1). En el ejemplo 7 del capítulo se detalla la construcción de un diagrama de media móvil.

7.9 DIAGRAMAS DE CONTROL PARA OBSERVACIONES DEPENDIENTES

Existen muchas situaciones en las que es necesario el control simultáneo de dos o más características relacionadas. Por ejemplo supóngase que se quiere controlar la linealidad de la operación en un equipo de medición en el que se relacionan lecturas propias del equipo (como áreas, absorbancias, etc.) con concentraciones de analitos. En estos procesos de control, la aplicación de las técnicas clásicas del CEP conlleva a la aparición continua de patrones anómalos sin que se pueda actuar sobre ellos en la mayoría de las veces (MONTGOMERY, 1991, cap. 7).

Los problemas de control de calidad en los que hay varias variables de interés relacionados se llaman a veces diagramas de control multivariados. En el anexo E7 se describen los

principios para la construcción de un diagrama de control divariado (MONTGOMERY, 1991, cap. 7).

En los ejemplos del capítulo 7, se presenta el procedimiento general para la construcción manual de los distintos tipos graficas de control.

7.10 EJERCICIOS DE CONSTRUCCION Y ANALISIS DE GRAFICAS DE CONTROL

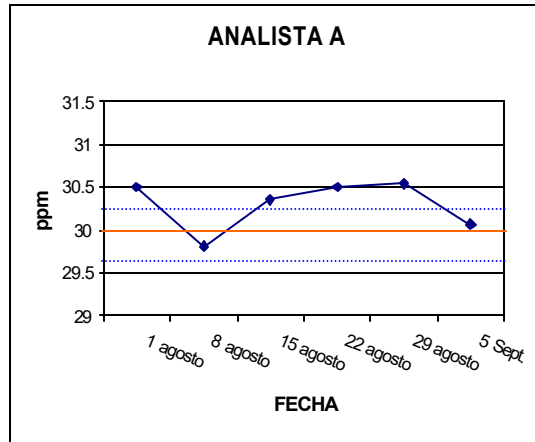
1. En un laboratorio se lleva la determinación rutinaria de un contaminante atmosférico para el cual $\sigma = 0.13$ ppm. Era importante mantener la concentración de ese contaminante en concreto por debajo de 30 ppm y para conseguirlo se analizaban cada semana varias docenas de muestras de aire. Como control cada semana se analizaba también una muestra de aire que contenía exactamente 30 ppm de ese contaminante. Los siguientes datos muestran el promedio de la determinación hecha sobre la muestra estándar obtenido por dos analistas durante un periodo de 6 semanas (BOX et al., 1989, cap. 17):

Muestra	Analista A	Analista B
1 agosto	30.50	30.25
8 agosto	29.80	29.90
15 agosto	30.35	30.10
22 agosto	30.50	30.25
29 agosto	30.55	30.40
5 Sept.	30.07	29.92

¿Esta en control el proceso de medición?

SOLUCION

Para investigar si el proceso de medición esta bajo control, podría recurrirse a la elaboración de gráficos de control para patrón dado es decir fijando el valor central como 30 ppm y el σ como 0.13 ppm, ya que son valores que como requisito tiene que cumplir la muestra estándar para que la medición del contaminante en la atmósfera sea efectiva. Si se elabora un grafico de control para los datos obtenidos por el analista A, el patrón de comportamiento del proceso de medición es el siguiente:



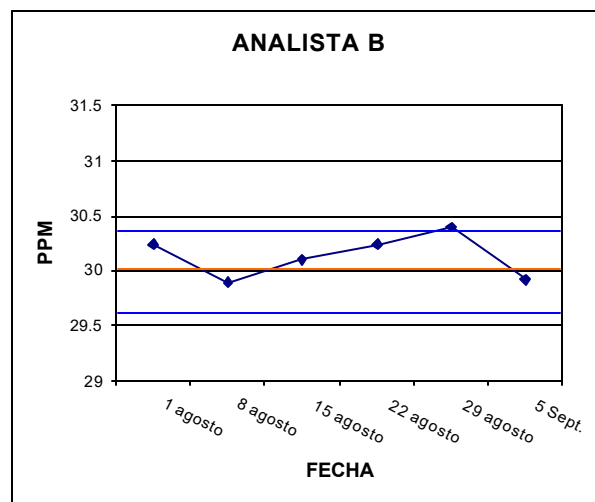
La línea central y de los límites de control vienen dados por:

$$LSC = \bar{X} + 3\sigma = 30 + 3(0.13) = 30.39$$

$$\text{LINEA CENTRAL: } \bar{X} = 30$$

$$LIC = \bar{X} - 3\sigma = 30 - 3(0.13) = 29.61$$

De acuerdo al gráfico puede observarse que el proceso de medición efectuado por analista A se encuentra fuera de control. Habría que analizar el comportamiento del proceso por parte del analista B, para identificar si el "fuera de control" del proceso es debido al trabajo del analista A. A continuación se presenta el gráfico de control para el analista B, los límites del gráfico continúan siendo los mismos ya que en ambos casos se trata de gráficos con patrón dado.



Como puede observarse para ambos gráficos de control, cambios de nivel similares ocurrían entre semana y semana para los dos analistas. Evidentemente, algún factor afectaba conjuntamente a los resultados de ambos analistas. Anteriormente

se había prestado atención a la humedad relativa del laboratorio y hallado que los resultados de los análisis estaban muy correlacionados con esta variable. Un pequeño número de experimentos especiales demostraron luego que la humedad relativa afectaba realmente a la respuesta medida, por lo que los procedimientos de ensayo analíticos se modificaron para eliminar la influencia de ese factor.

2. Con el objeto de dimensionar las instalaciones de una planta de tratamiento de aguas, debe llevarse el registro de la variación de contaminantes del agua residual saliendo del proceso, para lo cual se ha pensado medir esos contaminantes en términos de la demanda bioquímica de oxígeno (DBO) y de los sólidos en suspensión (SS) y como es su variación en 24h de proceso. Construya una gráfica de control que permita observar la variabilidad de la contaminación del agua residual en un período típico de proceso. A continuación se presentan los datos obtenidos:

DATOS Y SOLUCION:

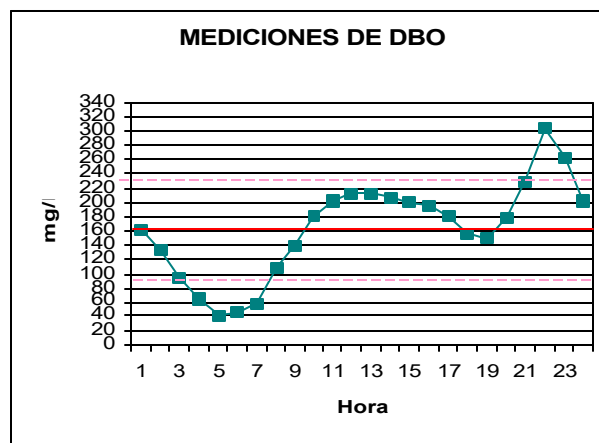
A continuación se presenta la tabla de datos:

HORA	DBO mg/l	SS mg/l	qi m ³ /s	Rango DBO	Rango SS
12	161	172	6.8		
1 a.m.	132	143	5.3	29	29
2	93	105	3.9	39	38
3	64	77	3.2	29	28
4	41	47	2.5	23	30
5	45	40	2.3	4	7
6	59	42	3.3	14	2
7	108	85	3.4	49	43
8	139	196	6.4	31	111
9	180	251	8.9	41	55
10	202	263	9.6	22	12
11	211	274	9.8	9	11
12	213	261	9.6	2	13
1 p.m.	208	249	9.4	5	12
2	200	225	8.7	8	24
3	195	195	8	5	30
4	182	161	7.5	13	34
5	156	147	7.3	26	14
6	150	145	7.5	6	2
7	179	169	8	29	24
8	230	198	8.9	51	29
9	305	206	9.1	75	8
10	262	201	8.7	43	5
11	203	180	7.8	59	21
TOTAL	3918	4032			
PROM.	163.25	168	164.9	26.61	25.30

Donde: DBO es la demanda bioquímica de oxígeno; SS son los sólidos en suspensión y q_i es el caudal medido para cada hora.

Los gráficos de control para cada variable vienen dados por:

- Para la demanda bioquímica de oxígeno (DBO):

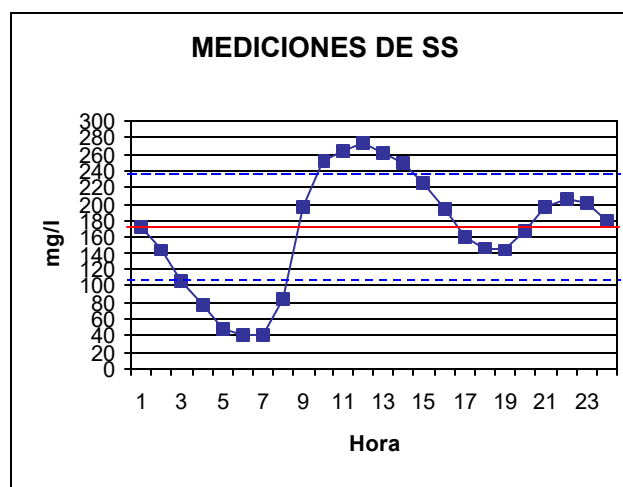


Si se utiliza la amplitud móvil de $n = 2$ observaciones, $d_2 = 1.128$, para los datos de la tabla los límites de control vienen dados por:

$$\bar{X} + 3 \frac{\bar{R}}{d_2} = 163.25 + 3 \frac{26.61}{1.128} = 234.02$$

$$\bar{X} - 3 \frac{\bar{R}}{d_2} = 163.25 - 3 \frac{26.61}{1.128} = 42.48$$

- Para los sólidos en suspensión (SS) se tiene:



Los límites de control vienen dados por:

$$\bar{X} + 3 \frac{\bar{R}}{d_2} = 168 + 3 \frac{25.304}{1.128} = 235.3$$

$$\bar{X} - 3 \frac{\bar{R}}{d_2} = 168 - 3 \frac{25.304}{1.128} = 100.70$$

El rango (amplitud) móvil para cada variable se encuentra calculado en la tabla inicial de datos en las columnas 6 y 7.

Como puede observarse de ambas figuras las características del agua residual que debe tratarse en la planta de tratamiento son extremadamente variables en las diferentes horas de proceso, puede deberse en parte a que el caudal o el nivel de producción varía durante los diferentes turnos del día.

Dado que para el diseño de la planta de tratamiento se necesita de un valor promedio del DBO y de los SS más representativo, sería conveniente hacer una estratificación en el estudio, considerando los distintos niveles de producción que puedan darse y luego, dimensionar la planta en base a los promedios más elevados, es decir, en base al horario de mayor productividad. Un indicador de esta productividad puede ser el caudal de agua. En bibliografía de tratamiento de aguas se recomienda calcular el promedio de la propiedad mediante la siguiente fórmula (METCALF y EDDY, 1996, cap. Tomo 1, cap. 3):

$$\text{característica media} = \frac{\sum_{i=1}^n x_i q_i}{\sum_{i=1}^n q_i}$$

Donde x_i es el valor de la característica medida en cada unidad del tiempo de registro y q_i es el caudal en cada unidad de tiempo de proceso considerada.

3. En un laboratorio de análisis químico se mide de rutina la conductancia del agua de alta pureza procesada en el mismo y empleada para la preparación de los reactivos utilizados en los diferentes métodos de ensayo. Los datos resultantes aparecen a continuación. Obtenga un diagrama de mediciones individuales y un diagrama para la amplitud móvil.

TABLA DE DATOS:

MUESTRA	CONDUCTANCIA (microohms)
1	0.0806
2	0.0814
3	0.0810
4	0.0820
5	0.0819
6	0.0815
7	0.0817
8	0.0810
9	0.0809
10	0.0812
11	0.0819
12	0.0815
13	0.0813
14	0.0816
15	0.0810

SOLUCION:

Para el diagrama de control de la medición de la conductancia en agua de alta pureza, los parámetros son:

$$LCS : \bar{X} + 3 \frac{\bar{R}}{d_2}$$

$$\text{Linea Central} : \bar{X}$$

$$LCI : \bar{X} - 3 \frac{\bar{R}}{d_2}$$

El cálculo de la amplitud móvil a partir de cada dos observaciones sucesivas se presenta en la tabla E73.1.

TABLA E73.1

MUESTRA	CONDUCTANCIA (microohms)	RANGO MOVIL (microohms)
1	0.0806	0.0008
2	0.0814	0.0004
3	0.0810	0.0010
4	0.0820	0.0001
5	0.0819	0.0004
6	0.0815	0.0002
7	0.0817	0.0007
8	0.0810	0.0001
9	0.0809	0.0003
10	0.0812	0.0007
11	0.0819	0.0004
12	0.0815	0.0002
13	0.0813	0.0003
14	0.0816	0.0006
15	0.0810	0.0008
PROMEDIO	0.0814	0.00044

Los límites de control para el gráfico de mediciones individuales vienen dados por:

$$LCS : \bar{X} + 3 \frac{\bar{R}}{d_2} = 0.08137 + 3 \frac{0.00044}{1.128} = 0.08254$$

$$\text{Linea Central} : \bar{X} = 0.08137$$

$$LCI : \bar{X} - 3 \frac{\bar{R}}{d_2} = 0.08137 - 3 \frac{0.00044}{1.128} = 0.0802$$

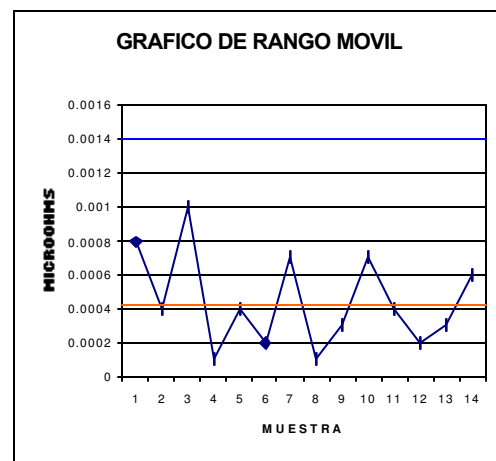
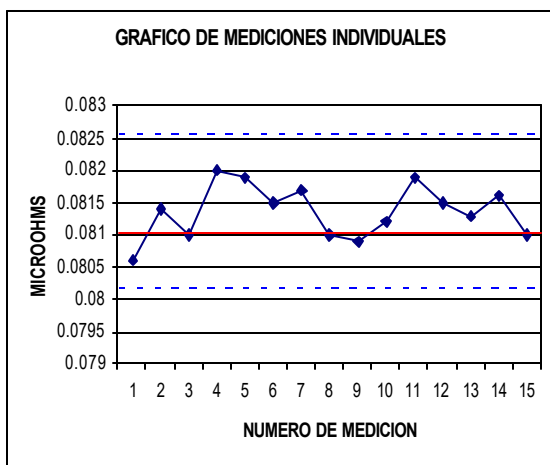
Los límites para el gráfico de rangos móviles son:

$$LCL : \bar{R}D_4 = 0.00044(3.267) = 0.001437$$

$$\text{Linea Central} : \bar{R} = 0.00044$$

$$LCI : \bar{R}D_3 = 0.00044(0) = 0$$

Los gráficos de control se presentan a continuación:



Como puede observarse de los gráficos de mediciones individuales y de rango móvil, si se analiza para los límites de 1σ y 2σ , no existe evidencia de fuera de control en el proceso de purificación del agua efectuado en el laboratorio.

- Un elemento de los procedimientos de control de calidad analítica para ensayos microbiológicos en aguas, es la determinación de la precisión del método, la cual se calcula, según los métodos normalizados, mediante los análisis duplicados de al menos 15 muestras de orígenes diferentes. Un laboratorio quiere acreditar el ensayo del "Recuento total de bacterias" y decide, como parte del proceso, hacer la determinación de la precisión del método. Con los datos proporcionados construya una gráfica de control de rangos para el control de la precisión en el futuro y

establezca el criterio de precisión para este ensayo en el laboratorio. NOTA: tome en consideración que las poblaciones bacterianas tienen distribución log-normal.

TABLA DE DATOS:

MUESTRA	CONTEO M 1	CONTEO M 2
1	3 3 7 6	4 0 0 0
2	1 0 0	1 0 0
3	3 0 0	2 0 0
4	1 3 0	1 6 7
5	6 9 3	6 5 7
6	2 0 0	2 0 0
7	6 0 0	8 0 0
8	3 1 6 0	2 8 9 8
9	2 0 5	1 9 0
10	2 6 9 8	2 9 6 0
11	1 3 3 9	1 4 5 3
12	6 0 9 9	5 0 5 0
13	1 1 8 3	1 8 9 7
14	6 3 8 7	5 7 9 7
15	1 1 9 0 0	1 3 5 4 2

SOLUCION:

La construcción del gráfico requiere primero de la evaluación del rango para cada muestra por duplicado. El hecho de que la variable "Conteo total de bacterias" se distribuya como una log-normal, implica que para la construcción del gráfico de control de rangos, se deberá primero transformar logarítmicamente a la variable original y luego calcular los rangos con la variable ya transformada. La transformación logarítmica para los datos por duplicado de las 15 muestras y la determinación de los rangos se presentan en la tabla E74.1

Una vez transformados los datos y calculado el rango para cada muestra se procederá a calcular los límites de control y a la construcción del gráfico de rangos.

- LIMITES DE CONTROL

Para $n = 2$:

$$LCS : \bar{RD}_4 = 0.069(3.267) = 0.225$$

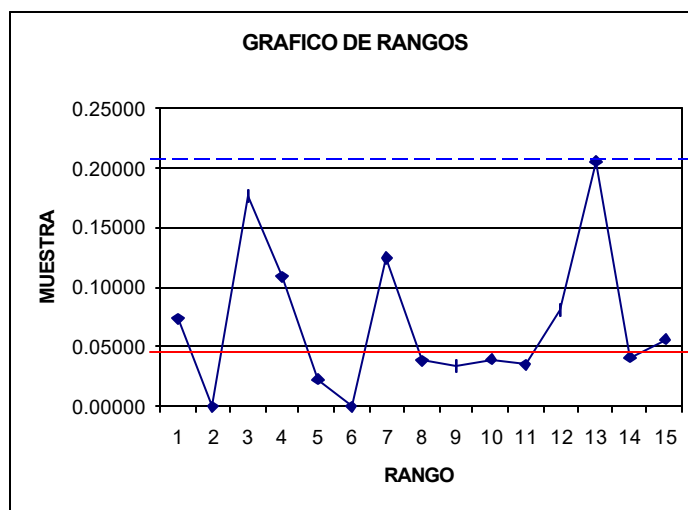
$$\text{Linea Central} : \bar{R} = 0.069$$

$$LCI : \bar{RD}_3 = 0.069(0) = 0$$

TABLA E74.1

MUESTRA	LOG M 1	LOG M 2	LOG M 1 - LOG M 2
1	3.528402438	3.602059991	0.07366
2	2	2	0.00000
3	2.477121255	2.301029996	0.17609
4	2.113943352	2.222716471	0.10877
5	2.840733235	2.81756537	0.02317
6	2.301029996	2.301029996	0.00000
7	2.77815125	2.903089987	0.12494
8	3.499687083	3.462098381	0.03759
9	2.311753861	2.278753601	0.03300
10	3.431041945	3.471291711	0.04025
11	3.126780577	3.162265614	0.03549
12	3.785258633	3.703291378	0.08197
13	3.072984745	3.278067331	0.20508
14	3.805296916	3.7632033	0.04209
15	4.075546961	4.131682809	0.05614

- GRAFICO DE CONTROL



Como puede observarse del gráfico de control, la precisión del método se encuentra bajo control.

- CRITERIO DE PRECISIÓN

Se sugiere analizar por duplicado el 10 por 100 de las muestras convencionales. Se transformarán los duplicados como se ha indicado y se calculará su rango. Si este es superior a 0.225 (que es el LCS de la gráfica de control), la variabilidad de los analistas es excesiva. Se determinará si es aceptable el incremento de la imprecisión; de no ser así, se determinarán todos los resultados analíticos obtenidos desde la última comprobación de la precisión. Se identificará el problema antes de seguir haciendo análisis.

5. Se desea llevar a cabo el control del proceso de producción de un medicamento, siendo la característica de calidad indicadora del proceso el porcentaje de disolución del principio activo de dicho medicamento. Se tomaron 10 muestras de tamaño 5. Calcule los límites de control para las gráficas de \bar{x} y R. A continuación, forme la gráfica y anote los datos e interprete resultados.

DATOS

Numero de muestra					
1	97	80	98	92	98
2	96	99	93	94	93
3	96	97	92	96	94
4	94	98	95	97	97
5	80	93	95	96	94
6	98	97	99	96	96
7	95	97	95	96	97
8	99	95	98	97	96
9	96	95	94	95	95
10	91	93	91	90	93

SOLUCION

En la tabla E75.1 se presenta el cálculo de los promedios y rangos

TABLA E75.1

MUESTRA	\bar{X}	R
1	93	18
2	95	6
3	95	5
4	96.2	4
5	91.6	16
6	97.2	3
7	96	2
8	97	4
9	95	2
10	91.6	3
PROMEDIO	94.76	6.3

En donde la \bar{X} de cada fila es calculada como: $\bar{X}_i = \frac{\sum_{j=1}^5 x_{ji}}{5}$ en donde el subíndice i indica el número de fila y j el número de columna. El rango de cada fila se calcula como el valor absoluto de la diferencia entre el mayor y el menor dato de cada subgrupo.

Los límites de control para el gráfico de promedios vienen dados a partir de: (n = 5)

$$\bar{X} + A_2\bar{R} = 94.76 + 0.572(6.3) = 98.36$$

$$\text{Linea Central} : \bar{\bar{X}} = 94.76$$

$$\bar{X} - A_2\bar{R} = 94.76 - 0.572(6.3) = 91.16$$

Los límites de control para el grafico de rangos son:

$$LCS : \bar{R}D_4 = 6.3(2.114) = 13.31$$

$$\text{Linea Central} : \bar{R} = 6.3$$

$$LCI : \bar{R}D_3 = 6.3(0) = 0$$

Los gráficos de control son los siguientes:

CARTA DE MEDIAS

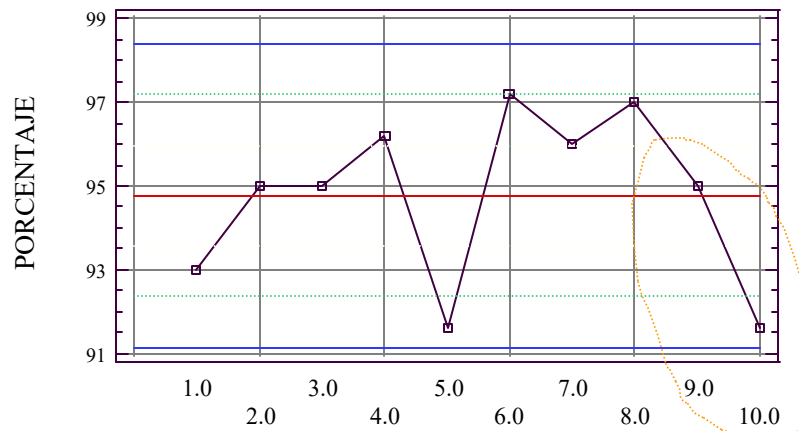
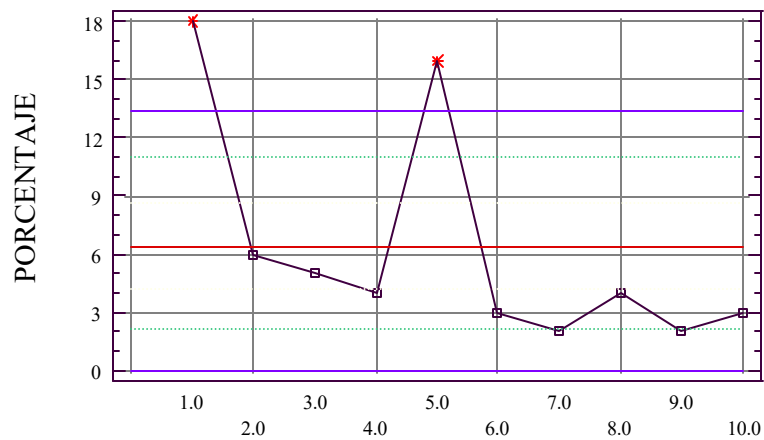


GRAFICO DE RANGOS



Observando los gráficos de medias y rangos pueden notarse los siguientes comportamientos que en general podrían indicar un estado de fuera de control:

- En el gráfico de medias a partir del subgrupo 8 se observa una tendencia de descenso, lo cual debe implicar revisar el proceso ya que si se continúa con esta tendencia en el porcentaje de disolución se corre el riesgo de romper con la garantía de calidad en las tabletas producidas, y puede llegarse a un punto en el que el porcentaje de disolución este por debajo del mínimo requerido (85%), que asegura una buena absorción del principio activo del medicamento por el cuerpo humano.
 - En el gráfico de rangos puede observarse que los subgrupos 2 y 5 sobrepasan los límites de 3σ . Esto indica que existe un cambio brusco en la variabilidad del proceso. Lo anterior puede también asociarse a que el mismo ensayo o análisis de disolución tiene una variabilidad asociada bastante grande, ya que pueden determinarse incertidumbres asociadas a las mediciones hechas con el método ya validado de alrededor del 5%.
6. Se desea construir una grafica de \bar{X} - R con el objeto de llevar a cabo el control de un proceso de medición. Para la construcción del grafico se acumularon 20 pares de determinaciones por duplicado de la concentración de analito en mg/l contenido en una muestra de control. Construya la gráfica de \bar{X} - R con los datos presentados a continuación y concluya sobre el control del proceso.

DATOS

MUESTRA	RESULTADO X_1	RESULTADO X_2	PROMEDIO	RANGO
1	0.501	0.491	0.496	0.010
2	0.490	0.490	0.490	0.000
3	0.479	0.482	0.480	0.003
4	0.520	0.512	0.516	0.008
5	0.500	0.490	0.495	0.010
6	0.510	0.488	0.499	0.022
7	0.505	0.500	0.502	0.005
8	0.475	0.493	0.484	0.018
9	0.500	0.515	0.508	0.015
10	0.498	0.501	0.500	0.003
11	0.523	0.516	0.520	0.007
12	0.500	0.512	0.506	0.012
13	0.513	0.503	0.508	0.010
14	0.512	0.497	0.504	0.015
15	0.502	0.500	0.501	0.002
16	0.506	0.510	0.508	0.004
17	0.485	0.503	0.494	0.018
18	0.484	0.487	0.486	0.003
19	0.512	0.495	0.504	0.017
20	0.509	0.500	0.504	0.009
PROMEDIO			$\bar{\bar{X}} = 0.500$	$\bar{R} = 0.00955$

El promedio de cada celda se calcula como: $\frac{X_1 + X_2}{2}$

Y el rango por celda como: $|X_1 - X_2|$

- CONSTRUCCION DE LAS CARTAS DE CONTROL DE \bar{X} - R

Límites de control para el gráfico de medias: para $n = 2$, $A_2 = 1.88$

$$\bar{X} + A_2\bar{R} = 0.500 + 1.88(0.00955) = 0.51795$$

$$\text{Linea Central: } \bar{\bar{X}} = 0.500$$

$$\bar{X} - A_2\bar{R} = 0.500 - 1.88(0.0095) = 0.48214$$

Límites de control para el gráfico de rangos: $n = 2$, $D_3 = 0$, $D_4 = 3.27$

$$LCS: \bar{R}D_4 = 0.0095(3.27) = 0.0311$$

$$\text{Linea Central: } \bar{R} = 0.0095$$

$$LCI: \bar{R}D_3 = 0.0095(0) = 0$$

GRAFICA DE MEDIAS EJEMPLO 6

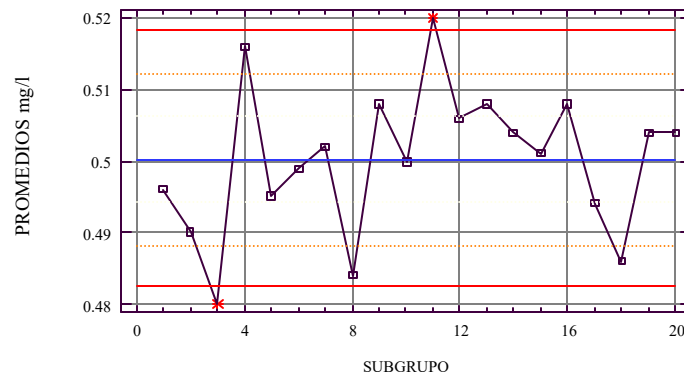
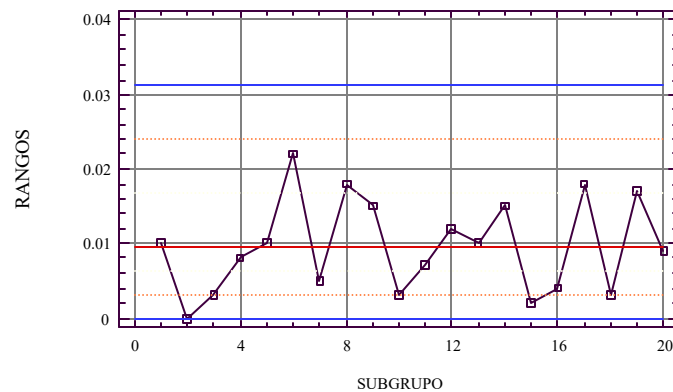


GRAFICO DE RANGOS EJEMPLO 6



Como puede observarse del gráfico de medias los subgrupos 3 y 11, están fuera del límite de 3 desviaciones estándar, por lo que para llevar el control del proceso es necesario recalcular los límites de control sin incluir los subgrupos antes mencionados.

El cálculo de los nuevos límites y la construcción de los gráficos se presentan a continuación:

- Límites de control

Límites de control para el gráfico de medias: para $n = 2$, $A_2 = 1.88$

$$\bar{X} + A_2\bar{R} = 0.5027 + 1.88(0.01) = 0.5192$$

$$\text{Linea Central : } \bar{\bar{X}} = 0.5027$$

$$\bar{X} - A_2\bar{R} = 0.5027 - 1.88(0.01) = 0.4814$$

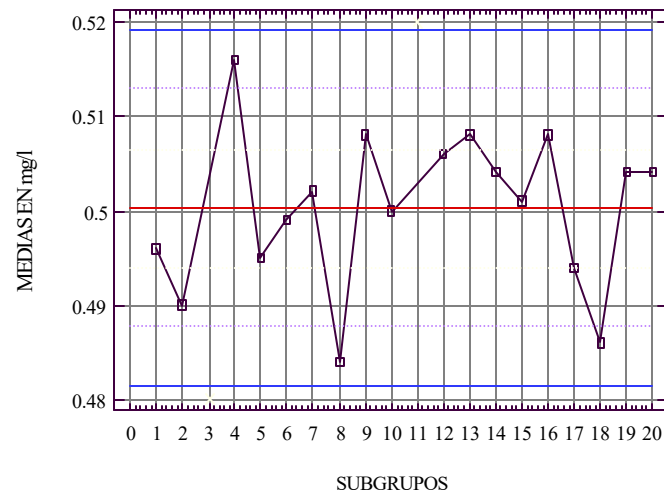
Límites de control para el gráfico de rangos: $n = 2$, $D_3 = 0$, $D_4 = 3.27$

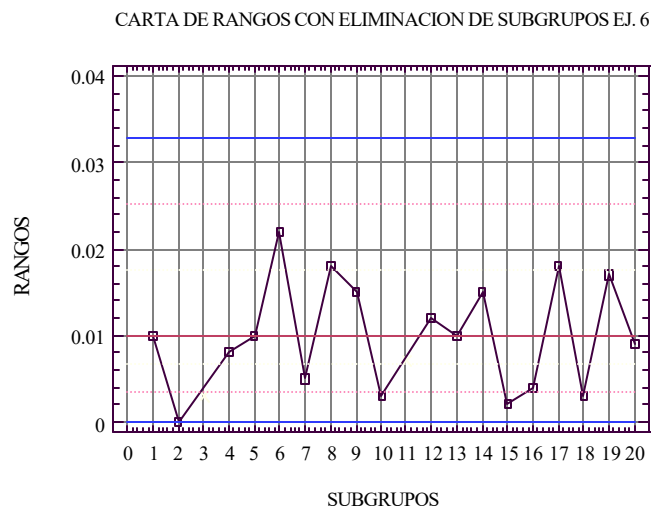
$$LCS : \bar{R}D_4 = 0.01(3.27) = 0.03287$$

$$\text{Linea Central : } \bar{R} = 0.01$$

$$LCI : \bar{R}D_3 = 0.01(0) = 0$$

GRAFICO DE MEDIAS EJEMPLO 6 SIN SUBGRUPOS 3 Y 11





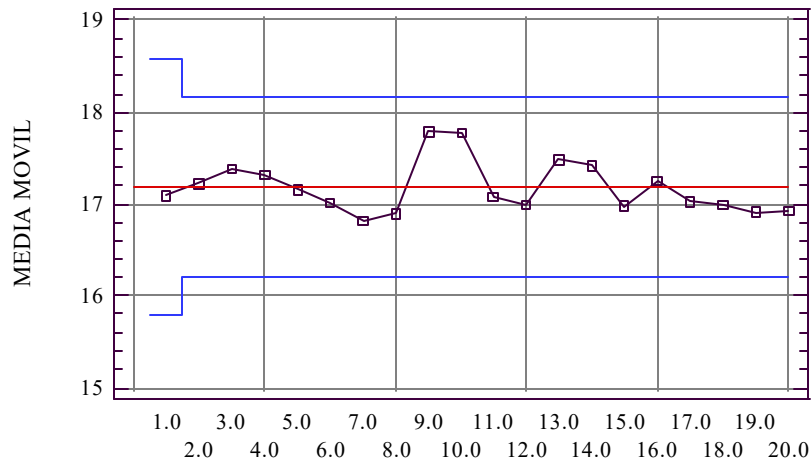
Los nuevos gráficos de \bar{X} y R indican que el proceso se encuentra bajo control y por lo tanto pueden ser usados para llevar el control del proceso de medición.

7. Se quiere construir un diagrama de control de media móvil a partir de los siguientes datos de mediciones individuales, considere $w = 2$ (APHA, 1996, cap. 1).

DATOS:

MUESTRA	VALOR DEL ENSAYO	NUMEROS DE MUESTRA INCLUIDOS	PROMEDIO MOVIL	RANGO MOVIL
1	17.09	-	17.09	-
2	17.35	1-2	17.22	0.26
3	17.40	2-3	17.38	0.05
4	17.23	3-4	17.32	0.17
5	17.09	4-5	17.16	0.14
6	16.94	5-6	17.02	0.15
7	16.68	6-7	16.81	0.26
8	17.11	7-8	16.90	0.43
9	18.47	8-9	17.79	1.36
10	17.08	9-10	17.78	1.39
11	17.08	10-11	17.08	0.00
12	16.92	11-12	17.00	0.16
13	18.03	12-13	17.48	1.11
14	16.81	13-14	17.42	1.22
15	17.15	14-15	16.98	0.34
16	17.34	15-16	17.25	0.19
17	16.71	16-17	17.03	0.63
18	17.28	17-18	17.00	0.57
19	16.54	18-19	16.91	0.74
20	17.30	19-20	16.92	0.76
			$\bar{X} = 17.18$	$\bar{R} = 0.5226$

GRAFICA DE MEDIA MOVIL



El promedio móvil para $w=2$, se calcula como se presenta a continuación:

$$M_t = \frac{x_t + x_{t-1}}{2} \text{ donde "t" en el subíndice indica la posición del promedio móvil.}$$

El rango móvil se calcula como la diferencia entre el mayor y menor de los datos de los subgrupos. El sigma del proceso viene estimado por: $\sigma_{\text{proceso}} = \frac{\bar{R}}{d_2}$, para $n=2$, $d_2 = 1.128$.

Los límites de control se calculan como: $\bar{X} \pm \frac{3\sigma}{\sqrt{nw}}$ para $t \geq 2$, para períodos entre $0 < t < 2$, los límites de control están dados por $\bar{X} \pm \frac{3\sigma}{\sqrt{nt}}$, en la tabla E77.1 se presenta el cálculo de los límites de control.

Según el diagrama de control de la figura, ningún punto cae fuera de los límites de control. Debe observarse que para el período inicial de $t < 2$, los límites de control son más amplios que su valor final en el estado estable. Las medias móviles que se encuentran separadas por menos de w períodos están altamente correlacionadas, y esto complica a menudo la interpretación de los patrones en el diagrama de control.

TABLA E77.1

MUESTRA	VALOR DEL ENSAYO	NUMEROS DE MUESTRA INCLUIDOS	PROMEDIO MOVIL	LIMITES DE CONTROL PARA M_t LCS-LCI
1	17.09	-	17.09	18.57-15.79
2	17.35	1-2	17.22	18.16-16.197
3	17.40	2-3	17.38	18.16-16.197
4	17.23	3-4	17.32	18.16-16.197
5	17.09	4-5	17.16	18.16-16.197
6	16.94	5-6	17.02	18.16-16.197
7	16.68	6-7	16.81	18.16-16.197
8	17.11	7-8	16.90	18.16-16.197
9	18.47	8-9	17.79	18.16-16.197
10	17.08	9-10	17.78	18.16-16.197
11	17.08	10-11	17.08	18.16-16.197
12	16.92	11-12	17.00	18.16-16.197
13	18.03	12-13	17.45	18.16-16.197
14	16.81	13-14	17.42	18.16-16.197
15	17.15	14-15	16.98	18.16-16.197
16	17.34	15-16	17.25	18.16-16.197
17	16.71	16-17	17.03	18.16-16.197
18	17.28	17-18	17.00	18.16-16.197
19	16.54	18-19	16.91	18.16-16.197
20	17.30	19-20	16.92	18.16-16.197

8. Para conteos ambientales usando la técnica de centelleo líquido, se requiere que las mediciones de fondo sean precisas. Cualquier inestabilidad en el instrumento provocará resultados de mediciones de fondo con excesiva variabilidad, lo cual hará imprecisa e inexacta cualquier otra medición. Cook et al. (1989) han estudiado la aplicabilidad de los estadísticos de la distribución de Poisson para observar la variabilidad en los conteos de fondo adquiridos con un equipo de centelleo utilizando analizadores de bajo nivel. A partir de los datos siguientes construya un grafico que permita controlar la estabilidad del equipo. Las muestras utilizadas para obtener los datos fueron muestras de fondo de bajo nivel de excitación. Cada muestra fue contada para diez ciclos y se hicieron diez determinaciones para cada ciclo. Cada ciclo para una muestra particular es considerado como un subgrupo.

DATOS:

CICLO	CONTEO PROMEDIO POR MINUTO (CPM)	DESV. STD DE LOS CONTEOS POR CICLO
1	0.94	0.116
2	0.74	0.095
3	0.81	0.145
4	0.86	0.070
5	0.83	0.145
6	0.86	0.140
7	0.76	0.105
8	0.82	0.110
9	0.80	0.091
10	0.86	0.160

SOLUCION

La carta de \bar{X} valora el nivel promedio del proceso, y la carta de s valora la variabilidad del proceso. Para la carta de s , los límites de control superior e inferior vienen definidos por:

$$LCS = \bar{S} + 3 \frac{S}{C_4} \sqrt{1 - C_4^2} \quad LCI = \bar{S} - 3 \frac{S}{C_4} \sqrt{1 - C_4^2}$$

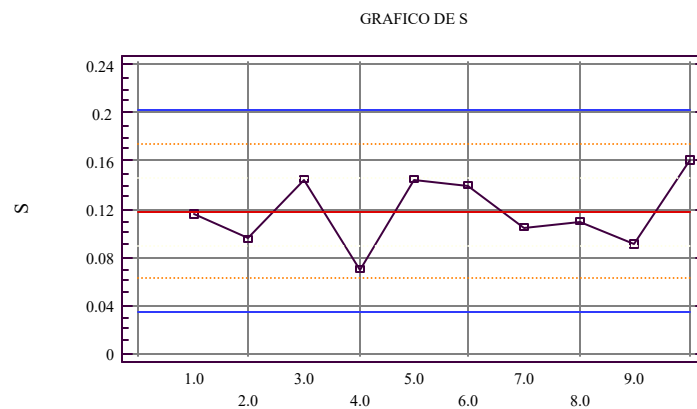
Donde para $n = 10$, $1/C_4 = 1.0281$ y $C_4 = 0.9727$. La desviación estándar del proceso viene dada a partir de:

$$\bar{S} = \frac{1}{m} \sum_{i=1}^m S_i = \frac{1}{10} (1.177) = 0.1177$$

Por tanto los límites de control para el gráfico de desviaciones estándar son:

$$LCS = 0.2019 \quad \text{y} \quad LCI = 0.0334$$

El gráfico se presenta a continuación:



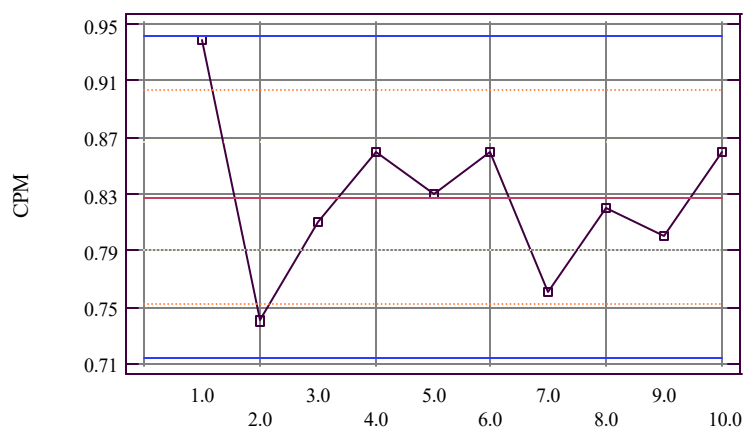
La figura anterior indica que la variabilidad del proceso esta en estado de control dado que ninguno de los puntos cae fuera de los límites de control. La carta de medias es construida como en los ejemplos anteriores, siendo los límites de control calculados a partir de:

$$\bar{X} \pm A_3S$$

De esta forma los límites de control superior e inferior para el diagrama de medias son: ($A_3 = 0.975$, para $n = 10$ y $\bar{X} = 0.7132$)

$$LCS = 0.9428 \text{ y } LCI = 0.828$$

GRAFICA DE MEDIAS



Como puede observarse del grafico de medias el conteo promedio de fondo esta bajo control.

Puede concluirse que existe estabilidad en el equipo de centelleo líquido.

REFERENCIAS BIBLIOGRAFICAS

- **APHA (1996)**. Manual de análisis de aguas y aguas de desecho. U.S.A.: AOAC.
- **Bartés A. P., Martorell Llabrés. X. T., Cintas. P. G. y Fernandez. L. P. (2000)**. Métodos estadísticos. Control y mejora de la calidad. México: Alfaomega Grupo Editor.
- **Bover W.J. et al. (1990)**. Manual on Quality Practices and Guidelines for Exxon Laboratorios. Version 2.0, Report Number CR.2BA.90. U.S.A.: Exxon Laboratories.
- **Box.G.E.P, Hunter.N.G y Hunter. J.S (1989)**. Estadística para investigadores. España: Reverté.
- **Duncan A. J. (1989)**. Control de Calidad y Estadística Industrial. Cuarta edición. México: Editorial Alfaomega.

- **Evans J.R. y Lindsay W.M. (1995).** Administración y Control de la Calidad. México: Grupo Editorial Iberoamérica S.A. de C.V.
- **Garfield F.M.(1990).** Principios de garantía de calidad para laboratorios analíticos. Segunda edición. España: AOAC.
- **Juran J.M. (1990).** Planificación para la calidad. Madrid, España: Ediciones DÍAZ DE SANTOS S.A. DE C.V.
- **Juran J.M y Gryna F. M. (1998).** Manual de control de calidad. (2 tomos). España: Mcgraw-Hill.
- **Kenett R. S. y Zacks S. (2000).** Estadística industrial moderna. Diseño y control de la calidad y la confiabilidad. México: International Thomson Editores.
- **Kuehl. R. O. (2001).** Diseño de experimentos. Principios estadísticos de diseño y análisis de investigación. México:International Thomson Editores.
- **Kume H.(1997).** Herramientas estadísticas básicas para el mejoramiento de la calidad. Colombia: Grupo editorial NORMA.
- **Lawson J., Madrigal J.L. y Erjavec J. (1992).** Estrategias experimentales para el mejoramiento de la calidad en la industria. México: Grupo Editorial Hibernoamérica.
- **Miller J. C. y Miller J. N. (1993).** Estadística para química analítica. Segunda edición. Wilmington, Delaware, E.U.A: Addison-Wesley Iberoamericana, S.A.
- **Montgomery D. C. (1991).** Control estadístico de calidad. México: Grupo Editorial Iberoamérica S.A. de C.V.
- **Pérez.C.(1999).** Control estadístico de calidad. México: ALFAOMEGA grupo editor S.A. de C.V.
- **Pérez.C.(2002).** Estadística aplicada a través de Excel. México: ALFAOMEGA grupo editor S.A. de C.V.
- **Peña P.E. et al. (1994).** Curso de Gerencia del control de calidad. El Salvador: Universidad Politécnica de Catalunya (UPC).
- **Prichard E. et al. (1995).** Quality in the Analytical Chemistry Laboratory. England: John Wiley & Sons Ltd.

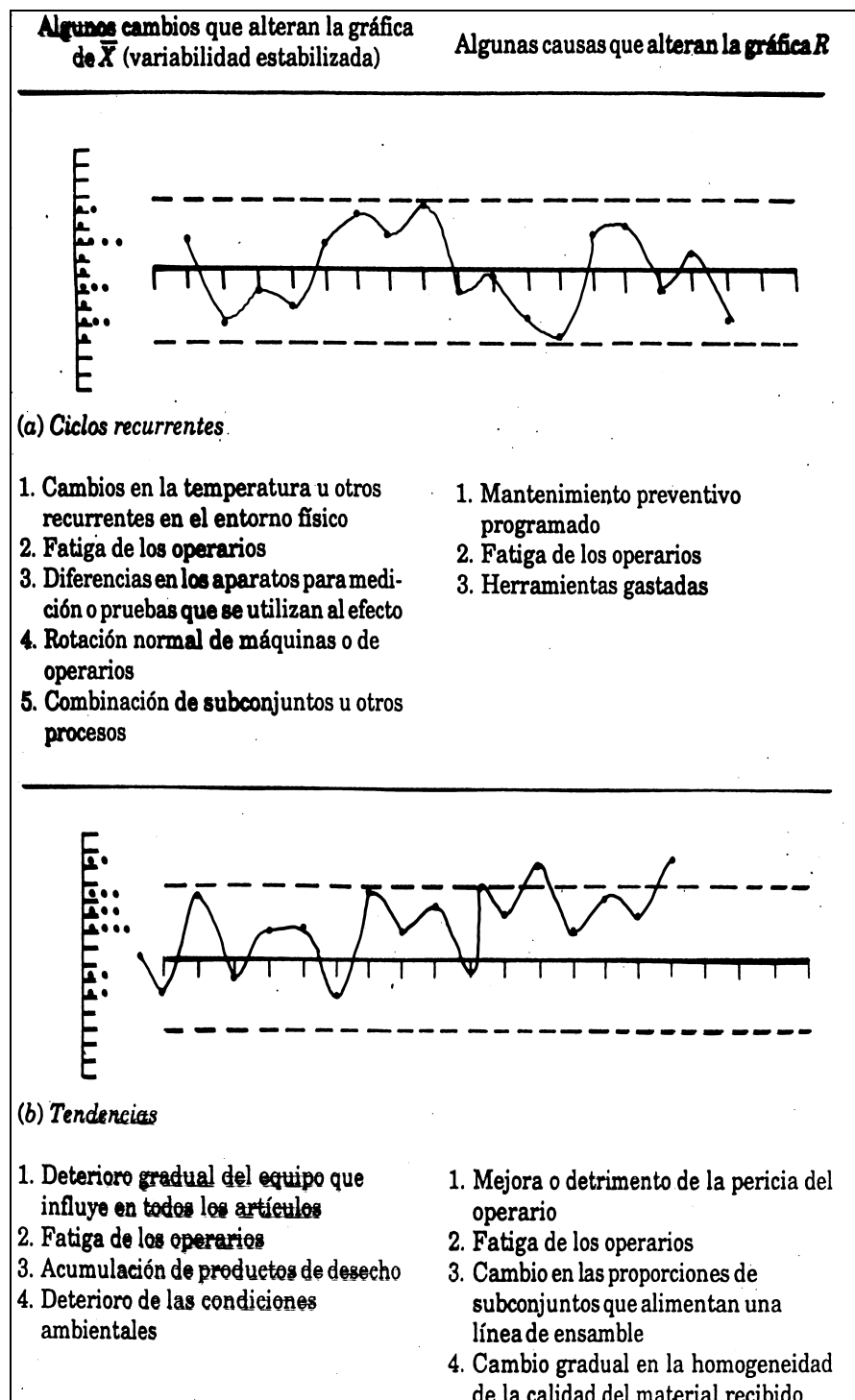
**ANEXO A7. FACTORES PARA CONSTRUIR DIAGRAMAS DE CONTROL PARA
VARIABLES**

(PEÑA P.E. et al., 1994, cap. 10)

n	A ₁	A ₂	B ₃	B ₄	C ₂	d ₂	D ₃	D ₄	E ₁	E ₂
2	3.759	1.880	0	3.267	0.564	1.128	0	3.268	5.318	2.660
3	2.394	1.023	0	2.568	0.724	1.693	0	2.574	4.146	1.772
4	1.880	0.729	0	2.266	0.798	2.059	0	2.282	3.760	1.457
5	1.596	0.577	0	2.089	0.841	2.326	0	2.114	3.568	1.290
6	1.410	0.483	0.030	1.970	0.869	2.534	0	2.004	3.454	1.184
7	1.277	0.419	0.118	1.882	0.888	2.704	0.076	1.924	3.378	1.109
8	1.175	0.373	0.185	1.815	0.903	2.847	0.136	1.864	3.323	1.054
9	1.094	0.337	0.239	1.761	0.914	2.970	0.184	1.816	3.283	1.010
10	1.028	0.308	0.284	1.716	0.923	3.078	0.223	1.777	3.251	0.975
11	0.973	0.285	0.321	1.679	0.930	3.173	0.256	1.744	3.226	0.946
12	0.925	0.266	0.354	1.646	0.935	3.258	0.284	1.717	3.205	0.921
13	0.884	0.249	0.382	1.618	0.941	3.336	0.308	1.692	3.188	0.899
14	0.848	0.235	0.406	1.594	0.945	3.407	0.324	1.671	3.174	0.881
15	0.817	0.223	0.428	1.572	0.949	3.472	0.348	1.652	3.161	0.864
16	0.788	0.212	0.448	1.552	0.952	3.532	-	-	-	-
17	0.762	0.203	0.466	1.534	0.955	3.588	-	-	-	-
18	0.738	0.194	0.482	1.518	0.958	3.640	-	-	-	-
19	0.717	0.187	0.497	1.503	0.960	3.689	-	-	-	-
20	0.698	0.180	0.510	1.490	0.962	3.735	-	-	-	-

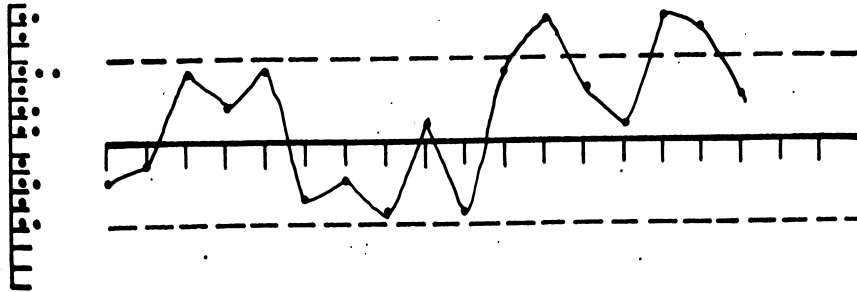
ANEXO B7 PLANTILLA PARA LA ELABORACION DE GRAFICAS \bar{X} -R

ANEXO C7 | L USTRACION DE PRUEBAS DE CAUSAS ASI GNABLES
(DUNCAN A.J., 1989)



Algunos cambios que alteran la gráfica de \bar{X} (variabilidad estabilizada)

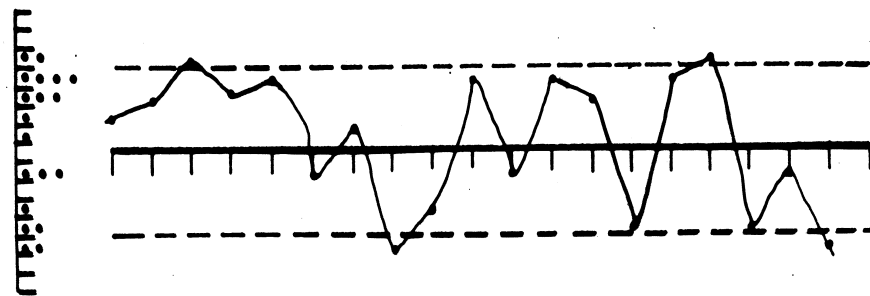
Algunas causas que alteran la gráfica R



(c) Cambios bruscos en el grado del proceso

1. Cambio en las proporciones de material o subconjuntos que vienen de diversas fuentes
2. Operario o máquina nuevos
3. Modificación del método o proceso de producción
4. Cambio en el aparato o método para inspección

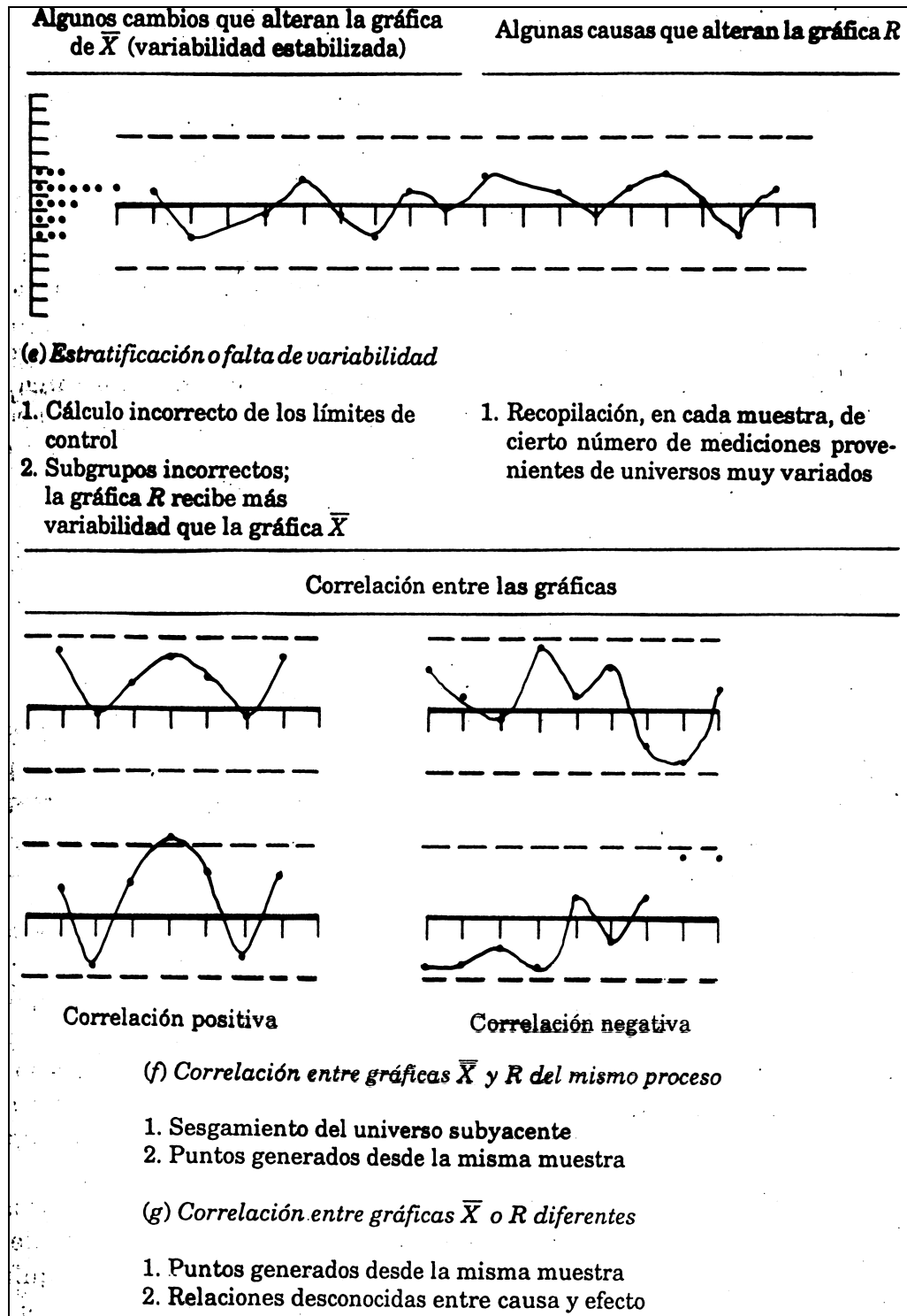
1. Cambio en el material
2. Cambio en el método
3. Cambio del operario



(d) Alta proporción de puntos cerca o fuera de los límites

1. Exceso de control
2. Grandes diferencias sistemáticas en la calidad del material
3. Grandes diferencias sistemáticas en el método o equipo para prueba
4. Llevar control de dos o más procesos en la misma gráfica

1. Mezcla de materiales de calidades notoriamente diferentes
2. Empleo de una sola gráfica de R para diversos operarios
3. Datos de procesos en diferentes condiciones, trazados en la misma gráfica



ACUSUMA” (PRAT B. A. et al., 2000, cap. 11)

A continuación se darán las pautas a seguir en la construcción de los gráficos acusuma. Para ello se referirá el caso particular en el que se quiere controlar la media de un proceso que, en el momento de la implantación del acusuma, está centrada en el valor nominal $\mu = T$.

1. Tomar muestras de tamaño n a intervalos de tiempo equidistantes y obtener la media y el recorrido de las mismas (para medidas individuales la media será cada observación y el rango será el rango móvil).
2. Calcular en cada instante la suma acumulada de las discrepancias de los valores obtenidos en 1, con el valor nominal $\mu = T$

$$S_i = \sum_{i=1}^k (\bar{x}_i - T)$$

3. Obtener una estimación de $\sigma_e = s_e = \frac{s}{\sqrt{n}}$. Esta puede ser obtenida en función del recorrido medio de un número suficientemente grande de muestras como:

$$s_e = \frac{\bar{R}/d_2}{\sqrt{n}}$$

4. Determinar el nivel de probabilidad α , o riesgo que se está dispuesto a asumir en la toma de decisiones, en cada uno de los lados del gráfico.
5. Definir el factor de escala del gráfico. Es recomendable que
1 unidad horizontal = $2s_c$ unidad vertical = A
6. Determinar el menor cambio D en la media que se quiere detectar y calcular

$$\delta = D/s_e$$

7. Obtener, a partir de δ , la distancia principal d

$$d = \frac{2 \ln(1 - \beta/\alpha)}{\delta^2} \quad \text{o, si } \beta \text{ es pequeña, } d = \frac{2 \ln(\alpha)}{\delta^2}$$

Donde β es la probabilidad de no detectar un cambio de D unidades en la media, y el ángulo θ

$$\theta = \arctg\left(\frac{D}{2A}\right)$$

8. A partir de d y θ construir la plantilla que define los límites de control, tal como se muestra en la figura D7-1:

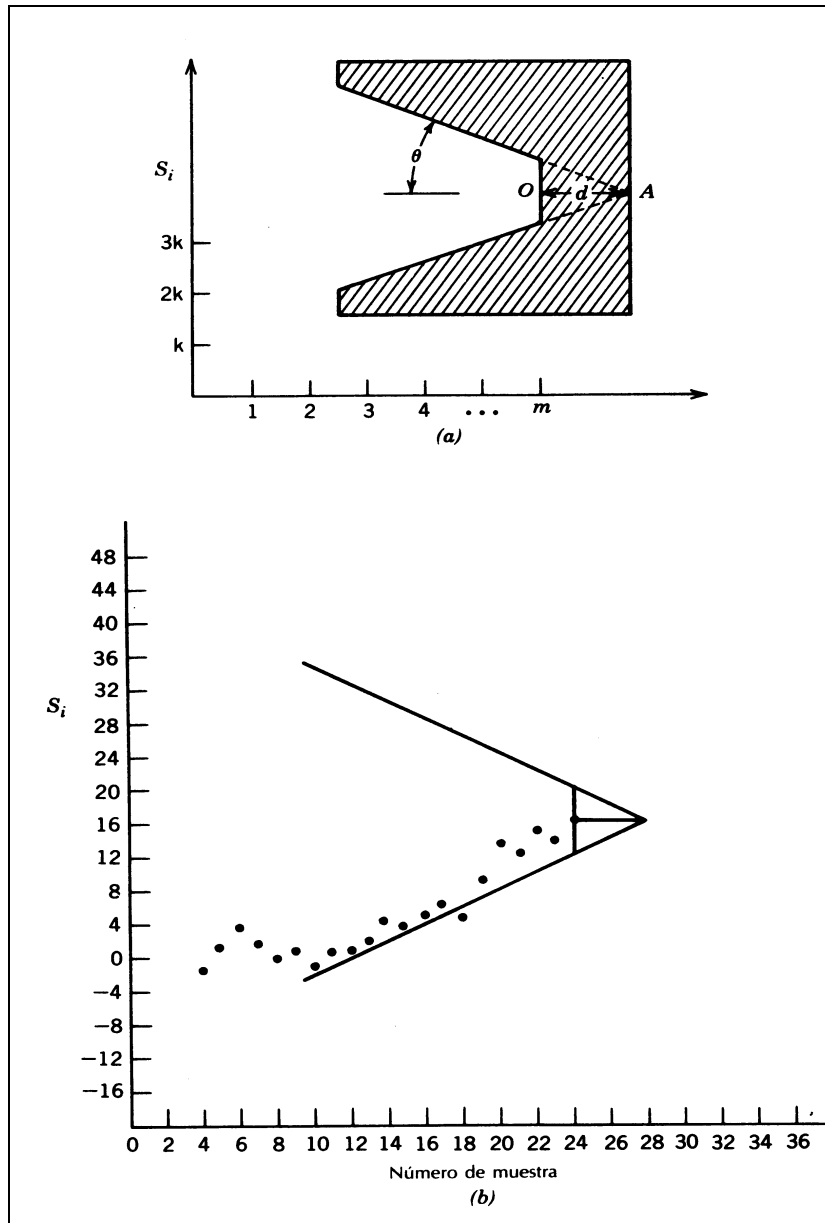


FIGURA D7-1 ESQUEMA DEL GRAFICO ACUSUMA

Para la interpretación del gráfico acusuma, el punto O en la plantilla se ha de colocar en el último valor de S_i obtenido. Si alguno de los puntos anteriores queda cubierto por la plantilla, se interpreta que alguna causa asignable ha entrado en el proceso y ha provocado un cambio en la media superior a D unidades. Además se ha de tener en cuenta que:

- El primer punto cubierto por la plantilla muestra el momento en el que el proceso ha dejado de estar bajo control;
- Si los puntos están cubiertos por encima de la plantilla es que la media del proceso ha disminuido;
- Si los puntos están cubiertos por debajo de la plantilla es que la media ha aumentado.

ANEXO E7 “PROCEDIMIENTO DE CONSTRUCCION DE DIAGRAMAS DE CONTROL BIVARIANTES” (MONTGOMERY, 1991, cap. 7)

Para formalizar un proceso de control bivalente supongamos que dos características de calidad X_1 y X_2 tienen una distribución normal divariada siendo \bar{X}_1 y \bar{X}_2 los valores medios de las características de calidad, y supóngase que se estiman las varianzas de X_1 y X_2 mediante las varianzas maestras S_1^2 y S_2^2 , respectivamente. La covarianza entre X_1 y X_2 es una medida de la dependencia de las dos variables, y se denota por S_{12} . Si \bar{X}_1 y \bar{X}_2 son las medias muestrales de las dos características, calculadas a partir de un subgrupo de tamaño n , entonces el estadístico:

$$T^2 = \frac{n}{S_1^2 S_2^2 - S_{12}^2} \left[S_2^2 (\bar{X}_1 - \bar{X}_1)^2 + S_1^2 (\bar{X}_2 - \bar{X}_2)^2 - 2S_{12} (\bar{X}_1 - \bar{X}_1)(\bar{X}_2 - \bar{X}_2) \right]$$

Se distribuirá como una distribución T^2 de Hotelling, con 2 y $n - 1$ grados de libertad. Si $T^2 > T_{\alpha,2,n-1}^2$, entonces por lo menos una de las variables está fuera de los límites de control, donde $T_{\alpha,2,n-1}^2$ es el punto porcentual superior de la distribución T^2 de Hotelling, con 2 y $n-1$ grados de libertad para un nivel o error de tipo I dado e igual a α . Es importante tener presente que la T^2 de Hotelling es aproximable por la F de Fisher Snedecor. Se pueden obtener los puntos porcentuales de la T^2 a partir de la distribución F, mediante la relación:

$$T_{\alpha,p,n-1}^2 = \frac{p(n-1)}{n-p} F_{\alpha,p,n-p}$$

donde p es el número de variables correlacionadas y n es el número de pares de puntos. Para representar gráficamente este procedimiento de control bivalente, considérese el caso en que \bar{X}_1 y \bar{X}_2 son independientes, es decir, $S_{12} = 0$. Es decir si $S_{12} = 0$, la ecuación que expresa T^2 define una elipse centrada en (\bar{X}_1, \bar{X}_2) con los ejes principales paralelos a los ejes \bar{X}_1 y \bar{X}_2 . Hacer T^2 igual a $T_{\alpha,2,n-1}^2$ implica que un par de medias maestras observadas (\bar{X}_1, \bar{X}_2) que se encuentran dentro de la elipse indican un estado de control estadístico, mientras un par de medias maestras observadas que caen fuera de la elipse muestran que el proceso está fuera de control. La figura obtenida se denomina a menudo **elipse de control**. Si las dos características son dependientes entonces $S_{12} \neq 0$, y la elipse de control correspondiente aparecerá como en la figura E7-1. Obsérvese que los ejes principales de la elipse ya no son paralelos a los ejes \bar{X}_1 y \bar{X}_2 cuando las dos variables son dependientes.

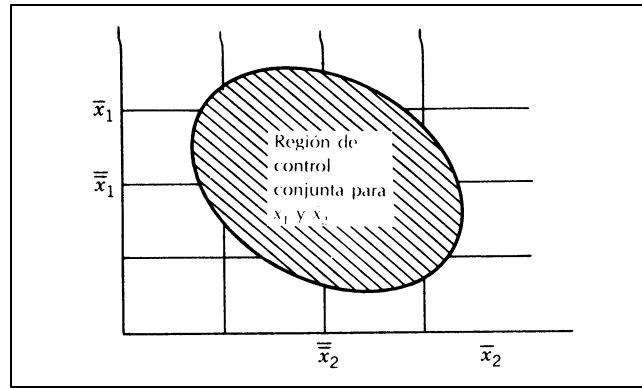


FIGURA E7-2 ELIPSE DE CONTROL

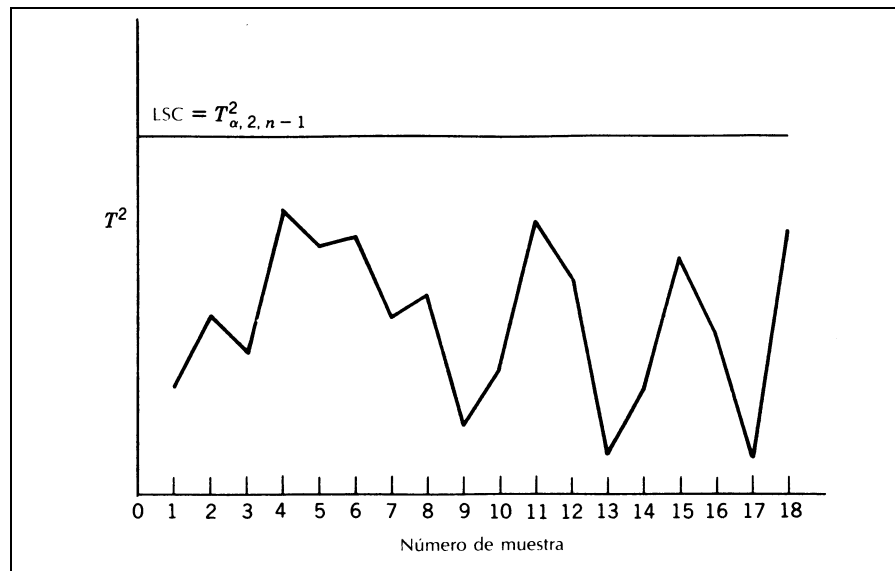


FIGURA E7-2 DIAGRAMA DE CONTROL T^2 HOTELLING PARA $P = 2$ CARACTERÍSTICAS DE CONTROL

CAPITULO VIII

MÉTODOS NO PARAMÉTRICOS PARA ANÁLISIS QUÍMICO

CONTENIDO

	Pág.
8.0 Métodos estadísticos no paramétricos	8-2
8.1 Prueba de la mediana	8-4
8.2 La prueba de signos	8-7
8.3 Prueba de rachas de Wald- Wolfwitz	8-9
8.4 Prueba de rangos y signos de Wilcoxon	8-10
8.5 Prueba de suma de rangos de Wilcoxon	8-12
8.6 Prueba de Mann- Whitney	8-14
8.7 Prueba de Kruskal Wallis	8-16
8.8 Problemas de aplicación de métodos no parametricos	8-17
REFERENCIAS BIBLIOGRAFICAS	8-26
ANEXO A8. Métodos no paramétricos de correlación lineal simple y bondad del ajuste	8-27
ANEXO B8. Valores críticos de la prueba de rachas de Wald- Wolfwitz	8-31
ANEXO C8. Valores críticos de la prueba de rangos y signos de Wilcoxon	8-32
ANEXO D8. Valores críticos de la prueba de suma de rangos de Wilcoxon- Mann-Withney	8-33

8.0 METODOS ESTADÍSTICOS NO PARAMÉTRICOS

Los métodos estadísticos paramétricos suponen que los datos que se analizan siguen una distribución normal (Gaussiana). La validez de esta hipótesis se basa en el teorema central del límite, que postula que la distribución muestral de la media puede ser aproximadamente normal aunque la población de referencia tenga una distribución muy diferente. La aproximación mejora a medida que el tamaño de la muestra aumenta.

Con frecuencia se presentan a los químicos analistas situaciones donde no pueden asumirse los supuestos requeridos por desconocerse la distribución de la variable estudiada, o bien, por ser la muestra muy pequeña de manera que incluso el teorema central del límite sería de escasa relevancia. O bien, aunque se conozca la distribución de la variable y sea válido teóricamente hacer supuestos sobre la distribución y sus parámetros, no es razonable utilizar una prueba paramétrica ya que hay poca certeza de que se cumplan en este conjunto de datos. También puede ocurrir que la variable no sea continua por lo que no se cumplen las restricciones establecidas para las pruebas paramétricas.

En cualquiera de los casos anteriores hay que buscar técnicas alternativas que permitan darle solución a estas situaciones de forma eficiente. Surge así la necesidad de desarrollar una serie de técnicas estadísticas que tengan un mínimo de restricciones. A estas técnicas se les conoce como: Métodos no paramétricos.

Según WONNACOTT (1973) existen dos indicaciones para preferir las pruebas no paramétricas:

1. Cuando la prueba clásica correspondiente no es válida.
2. En aplicaciones en donde la prueba clásica es razonablemente válida, pero un estimador no paramétrico puede ser más eficiente.

Entre las ventajas del uso de métodos no paramétricos se encuentran las siguientes (PRIA, 2001, cap. 1):

1. Tienen mayor eficiencia que los métodos paramétricos en distribuciones asimétricas, o sea cuando hay valores atípicos o datos aberrantes.
2. Tienen validez en el sentido de que su nivel de confiabilidad es realmente el especificado en la mayoría de las pruebas.

3. Generalmente son de cómputo mucho más fácil que las técnicas de las estadísticas clásicas.
4. Son aplicables en situaciones donde los procedimientos clásicos no son aplicables.
5. Aún cuando se cumplan los requisitos para realizar una prueba paramétrica, si la muestra es pequeña la eficiencia relativa de la prueba no paramétrica es alta.

Entre las desventajas de la aplicación de los métodos no paramétricos se citan las siguientes (PRIA, 2001, cap. 1):

1. Fundamentalmente cuando las muestras son muy grandes las pruebas no paramétricas tienen una eficiencia relativa baja con relación a las paramétricas cuando se cumplen los supuestos.
2. Las hipótesis que se plantean en las pruebas no paramétricas son menos precisas, lo que hace que la interpretación de los resultados sea más ambigua.
3. Su aplicación en muestras grandes se hace muy laboriosa.
4. Para un problema particular pueden existir varias pruebas, por lo que en ocasiones se hace difícil seleccionar la mejor.

En general, este tipo de pruebas deben utilizarse cuando no se cumplan los supuestos para hacer una prueba paramétrica y cuando se duda sobre el cumplimiento de los supuestos en una muestra pequeña. La principal ventaja de las pruebas no paramétricas consiste en que pueden efectuarse inferencias exactas cuando las suposiciones fundamentales de los métodos estándar no pueden cumplirse en su totalidad; su principal desventaja radica en que exigen demasiada información y tienen una eficiencia menor cuando todas las suposiciones se satisfacen. Sin embargo, si se afirma que la eficiencia de cierto método no paramétrico es del 80%, se debe entender en realidad que es relativamente valioso, dado que la eficiencia del correspondiente método "estándar" es poco menos del 100% si todas las suposiciones no se cumplen con exactitud. En general es verdad que *cuanto menos se supone, tanto menos se puede inferir de un conjunto de datos*, pero también es cierto que *cuanto menos se supone, tanto más se amplía la aplicabilidad de un método* (FREUD y WALPOLE, 1990, cap. 16).

Es posible comparar la confiabilidad del uso de una prueba no paramétrica con respecto al uso de una prueba paramétrica homóloga a través del concepto de *potencia-eficiencia* de

una prueba, que en general, se refiere al aumento porcentual del tamaño de muestra que se necesita para obtener los mismos resultados con una prueba X, no tan poderosa como otra que es la prueba Y, que es la más potente que se conoce para resolver un problema específico cuando se cumplen los supuestos para efectuar la misma. La *potencia-eficiencia* de la prueba X puede calcularse de la forma siguiente (PRIA, 2001, cap. 1):

$$\text{Potencia – Eficiencia de la prueba X} = \frac{N_Y}{N_X} \times 100$$

En donde N_Y es el tamaño de muestra requerido para la prueba Y, y, N_X el tamaño de muestra requerido para la prueba X.

Existe una gran variedad de pruebas no paramétricas tanto para el análisis de variables cualitativas como cuantitativas. En el cuadro 8.1 se presenta un listado de dichas pruebas y de su aplicación.

En las secciones de la 8.1 a la 8.7 se describen con detalle las pruebas no paramétricas asociadas a variables de tipo cuantitativo de mayor aplicación en análisis químico.

Además de las pruebas expuestas en la tabla 8.1, existen pruebas no paramétricas para correlación y análisis de regresión, cuyas bases se presentan en el anexo A8.

8.1 PRUEBA DE LA MEDIANA

Esta prueba permite determinar si dos muestras independientes difieren con relación a sus medianas, o sea permite determinar si dos muestras independientes provienen de poblaciones con la misma mediana siempre que la variable esté al menos en escala ordinal (PRIA, 2001, cap. 1). Existe una generalización de esta prueba que permite la comparación de las medianas de tres o más muestras independientes que no será objeto de estudio en esta sección.

La hipótesis a probar es:

$$\begin{aligned} H_0: \tilde{X}_1 &= \tilde{X}_2 \\ H_1: \tilde{X}_1 &\neq \tilde{X}_2 \end{aligned}$$

En donde \tilde{X}_1 y \tilde{X}_2 representan las medianas de las dos muestras que se están comparando.

CUADRO 8.1 DESCRIPCIÓN DE PRUEBAS NO PARAMÉTRICAS

PRUEBA	TIPO DE VARIABLE	APLICACIÓN
Prueba χ^2 de independencia	Cualitativa	Esta prueba permite medir la significación de la asociación entre dos variables de clasificación.
Prueba χ^2 de homogeneidad	Cualitativa	Cuando se tienen varias muestras y se desea determinar si son homogéneas con relación a la distribución en las mismas de una variable cualitativa.
Prueba Kolmogorov-Smirnov	Cuantitativa	Es una prueba no paramétrica que se utiliza para diferencias entre distribuciones acumuladas, es, pues, una prueba de bondad de ajuste.
Prueba de la mediana	Cuantitativa	Esta prueba permite determinar si dos muestras independientes difieren con relación a sus medianas, o sea permite determinar si dos muestras independientes provienen de poblaciones con la misma mediana.
Prueba de signos	Cuantitativa	Esta prueba permite la comparación de la mediana de una serie de datos con un valor especificado. También permite indicar la existencia de tendencias.
Prueba de rangos y signos de Wilcoxon	Cuantitativa	Permite probar la aleatoriedad de una secuencia de datos. También permite probar la simetría de una distribución. Otra aplicación de esta prueba es comparar la distribución de una serie de datos con un valor especificado.
Prueba de la suma de rangos de Wilcoxon	Cuantitativa	Constituye la base para el resto de pruebas que utilizan rangos y permite determinar si dos muestras proceden de la misma distribución, las muestras deben de ser del mismo tamaño y no necesariamente independientes.
Prueba de U-Man Whitney	Cuantitativa	Esta prueba se utiliza para resolver el mismo caso que resuelve la prueba de suma de rangos de Wilcoxon con muestras no necesariamente del mismo tamaño.
Prueba de Kruskal Wallis	Cuantitativa	Es una generalización de la prueba de suma de rangos de Wilcoxon, permitiendo comparar más de dos muestras con el propósito de conocer si proceden de la misma población o si hay diferencias entre las medidas de tendencia central de más de dos poblaciones.

Debe señalarse que en el caso de que las distribuciones de ambas poblaciones sean simétricas, la mediana debe coincidir con la media aritmética, convirtiéndose ésta en una prueba de comparación de dos medias también (PRIA, 2001, cap. 1).

Como se está asumiendo que ambas muestras provienen de poblaciones con la misma mediana, se determina el valor de la mediana general, o sea se combinan los valores de ambos grupos para determinar el valor de la mediana.

Una vez determinado éste valor se determina en cada una de las muestras cuántas observaciones tienen valores mayores que la mediana general y cuántas tienen valores inferiores a la misma.

Según SIEGEL (1972) si hay valores que coincidan con los de la *mediana general* en las muestras estudiadas, se procederá de la siguiente forma:

1. Si la muestra es grande y un número pequeño de observaciones coincide con este valor pueden excluirse del análisis estas observaciones.
2. Si la muestra no es grande o si el número de observaciones que coincide con este valor es grande entonces pueden contarse estos valores junto con los valores menores que la mediana.

La información se resumirá en una tabla de contingencia de la forma siguiente (PRIA, 2001, cap. 1):

Valores/ Mediana	Muestra 1	Muestra 2	Total
Observaciones > Mediana General	a	b	a+b
Observaciones ≤ Mediana General	c	d	c+d
Total	a+ c	b+ d	N

Para la prueba de hipótesis se necesita un estadígrafo con una distribución conocida para determinar si existen diferencias significativas entre las medianas de ambas poblaciones. Para ello se utilizará la expresión (PRIA, 2001, cap. 1):

$$\chi_o^2 = \frac{n \left(|ad - bc| - \frac{n}{2} \right)^2}{(a + b)(c + d)(a + c)(b + d)}$$

Este estadístico bajo el supuesto de que no existen diferencias entre las medianas de ambas poblaciones se distribuye χ^2 con un grado de libertad.

Si $\chi_o^2 > \chi_{tabulada}^2$ con un grado de libertad para un nivel de significación determinado, se rechaza la hipótesis nula.

Al usar el estadígrafo χ^2 hay que tener en cuenta que se deben cumplir las restricciones para la prueba χ^2 . En este caso, como hay una tabla de contingencia de 2 x 2, debe cumplirse que todos los valores esperados sean mayores o iguales a 5, en caso de no cumplirse esta restricción se deberá utilizar la prueba de las probabilidades exactas de Fisher (PRIA, 2001, cap. 1).

Según SIEGEL y MOOD (1954) se ha demostrado que si la prueba de la mediana se aplica a datos que pueden analizarse adecuadamente por una prueba paramétrica más poderosa, la prueba t en éste caso, su potencia-eficiencia sería de cerca de un 95% cuando $n_1 + n_2 = 6$. Este valor iría disminuyendo a medida que aumenta el tamaño muestral llegando a tener una eficiencia asintótica eventual de 63%.

8.2 LA PRUEBA DE SIGNOS

La prueba de signos es uno de los métodos no paramétricos más simples. La prueba t supone que los datos se distribuyen normalmente. La prueba del signo prescinde de tal hipótesis y es mucho más fácil de realizar. Se puede utilizar de diferentes formas, la más simple se describe a continuación:

Para probar la hipótesis nula $\mu = \mu_o$ contra una alternativa apropiada, basándose en una muestra aleatoria de tamaño n , se reemplaza cada valor muestral mayor que μ_o por un signo más y cada valor muestral menor que μ_o por un signo menos (MILLER y FREUND, 1986, cap. 10).

Se ignoran por completo aquellos valores que son iguales a μ_o . Para contrastar si la preponderancia de signos menos, es significativa se utiliza la ley de la binomial acumulada. Esta ley establece que la probabilidad de que aparezcan r signos menos entre n signos está dada por (MILLER y MILLER, 1993, cap. 6):

$$p(r) = nC_r p^r q^{n-r}$$

donde:

nC_r : indica el número de combinaciones de r elementos de un total de n elementos.

p : es la probabilidad de que aparezca un signo menos en uno de los resultados.

q : es la probabilidad de que no aparezca un signo menos en uno de los resultados, es decir, $q = 1 - p$.

Si la probabilidad experimental es menor que un nivel de significación α , la hipótesis nula debe rechazarse. Es decir, existe evidencia como para rechazar que los datos proceden de una población con $\mu = \mu_0$ (MILLER y MILLER, 1993, cap. 6).

La prueba de signos puede utilizarse también como una alternativa no paramétrica de la prueba t por parejas para comparar dos grupos de resultados para las mismas muestras. Así, si se examinan n muestras con cada uno de los dos métodos, A y B se puede contrastar si los dos proporcionan lecturas significativamente diferentes (MILLER y MILLER, 1993, cap. 6).

Calculando para cada muestra la diferencia en los resultados, es decir (MILLER y MILLER, 1993, cap. 6):

$$[(\text{resultado obtenido por el método A}) - (\text{resultado obtenido por el método B})]$$

La hipótesis nula será que los dos métodos no proporcionan resultados significativamente diferentes. En la práctica esto significará que la probabilidad de obtener un signo más (o un signo menos) es 0.5. La probabilidad o frecuencia de un número signos positivos sigue una distribución binomial de parámetros $n, p = 0.5$ (MILLER y MILLER, 1993, cap. 6).

El resumen de la prueba del signo para comparar dos poblaciones se presenta en el cuadro 8.2.

Cuando el número de pares en un experimento de diferencias por parejas es grande, se puede aproximar la distribución de probabilidad para x , el número de veces que x_A es mayor que x_B , por una distribución normal. Por lo tanto, con $p = P(X_A > X_B)$ se puede probar $H_0: p = 0.5$ utilizando la prueba para una proporción binomial. (Ver cuadro 3.14, capítulo 3).

Otro uso de la prueba de signos es indicar una tendencia. En esta prueba al primer resultado del primer grupo se le resta el primer resultado del segundo grupo, al segundo

resultado del primer grupo se le resta el segundo resultado del segundo grupo y así sucesivamente. Se obtiene de esta forma una serie de signos asociados a las diferencias obtenidas e ignorando los ceros se calcula la probabilidad binomial de obtener r pares de signos iguales. La hipótesis nula de que no existe ninguna tendencia en los resultados, se rechaza si la probabilidad calculada es menor que el nivel de significación α de la prueba de dos colas (MILLER y MILLER, 1993, cap. 6).

CUADRO 8.2 RESUMEN DE LA PRUEBA DEL SIGNO APLICADA AL ANÁLISIS DE DOS MUESTRAS

1. Hipótesis nula: $H_0: P(X_A - X_B) = 0.5$	
2. Hipótesis alternativa	
Prueba de una cola	Prueba de dos colas
$H_1: P(X_A - X_B) > 0.5$ (o bien $P(X_A - X_B) < 0.5$)	$H_1: P(X_A - X_B) \neq 0.5$
3. Estadístico de contraste: x = número de pares de observaciones para los cuales X_A excede a X_B .	
4. Región de rechazo:	
Prueba de una cola	Prueba de dos colas
<i>Para:</i> $P(X_A - X_B) > 0.5$	Rechace H_0 para valores muy grandes o muy pequeño de x .
Rechace H_0 para valores muy grandes de x .	Basta observar si la frecuencia observada de signos positivos se desvía significativamente de la esperada $n/2$, en cuyo caso se rechaza H_0 .
<i>Para:</i> $P(X_A - X_B) < 0.5$	
Rechace H_0 para valores muy pequeños de x .	
<i>Suposiciones: Las observaciones x_A y x_B se seleccionan aleatoria e independientemente por pares. Cuando los datos de las observaciones de x_A son iguales a las de x_B se eliminan y con esto se reduce el número de pares n.</i>	

8.3 PRUEBA DE RACHAS DE WALD – WOLFWITZ.

Las rachas son una secuencia de signos positivos seguida de una de signos negativos, y luego otra de signos positivos. En algunos casos podría tener interés no sólo investigar si

las observaciones generan signos positivos o negativos, si no también si éstos aparecen en una secuencia aleatoria.

En el caso de ajuste de curvas, una secuencia no aleatoria de signos positivos y negativos conducirá a un número más pequeño de rachas que una aleatoria. El método de Wald – Wolfwitz prueba si el número de rachas es suficientemente pequeño para que se rechace la hipótesis nula de una distribución aleatoria de los signos (MILLER y MILLER, 1993, cap. 6).

El número de rachas en los datos experimentales se compara con los números de la tabla del anexo B8, de “prueba de rachas de Wald – Wolfwitz”, que se presenta para un nivel de significación α . En esta tabla los valores de N representan el número de signos positivos y M es el número de signos negativos. Si el número experimental de rachas es más pequeño que el valor tabulado, entonces se puede rechazar la hipótesis nula.

En relación a la prueba de Wald-Wolfwitz, existen otros puntos de interés, por ejemplo, permite comparar la mediana de una serie de datos con un valor especificado, al crear una serie de signos a través de la comparación de cada valor de la serie con el valor especificado y asignando un signo negativo a los datos inferiores al valor especificado y un signo positivo a los superiores a dicho valor. Ignorando el número de ceros, se determina el número de rachas obtenidas y se compara su valor con el propuesto en la tabla del anexo B8.

También cabe resaltar que se puede encontrar números inusualmente grandes de rachas cortas, así como también números inusualmente pequeños de rachas largas. Así por ejemplo, como en el caso de 6 signos positivos y 6 signos negativos en el orden + - + - + - + - + - + -, se podría sospechar que existe una secuencia no aleatoria. La tabla muestra que, con $N = M = 6$, un total de 11 ó 12 rachas indica que la hipótesis nula de un orden aleatorio se debería rechazar y sospechar una cierta periodicidad en los datos (MILLER y MILLER, 1993, cap. 6).

8.4 PRUEBA DE RANGOS Y SIGNOS DE WILCOXON

La importancia de la prueba de signos radica en los supuestos mínimos que se hacen sobre los datos experimentales. No se supone que la población de la cual se toma la muestra sea normal, ni incluso que sea simétrica. La única información a priori necesaria es el valor de la mediana. Una cierta desventaja de la prueba de signos es que no utiliza

toda la información disponible. Sólo es necesario saber si una medición individual es más grande o más pequeña que la mediana, y la magnitud de esta desviación no se utiliza en absoluto (MILLER y MILLER, 1993, cap. 6).

En muchos casos un analista tendrá suficientes razones para creer que sus mediciones se distribuyen simétricamente pero no desea suponer que siguen una distribución normal. Este supuesto de datos simétricos, y la consecuencia de que la media y la mediana de la población sean iguales, permite desarrollar pruebas de significación más potentes. Wilcoxon contribuyó con importantes avances al respecto, y su prueba de rangos y signos tienen varias aplicaciones (MILLER y MILLER, 1993, cap. 6).

La comparación de la distribución de una serie de datos con un valor de referencia se realiza mediante la obtención de las diferencias entre cada dato de la muestra y el valor de referencia (conservando los signos). Los valores absolutos de estas diferencias se ordenan posteriormente de menor a mayor y a continuación se incorporan sus signos, los números entonces se jerarquizan, en este proceso se mantienen sus signos pero se les asignan números que indican su orden o rango.

Luego, asignando con X a la suma de los rangos positivos y con Y a la suma de rangos negativos, se selecciona la menor de estas cifras (X ó Y) y se toma como el estadístico de contraste. El teorema binomial dará la probabilidad de que aparezca este número. Si los datos provienen de una población con una mediana igual al valor especificado, se esperaría que la suma de rangos positivos y negativos sea aproximadamente igual.

La probabilidad de que aparezca una suma concreta está dada en la tabla que se presenta en el anexo C8. En esta prueba se rechaza la hipótesis nula si el valor tabulado es menor o igual que el valor experimental, es decir, situación opuesta de la observada en la mayoría de las pruebas de significación.

Una ventaja importante de la prueba de rangos y signos es que también se puede utilizar para datos por parejas, ya que las diferencias entre los datos de las dos series se pueden transformar en el tipo de datos como en el caso anterior. De esta manera se puede utilizar este método no paramétrico como una alternativa a la prueba t por parejas (MILLER y MILLER, 1993, cap. 6).

Si no hay diferencia entre las dos series de datos, entonces se esperaría que las diferencias entre los resultados para cada muestra, [(resultado muestra 1) – (resultado muestra 2)] deberán distribuirse en torno a cero. Cuando hay posiciones empatadas, el problema se resuelve asignando posiciones promedio a los valores empatados, con signos

adecuados. Si la ordenación es correcta, al calcular la suma de todos los valores sin tomar en cuenta el signo, debe ser la misma que la suma de los ocho primeros números naturales (MILLER y MILLER, 1993, cap. 6).

Si la suma de los valores positivos es X y la de los valores negativos es Y , el estadístico de prueba será la menor de estas sumas. Este valor se compara con el valor obtenido de la tabla del anexo C8, rechazando la hipótesis nula si el valor del estadístico de prueba es menor que el especificado en dicha tabla.

Es de hacer notar que la prueba de rangos y signos es un método simple y de mucha utilidad, aunque su principal limitación radica en que no se puede aplicar a conjuntos de datos muy pequeños, para el caso de prueba de dos colas con un nivel de significación $p = 0.05$, n tiene que ser al menos 6 (MILLER y MILLER, 1993, cap. 6).

8.5 PRUEBA DE SUMA DE RANGOS DE WILCOXON

La prueba de rangos y signos descrita anteriormente es válida para el estudio de conjuntos individuales de mediciones y para conjuntos de datos por parejas que se pueden reducir con facilidad a conjuntos individuales. Sin embargo, en muchos casos es necesario comparar dos muestras independientes que no se puedan reducir a un conjunto único de datos, ya que pueden contener diferentes números de mediciones. Existen varias pruebas no paramétricas que permiten abordar estas situaciones (MILLER y MILLER, 1993, cap. 6). Sean A y B , dos muestras con M y N observaciones respectivamente, donde se cumple que $M \geq N$, la prueba de suma de rangos de Wilcoxon para determinar si ambas muestras proceden de la misma distribución se describe a continuación (PRIA, 2001, cap. 3):

1. Se ordenan todas las observaciones de ambas muestras, como si fuera una sola muestra, en orden ascendente y se asignan los rangos a los valores ordenados.
2. Se identifican los valores que pertenecen a cada muestra.
3. Se determina el estadígrafo que en esta prueba es: $T_o =$ Suma de rangos de $B = (\sum \text{rangos } B)$ (en donde B es la muestra más pequeña)
4. *La regla de decisión será:*
 - a) Se plantea la Hipótesis que se adecue a la situación que se necesita resolver, y se aplica la regla de decisión de acuerdo a lo que se presenta en el cuadro 8.3:

CUADRO 8.3 APLICACIÓN DE LA REGLA DE DECISIÓN PARA LA PRUEBA DE SUMA DE RANGOS DE WILCOXON

Hipótesis	Regla de decisión rechazar H_0 si:	α más usados
$H_0 : Me_B = Me_A$ $H_1 : Me_B \neq Me_A$	$T_o \leq T_I$ ó $T_o \geq T_S$	0.025 0.05
$H_0 : Me_B \geq Me_A$ $H_1 : Me_B < Me_A$	$T_o \leq T_I$	0.05 0.01
$H_0 : Me_B \leq Me_A$ $H_1 : Me_B > Me_A$	$T_o \geq T_S$	0.05 0.01

Donde T_I y T_S son los valores obtenidos en la tabla de Valores críticos para la estadística de prueba de la Suma de Rangos de Wilcoxon (ver anexo D8), considerando un tamaño de n_A y n_B y un nivel de significación dado. Esta tabla sirve para trabajar cuando el tamaño de la muestra llega hasta 25 la muestra menor y 50 la mayor. Me_B y Me_A representan los parámetros de tendencia central de las distribuciones de ambas muestras.

Cuando el tamaño de la muestra menor excede las 25 observaciones puede trabajarse con la aproximación a la distribución normal (PRIA, 2001, cap. 3).

Otra forma de rechazar la hipótesis nula es cuando se está trabajando con un programa estadístico que calcule T_o y p que es la probabilidad de error de tipo I asociada a ese valor. En éste caso si el valor de p es menor que el α prefijado se rechaza la hipótesis nula.

Deben aclararse los aspectos siguientes (PRIA, 2001, cap. 3):

1. Si $n_A = n_B$ se seleccionará el estadígrafo Σ rangos de la muestra tomando en consideración las hipótesis alternativas, de la forma siguiente:

Si la Hipótesis Alternativa es:	Se seleccionará la muestra:
$H_1 : Me_B \neq Me_A$	Cualquiera de las dos tiene igual solución.
$H_1 : Me_B < Me_A$	T_o = la que tenga mayor suma de rangos.
$H_1 : Me_B > Me_A$	T_o = la que tenga menor suma de rangos.

2. Las hipótesis deben plantearse tomando como primer parámetro de referencia para contrastar el de la muestra más pequeña, en el caso de que sean de igual tamaño podrán plantearse de cualquier forma.

3. En caso de existir observaciones de igual valor se asignarán rangos promedios, por ejemplo si los tres primeros valores son:

VALORES	15	15	15
LUGARES	1	2	3
RANGOS	$(1+2+3)/3 = 2$	$(1+2+3)/3 = 2$	$(1+2+3)/3 = 2$

4. Potencia - Eficiencia:

Según Siegel, si la prueba de la Suma de Rangos de Wilcoxon se aplica a datos que pueden analizarse adecuadamente por una prueba paramétrica más poderosa, la prueba t de Student en éste caso, su potencia-eficiencia sería de cerca de un 95.5% y se acerca a 95% para muestras de tamaño moderado, por lo que es una excelente alternativa ante la prueba t .

Debe destacarse que esta prueba es más potente que la prueba de la Mediana, pues esta última utiliza solamente la información de cómo están ubicadas las observaciones de cada muestra con relación al valor de la *mediana general*, en cambio la prueba de *suma de rangos de Wilcoxon* utiliza además, la información relativa a la ubicación de cada observación en las muestras, que se resume en el estadígrafo de la *suma de rangos*.

8.6 PRUEBA MANN-WHITNEY

Otra prueba que se utiliza para resolver el mismo caso que resuelve la prueba de la Suma de Rangos de Wilcoxon es la prueba de Mann-Whitney. El procedimiento seguido en ambas es muy parecido. Las hipótesis que se contrastan son las mismas y el estadígrafo utilizado se parece aunque por supuesto no es igual. A continuación se describe el procedimiento para el contraste de hipótesis mediante el uso de la prueba de Mann-Whitney:

Inicialmente se identifican ambas muestras A y B, con M y N observaciones respectivamente, donde se cumple que $M \geq N$. Todas las observaciones de ambas muestras se ordenan, como si fuera una sola muestra, en orden ascendente y se asignan los rangos a los valores ordenados. Posteriormente, se identifican los valores que

pertencen a cada muestra y se calculan las *sumas de rangos* de cada muestra y se define S , que es la *suma de rangos* de menor valor (PRIA, 2001, cap. 3).

El estadístico de contraste para esta prueba es (PRIA, 2001, cap. 3):

$$T_o = S - \frac{N(N+1)}{2}$$

La *regla de decisión* vendrá dada en función del planteamiento de la hipótesis que se adecue a la situación que se desea resolver, como se muestra en el cuadro 8.4 (PRIA, 2001, cap. 3):

CUADRO 8.4 APLICACIÓN DE LA REGLA DE DECISION PARA LA PRUEBA DE MANN-WHITNEY

Hipótesis	Regla de decisión rechazar H_0 si:	α más usados
$H_0 : Me_B = Me_A$ $H_1 : Me_B \neq Me_A$	$T_o \leq W_{\alpha/2}$ o $T_o \geq W_{1-\alpha/2}$	0.025 0.005
$H_0 : Me_B \geq Me_A$ $H_1 : Me_B < Me_A$	$T_o \leq W_\alpha$	0.05 0.01
$H_0 : Me_B \leq Me_A$ $H_1 : Me_B > Me_A$	$T_o \geq W_{1-\alpha}$ donde $W_{1-\alpha} = NM - W_\alpha$	0.05 0.01

Donde W_α y $W_{\alpha/2}$ son los valores críticos obtenidos en la tabla de valores críticos para la estadística de prueba de Mann-Whitney presentada en el anexo D8, en la que se consideran tamaños muestrales n_A , n_B y un nivel de significación α . Esta tabla sirve para trabajar cuando los tamaños muestrales son de 20 para la muestra menor y 40 para la mayor.

Cuando el tamaño de la muestra menor excede las 20 observaciones puede trabajarse con la aproximación a la distribución normal (ver capítulo 3.0).

Otra forma de rechazar la hipótesis nula es cuando se está trabajando con un programa estadístico que calcule T_o y p que es la probabilidad de error de tipo I asociada a ese valor, en este caso se rechaza la hipótesis nula si p es menor al nivel de significación α tomado para la prueba.

8.7 PRUEBA DE KRUSKAL WALLIS

Cuando se presenta el problema de comparar más de dos muestras con el propósito de conocer si proceden de la misma población, o bien, comparar si existen diferencias entre las medidas de tendencia central de más de dos poblaciones y no se justifica la suposición de normalidad y de igualdad de varianzas, el químico analista podrá recurrir a un procedimiento alternativo al de la prueba F del análisis de la variancia y que no depende de esta suposición. Kruskal y Wallis (1952) desarrollaron un procedimiento como alternativa para dar solución a este problema, conocido como *la prueba de Kruskal – Wallis*.

La *prueba de Kruskal–Wallis* constituye una alternativa no paramétrica al análisis de varianza usual y se considera como una extensión del procedimiento de *suma de rangos de Wilcoxon* como se verá en su desarrollo (MILLER y FREUND, 1989, cap. 10).

La hipótesis nula para la prueba de Kruskal–Wallis es que no existe diferencia entre los tratamientos ($\mu_1 = \mu_2 = \dots = \mu_a$), mientras que la hipótesis alternativa es que exista diferencia entre al menos un par de tratamientos ($(\mu_i \neq \mu_j)$).

Para realizar la prueba de Kruskal–Wallis los datos pueden agruparse como se presenta en la tabla a continuación (MONTGOMERY, 1991, cap. 4):

Repeticiones	Tratamientos			
	A	B	C	...
1	X_{1A}	X_{1B}	X_{1C}	...
2	X_{2A}	X_{2B}	X_{2C}	...
⋮	⋮	⋮	⋮	...
n	X_{nA}	X_{nB}	X_{nC}	...

Donde X_{ij} son las observaciones
 $i = 1, 2, \dots, n$
 $j = A, B, C, \dots$
 $N = n_A + n_B + n_C \dots$

Las observaciones se organizan posteriormente, en orden ascendente y se asignan las posiciones (o rangos) con el rango 1 correspondiente a la observación más pequeña. En caso de empate (varias observaciones con el mismo rango o posición), se asigna el rango promedio a cada observación empatada (MONTGOMERY, 1991, cap. 4).

El estadístico de prueba de Kruskal–Wallis está dado por la expresión (MONTGOMERY, 1991, cap. 4):

$$H = \frac{1}{S^2} \left[\sum_{i=1}^a \frac{R_i^2}{n_i} - \frac{N(N+1)^2}{4} \right]$$

En donde R_i es la suma de los rangos de las observaciones del i -ésimo tratamiento y n_i es el número de observaciones del i -ésimo tratamiento; N es el número total de observaciones y (MONTGOMERY, 1991, cap. 4)

$$S^2 = \frac{1}{N-1} \left[\sum_{i=1}^a \sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N-1)^2}{4} \right]$$

Debe notarse que S^2 es igual a la varianza de los rangos. Si no hay empate, $S^2 = N(N+1)/12$ y el estadístico de prueba se simplifica a (MONTGOMERY, 1991, cap. 4):

$$H = \frac{12}{N-1} \left[\sum_{i=1}^a \frac{R_i^2}{n_i} - 3(N+1) \right]$$

Si n_i es razonablemente grande, como sería el caso de $n_i \geq 5$, entonces H tiene una distribución aproximadamente χ_{a-1}^2 (ver tabla del anexo C3) si la hipótesis nula es verdadera. Por lo tanto, si (MONTGOMERY, 1991, cap. 4)

$$H \geq \chi_{\alpha, a-1}^2$$

hay que rechazar la hipótesis nula.

Si H se calcula mediante un programa estadístico en una computadora se obtendrá el valor de H con la probabilidad exacta de error tipo I asociada a ese valor.

Debe señalarse que la potencia-eficiencia de esta prueba comparada con la prueba paramétrica más poderosa, la prueba F , considerando que se cumplen los supuestos para la misma es de un 95.5% (PRIA, 2001, cap. 3).

8.8 PROBLEMAS DE APLICACIÓN DE METODOS NO PARAMETRICOS

1. En relación al ejercicio 1 de la sección 3.5.2 (cap. 3), que se refiere a la determinación del tanto por ciento del níquel de una muestra particular de acero de referencia del NIST por un nuevo método espectrofotométrico, determinar mediante la prueba del signo para los niveles de significación del 0.01 y 0.05 si $\mu = 1.12\%$.

Datos: 1.10, 1.08, 1.09, 1.12, 1.109

Solución:

a) Hipótesis nula: $\mu = 1.12\%$
 Hipótesis alterna: $\mu \neq 1.12\%$

b) Nivel de significación: $\alpha = 0.01$ y 0.05

c) Cálculo:

Reemplazando cada valor mayor que 1.12% por un signo positivo y cada valor menor que 1.12% por un signo menos, los cinco valores generados son:

- - - 0 -

d) Criterio: el criterio de decisión se basa en el número signos positivos (X) o en el número de signos negativos (Y). Utilizando el número de signos menos frecuente, se rechaza H_0 si la probabilidad de obtener X ó Y o más signos es menor o igual a 0.01 ó 0.05 (dependiendo del nivel de significación dado a la prueba.

Así, para $X = 0$ que es el *número de signos positivos*, el cálculo de la probabilidad binomial para $X \geq 0$ con parámetros $n = 4$ y $p = 0.5$ es:

$$P(X \geq 0; n = 4, p = 0.5) = \binom{4}{0} 0.5^0 (1 - 0.5)^{4-0} = 0.0625$$

Para un nivel de significación del 10%, puede rechazarse la hipótesis nula de que $\mu = 1.12\%$, dado que $P(=0.0625) < 0.1$. Mientras que si se usa un nivel de significación del 5% no existe evidencia como para rechazar H_0 . Se observa, en el primer caso, que se coincide con los resultados obtenidos en la prueba paramétrica, mientras que en el segundo caso, no. La prueba *paramétrica* puede volverse más concluyente en la medida que se aumente el número de mediciones.

2. Los siguientes quince datos son mediciones del punto de ebullición de un compuesto de silicio en °C: 166, 141, 136, 153, 170, 162, 155, 146, 183, 157, 148, 132, 160, 175, 150. Para un nivel de significación de 0.05, utilizar las pruebas del signo y de rangos y signos de Wilcoxon para probar la hipótesis de que $\mu = 158^\circ\text{C}$.

Solución:*Prueba del signo*

a) Hipótesis nula: $\mu = 158^\circ\text{C}$
 Hipótesis alterna: $\mu \neq 158^\circ\text{C}$

b) Nivel de significación: $\alpha = 0.05$

c) Cálculo:

Reemplazando cada valor mayor que 158°C por un signo positivo y cada valor menor que 158°C por un signo menos, los quince valores generados son:

+ - - - + + - - + - - - + + -

d) Criterio: el criterio de decisión se basa en el número signos positivos (X) o en el número de signos negativos (Y). Utilizando el número de signos menos frecuente, se rechaza H_0 si la probabilidad de obtener X ó Y o más signos es menor o igual 0.05 (dependiendo del nivel de significación dado a la prueba). El signo menos frecuente es el signo más (+) y su número es $X = 6$, el cálculo de la probabilidad binomial para $X \geq 6$, con parámetros $n = 15$ y $p = 0.5$ es:

$$P(X \geq 6; n = 15, p = 0.5) = \sum_{x=0}^6 \binom{15}{x} (0.5)^x (0.5)^{15-x} = 0.3036$$

Dado que 0.3036 es mayor que el nivel de significación α de 0.05, no existe evidencia como para rechazar que el punto de ebullición del compuesto de silicio sea de 158°C.

Suma de rangos y signos de Wilcoxon

a) Hipótesis nula: $\mu = 158^\circ\text{C}$
Hipótesis alterna: $\mu \neq 158^\circ\text{C}$

b) Nivel de significación: $\alpha = 0.05$

c) Cálculo:

Comparando cada dato obtenido con 158°C, se obtiene:

8 -17 -22 -5 12 4 -3 -12 25 -1 -10 -26 2 17 -8

Ordenando los datos de menor a mayor sin tomar en cuenta los signos se tiene:

1 2 3 4 5 8 8 10 12 12 17 17 22 25 26

Colocando los rangos y los signos asociados a cada observación se tiene:

-1 2 -3 4 -5 6.5 -6.5 -8 9.5 -9.5 11.5 -11.5 -13 14 -15

d) Criterio: el criterio de decisión se basa en la suma de los rangos con signo positivo (X) o en la suma de los rangos con signo negativo (Y). Utilizando la suma menor, se rechaza H_0 si al compararla con el valor obtenido de la tabla del anexo B8 se observa que es menor o igual que este.

En este problema:

$$\begin{aligned} X \text{ (suma de rangos con signo positivo)} &= 47.5 \\ Y \text{ (suma de rangos con signo negativo)} &= 72.5 \end{aligned}$$

Por lo que el estadístico de contraste es $X = 47.5$, de la tabla del anexo B8, se obtiene el valor de 25 para $n = 15$ y $\alpha = 0.05$. Dado que el valor del estadístico de contraste es mayor que el valor leído en la tabla puede concluirse que no existe evidencia significativa al 5% como para afirmar que el punto de ebullición del compuesto de silicio difiere de 158°C .

Como puede observarse este resultado, coincide con el generado a partir de la aplicación de la prueba del signo.

3. En relación al ejercicio 3 de la sección 3.2.5(cap. 3) que se refiere a la determinación de la homogeneidad de una muestra patrón de cloruros mediante análisis de porciones de material, tomadas en la superficie y en el fondo del contenedor, en el que los datos obtenidos fueron:

| % de Cloruros en la superficie | % de cloruros en el fondo |
|--------------------------------|---------------------------|
| 26.32 | 26.28 |
| 26.33 | 26.25 |
| 26.38 | 26.38 |
| 26.39 | |

A partir de la prueba de la mediana, determinar si existe homogeneidad en el material a un nivel de significación α del 5%.

Solución:

- a) *Hipótesis nula:* hay homogeneidad en el contenido de cloruros de la muestra patrón, es decir: $Me_{\text{fondo}} = Me_{\text{superficie}}$
- b) *Hipótesis alterna:* no hay homogeneidad en el contenido de cloruros de la muestra patrón, es decir: $Me_{\text{fondo}} \neq Me_{\text{superficie}}$
- c) *Nivel de significación(α):* 0.05
- d) *Procedimiento de prueba:*
- El primer paso consiste en ordenar la serie de datos en orden ascendente como si procedieran de la misma muestra, para los datos del ejercicio se tiene:

26.25 26.28 26.32 26.33 26.38 26.38 26.39

- La mediana de la serie anterior es: 26.33
- El resumen de la información del problema para la prueba de la mediana se presenta en la tabla a continuación:

| Valores/ mediana | Muestra 1 | Muestra 2 | Total |
|------------------------|-----------|-----------|-------|
| Obs. > mediana general | 2(a) | 2(b) | 4 |
| Obs. ≤ mediana general | 2(c) | 1(d) | 3 |
| Totales | 4 | 3 | 7 |

- El cálculo del estadístico para el contraste es el siguiente:

$$\chi_o^2 = \frac{n \left(|ad - bc| - \frac{n}{2} \right)^2}{(a+b)(c+d)(a+c)(b+d)}$$

$$\chi_o^2 = \frac{7 \left(|2 \times 1 - 2 \times 2| - \frac{7}{2} \right)^2}{4 \times 3 \times 4 \times 3} = \frac{15.75}{144} = 0.109$$

- e) *Criterio de decisión: si $\chi_o^2 \geq \chi_{tabulado}^2$ con un grado de libertad y para un $\alpha = 0.05$ (ver anexo C3, tabla de χ^2) se rechaza la hipótesis nula. Dado que $0.109 < 3.84$ (ver anexo C3), no existe evidencia significativa al 5% como para rechazar la hipótesis nula. Por lo que se concluye que no existe evidencia a un nivel de significación de 0.05 de que las medianas de las distribuciones de los valores de cloruro varíen entre la superficie y el fondo del contenedor. Como puede observarse este resultado coincide con el obtenido mediante la prueba paramétrica.*

4. En el desarrollo de un nuevo método para la determinación de niveles de alcohol en la sangre, se analizó cinco veces una muestra de sangre, con los resultados siguientes: 64.5, 66.0, 63.9, 65.1 y 64.0 mg/100 ml. El método de análisis estándar aplicado a la misma muestra proporciona los resultados 66.2, 65.8, 66.3, 65.6 mg/100ml. Utilizando las pruebas de U-Mann Whitney y la de suma de rangos de Wilcoxon, probar si los métodos difieren significativamente.

Solución:

Prueba de U-Mann Whitney

- a) *Hipótesis nula: $Me_{\text{método estándar}} = Me_{\text{método nuevo}}$*

b) *Hipótesis alterna:* $Me_{\text{método estándar}} \neq Me_{\text{método nuevo}}$

c) *Nivel de significación*(α): 0.05

d) *Procedimiento de prueba:*

- El primer paso consiste en asignar los rangos a los datos de ambas muestras como si todos fueran una sola muestra, de la siguiente forma:

| Método nuevo | Rango | Método estándar | Rango |
|--------------|-------|-----------------|-------|
| 65.5 | 4 | 66.2 | 8 |
| 66.0 | 7 | 65.8 | 6 |
| 63.9 | 1 | 66.3 | 9 |
| 65.1 | 3 | 65.6 | 5 |
| 64.0 | 2 | | |

- La suma de rangos de cada muestra es:

$$\begin{aligned} \text{Muestra del nuevo método} &= \mathbf{17} \\ \text{Muestra del método estándar} &= \mathbf{28} \end{aligned}$$

- El estadístico viene dado por tanto como:

$$T_o = S - \frac{N(N+1)}{2} = 17 - \frac{5(5+1)}{2} = 2$$

Para $S = 17$ (menor de las sumas de rangos) y $N = 5$ (número de datos para la muestra con la menor suma de rangos).
De esta forma $T_o = 2$

e) *Criterio de decisión:*

El valor de $W_{\alpha/2, n1, n2}$ obtenido de la tabla del anexo D8 es:

$$W_{0.05, 4, 5} = 1$$

Dado que $T_o = 2 > W_{0.05, 4, 5} = 1$, no existe evidencia a un nivel de significación del 5% como para rechazar la hipótesis nula, o, en otras palabras, se concluye que no existe evidencia significativa al 5% como para rechazar que los dos métodos producen resultados equivalentes.

Prueba de Suma de Rangos de Wilcoxon:

La variante de esta prueba con respecto a la prueba de U-Mann Whitney

se encuentra únicamente en el estadístico de contraste empleado, pues en la prueba de Suma de rangos de Wilcoxon, se calculan dos estadísticos que vienen dados como:

$$T_1 = S_1 - \frac{N_1(N_1 + 1)}{2} = 17 - \frac{5(5 + 1)}{2} = 2$$

$$T_2 = S_2 - \frac{N_2(N_2 + 1)}{2} = 28 - \frac{4(4 + 1)}{2} = 18$$

Criterio de decisión:

El valor de $W_{\alpha/2, n_1, n_2}$ obtenido de la tabla del anexo D8 es:

$$W_{0.05, 4, 5} = 1$$

Dado que la menor de estas sumas es 2 y supera al valor tabulado de 1, no existe evidencia significativa al 5% como para rechazar H_0 .

5. En relación al ejercicio 6 de la sección 3.5.2, de la comparación de los resultados obtenidos en la aplicación de un nuevo método de determinación de Indio en Zinc y sus aleaciones y el valor reportado en el patrón certificado. Estudiar con un nivel de significación del 5%, mediante la prueba de rangos y signos de Wilcoxon, si existe diferencia entre las mediciones con el nuevo método y el valor certificado del patrón. La tabla de datos se presenta a continuación:

| Patrón | Observación | Certificado |
|---------------|--------------------|--------------------|
| CRM322 | 2 | 3 |
| CRM323 | 6 | 5 |
| CRM324 | 15 | 16 |
| CRM325 | 44 | 46 |
| CRM352 | 2 | 3 |
| CRM353 | 1 | 3 |
| CRM354 | 7 | 10 |
| CRM357 | 6 | 3 |
| CRM358 | 12 | 7 |
| CRM359 | 16 | 16 |
| CRM360 | 29 | 30 |

Solución:

- a) *Hipótesis nula:* $\mu_d = 0$
 b) *Hipótesis alterna:* $\mu_d \neq 0$

c) Nivel de significación(α): 0.05

d) Procedimiento de prueba:

En la tabla a continuación se presentan en la primera columna las diferencias calculadas para cada par de datos, en la segunda columna se presentan estas diferencias ordenadas en orden ascendente sin tomar en cuenta el signo y en la columna tres se presenta el rango asignado a cada diferencia:

| $d_i = \text{obs- Cert.}$ | orden | rango |
|---------------------------|-------|-------|
| -1 | 0 | 1 |
| 1 | -1 | -4 |
| -1 | -1 | -4 |
| -2 | -1 | -4 |
| -1 | -1 | -4 |
| -2 | 1 | 4 |
| -3 | -2 | -7.5 |
| 3 | -2 | -7.5 |
| 5 | -3 | -9.5 |
| 0 | 3 | 9.5 |
| -1 | 5 | 11 |

La suma de los rangos con signo positivo y negativo son respectivamente:

$$X (\text{suma de rangos con signo } +) = 25.50$$

$$Y (\text{suma de rangos con signo } -) = 40.50$$

Por lo tanto el estadístico de prueba es: $X = 25.5$

e) Criterio de decisión:

Al comparar el estadístico de prueba con el valor obtenido de la tabla del anexo C8, para $n = 11$, se observa que $X = 25.5 > 10$, por lo tanto no existe evidencia de que la mediana (media) de la diferencia sea distinta de cero, es decir, no existe evidencia significativa al 5% de la diferencia entre las mediciones de la concentración del patrón por el nuevo método y valor certificado del mismo.

6. Considere el ejercicio 1 de la sección 4.1.10 (cap. 4), en el que se quiere comparar el trabajo de cuatro analistas de un laboratorio en el ensayo de determinación del % de alcohol metílico en muestras de un producto químico, mediante la técnica de cromatografía líquida de alta resolución (HPLC). Los analistas reportaron los resultados siguientes:

| Analista | % de alcohol metílico | | |
|----------|-----------------------|-------|-------|
| 1 | 84.99 | 84.02 | 84.38 |
| 2 | 85.15 | 85.13 | 84.88 |
| 3 | 84.72 | 84.48 | 85.16 |
| 4 | 84.20 | 84.10 | 84.55 |

Mediante la prueba no paramétrica de Kruskal-Wallis, determinar si el trabajo de los analistas difiere significativamente. Use un nivel de significación de 0.05.

Solución:

- f) *Hipótesis nula:* los cuatro analistas trabajan de forma equivalente, es decir: $Me_1 = Me_2 = Me_3 = Me_4$
- g) *Hipótesis alterna:* al menos dos analistas trabajan de forma diferente:
- h) *Nivel de significación*(α): 0.05
- i) *Procedimiento de prueba:*

- El primer paso consiste en ordenar la serie de datos en orden ascendente como si procedieran de la misma muestra, luego se establecen los rangos. La asignación de rangos se presenta en la siguiente tabla:

| Muestra | Analista 1 | Analista 2 | Analista 3 | Analista 4 |
|--------------|------------|------------|------------|------------|
| 1 | 9 | 11 | 7 | 3 |
| 2 | 1 | 10 | 5 | 2 |
| 3 | 4 | 8 | 12 | 6 |
| $\sum R_j =$ | 14 | 29 | 24 | 11 |

- El cálculo del estadístico para el contraste es el siguiente:

Como no hay empates, el estadístico de prueba será:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^a \frac{R_i^2}{n_i} - 3(N+1)$$

$$H = \frac{12}{12 \times 13} \left(\frac{14^2}{3} + \frac{29^2}{3} + \frac{24^2}{3} + \frac{11^2}{3} \right) - 3 \times 13 = 5.46$$

Criterio de decisión: si $H > \chi^2_{\alpha, a-1}$ se rechaza la hipótesis nula. En este caso $\chi^2_{0.05,3} = 7.81$ (de tabla del anexo B3). Dado que $5.46 < 7.81$, no existe evidencia significativa al 5% como para rechazar la hipótesis nula, por lo que se concluye que no existe evidencia a un nivel de significación de 0.05, como para afirmar que los analistas trabajan de forma diferente.

REFERENCIAS BIBLIOGRÁFICAS

- **Freund, J.E. y Walpole, R.E. (1990).** Estadística matemática con aplicaciones. Cuarta edición. México: PRENTICE HALL HISPANOAMERICANA S.A.
- **Mendenhall, W. (1990).** Estadística para administradores. Segunda edición. México: Grupo Editorial Iberoamérica.
- **Miller, I. y Freund, J.E. (1986).** Probabilidad y estadística para ingenieros. Tercera edición. México: PRENTICE HALL HISPANOAMERICANA S.A.
- **Miller, J. C. y Miller, J. N. (1993).** Estadística para química analítica. Segunda edición. Wilmington, Delaware, E.U.A: Addison Wesley Iberoamericana, S.A.
- **Pérez, C. (2001).** Técnicas estadísticas con SPSS. España: Pearson Education.
- **Pria Barros, M. C. (2001).** Métodos no paramétricos. [En línea]. Cuba: Universidad de La Habana. Disponible en Word Wide Web: [http://www.Vcl.sld.cy/75 cm/facmedic/ webosalud/materiales/mnoparam. html](http://www.Vcl.sld.cy/75%cm/facmedic/webosalud/materiales/mnoparam.html) [citado 10 de enero 2003]
- **Walpole, R. E. y Myers, R. H. (1992).** Probabilidad y estadística. Cuarta edición. México: Mc GRAW-HILL.
- **Wayne, W. D. (1992).** Bioestadística. Base para el análisis de las ciencias de la salud. Cuarta edición. México: Limusa.

**ANEXO A8: METODOS NO PARAMETRICOS DE CORRELACION,
REGRESION LINEAL SIMPLE Y BONDAD DE AJUSTE
(PRIA, 2001, cap. 4)**

1. COEFICIENTE DE CORRELACION DE SPEARMAN

En muchas ocasiones es importante conocer si dos variables están relacionadas y si lo están, evaluar cuál es la intensidad de dicha relación.

En las estadísticas clásicas la medida de correlación usada es el coeficiente de Pearson que se estima por el método de mínimos cuadrados que requiere:

1. Que las variables estén medidas en una escala continua
2. Si se quiere medir su significación estadística es necesario que estas variables tengan una distribución normal bivariada.

Cuando estos supuestos no se cumplen pueden utilizarse medidas de correlación no paramétricas, que permiten además establecer una correlación existente, determinar su significación estadística.

Existen varios coeficientes de correlación no paramétricos, el que se expone aquí es el coeficiente de correlación de Spearman, que fue la primera estadística de rangos que se desarrolló y que requiere que ambas variables estén medidas al menos en una escala ordinal de forma que las unidades experimentales puedan colocarse en dos series y que las observaciones se hayan extraído aleatoriamente de una población.

Sean n observaciones de dos variables aleatorias x e y , medidas ambas al menos, en escala ordinal de forma simultánea en una misma unidad de observación. Los datos así obtenidos pueden presentarse de la forma siguiente:

| Elementos | Variable | |
|-----------|----------|-------|
| | X | Y |
| 1 | x_1 | y_1 |
| 2 | x_2 | y_2 |
| 3 | x_3 | y_3 |
| ⋮ | ⋮ | |
| n | x_n | y_n |

Cada elemento queda representado por un par de valores de la variables x e y , es decir por el punto (x_i, y_i) , que posteriormente puede ser llevado a un gráfico de dispersión que se utiliza para explorar la relación lineal entre las variables.

A cada valor de la variable x e y se asignan R_{1j}, R_{2j} respectivamente, convirtiéndose la información en una matriz de rangos como la siguiente:

| Individuos | Variable x | Rango de x | Variable y | Rango de y | $d_{ij} = R_{1j} - R_{2j}$ |
|------------|------------|------------|------------|------------|----------------------------|
| 1 | x_1 | R_{11} | y_1 | R_{21} | $R_{11} - R_{21}$ |
| 2 | x_2 | R_{12} | y_2 | R_{22} | $R_{12} - R_{22}$ |
| 3 | x_3 | R_{13} | y_3 | R_{23} | $R_{13} - R_{23}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| N | x_n | R_{1n} | y_n | R_{2n} | $R_{1n} - R_{2n}$ |

d_{ij} mide las discrepancias que hay entre los rangos de x e y para cada individuo. En el caso de existir observaciones de igual valor deben asignarse rangos promedios, por ejemplo si los 3 primeros valores son:

| | | | |
|----------------|---------------|---------------|---------------|
| Valores | 15 | 15 | 15 |
| Lugares | 1 | 2 | 3 |
| Rangos | $(1+2+3)/3=2$ | $(1+2+3)/3=2$ | $(1+2+3)/3=2$ |

La fórmula del coeficiente de correlación de Rangos de Spearman se deduce a partir de la del coeficiente de correlación de Pearson, sustituyendo los valores de las observaciones de cada variable por la de sus rangos, llegándose a obtener la siguiente expresión:

$$R_s = 1 - \frac{6 \sum_{i=1}^n di^2}{n^3 - n}$$

De esta forma R_s brinda una estimación del grado de asociación entre las variables x e y . Debe recordarse que cuando se analiza un coeficiente de correlación se deben considerar dos aspectos:

1. Su magnitud: oscila entre 0 y 1, valores cercanos a 0 significan que no hay relación lineal entre ambas variables, mientras que los valores cercanos a 1 significan que la relación lineal entre las variables es muy intensa.
2. Su sentido. Está dado por el signo + o -, el signo - significa que existe una relación inversamente proporcional; es decir, que a valores altos de una variable corresponden valores bajos de la otra. El signo + significa que existe una relación directamente proporcional entre las variables o sea que a valores altos de una se corresponden valores altos de la otra y viceversa.

Las hipótesis a plantear para resolver la situación descrita son las siguientes:

| Hipótesis | Rechazar H_0 Si | α más usados |
|--------------------|-----------------------------|---------------------|
| $H_0 : R_s = 0$ | $-R_s \leq -R_{tabulada}$ o | 0.025 |
| $H_1 : R_s \neq 0$ | $R_s \geq R_{tabulada}$ | 0.005 |
| $H_0 : R_s \geq 0$ | $-R_s \leq -R_{tabulada}$ | 0.05 |
| $H_1 : R_s < 0$ | | 0.01 |
| $H_0 : R_s \leq 0$ | $R_s \geq R_{tabulada}$ | 0.05 |
| $H_1 : R_s > 0$ | | 0.01 |

Donde $-R_{tabulada}$ y $R_{tabulada}$ son los valores críticos obtenidos en la tabla de *valores críticos para el coeficiente de correlación de rangos de Spearman R_s* , ver tabla F.8.1 al final del anexo, para tamaños de muestra que oscilen entre 4 y 30 y un nivel de significación dado. Cuando el tamaño de la muestra excede las 30 observaciones puede trabajarse entonces con la aproximación a la distribución normal:

$$Z = R_s \sqrt{n-1} \quad \text{que se distribuye Normal } (0, 1)$$

Otra forma de rechazar la hipótesis nula es cuando se está trabajando con un programa estadístico entonces se calcula R_s y p que es la probabilidad de error de tipo I asociada a ese valor. En este caso si el valor de p es menor que el α prefijado se rechaza la hipótesis nula.

Según Siegel, si la prueba de correlación de rangos de Spearman se aplica a datos que pueden analizarse adecuadamente por una prueba paramétrica más poderosa, como la prueba del coeficiente de correlación de Pearson en éste caso, su potencia eficiencia sería de cerca de un 91%.

2. METODOS DE REGRESION NO PARAMETRICOS

En muchas ocasiones existe interés en realizar aproximaciones no paramétricas cuando se quiere ajustar una línea recta a un conjunto de puntos. Theil desarrolló dos métodos: el método completo de Theil y el método incompleto de Theil, denominado de esta forma para distinguirlo del otro método y por ser quizá el más simple, el cual se expone a continuación. Theil supone que un conjunto de puntos $(x_1, y_1), (x_2, y_2), \dots$, va a ser ajustado por una recta de la forma $y = bx + a$.

Para los cálculos Theil desarrolla los siguientes pasos:

1. Disponer los puntos en orden de x creciente. Si el número de puntos x , es impar, el punto medio, es decir, el valor mediano de x se borra, el cálculo simple exige un número par de puntos.
2. Para cualquier par de puntos $(x_i, y_i), (x_j, y_j)$ donde $x_j < x_i$, la pendiente, b_{ij} , de la línea que une los puntos es:
$$b_{ij} = \frac{y_i - y_j}{x_j - x_i}.$$

Se calculan las pendientes b_{ij} para el par de puntos (x_1, y_1) y el inmediatamente posterior al valor mediano de la x , para (x_2, y_2) y el segundo punto después del valor mediano de la x , y así sucesivamente hasta que se calcula la pendiente para la línea que une el punto inmediatamente anterior a la mediana de x con el último punto.

3. Las estimaciones de la pendiente, se disponen en orden creciente y su valor mediano, es la pendiente estimada de la línea recta.
4. Con este valor de b , los valores a_i se estiman para cada punto con la ayuda de la ecuación $y = bx + a$.
5. Las estimaciones de a se ordenan en forma creciente y el valor mediano se elige con la mejor estimación.

3. LA PRUEBA DE BONDAD DEL AJUSTE DE KOLMOGOROV

La bondad del ajuste surge cuando se necesita probar la procedencia de una muestra de observaciones de una distribución completa, como puede ser la distribución normal. La prueba Chi-cuadrada es utilizada para este propósito cuando los datos se presentan como

frecuencias, aunque la prueba normalmente no se utiliza para menos de cincuenta observaciones, y es difícil usarla con datos continuos y aislados.

El método de Kolmogorov tiene dos aplicaciones frecuentes la *prueba de aleatoriedad* y la *prueba de normalidad de una distribución*.

Al utilizar la prueba de bondad del ajuste de Kolmogorov se hace una comparación entre alguna función de distribución acumulada y teórica ($F_{T(x)}$), y la función de distribución acumulada de una muestra, $F_{s(x)}$ el estadístico a utilizar es:

$$D = \text{Máx.} |F_{s(x)} - F_{T(x)}|.$$

La hipótesis nula se rechaza al nivel de significación α si el valor calculado D excede al valor mostrado en la tabla A8-1 al final del anexo, para $1 - \alpha$.

TABLA A8-1 VALORES CRITICOS DE D PARA LA PRUEBA DE KOLMOGOROV SMIRNOFF
(MILLER y FREUND, 1986, pág. 542)

| Tamaño muestral
n | $D_{.10}$ | $D_{.05}$ | $D_{.01}$ |
|------------------------|-----------|-----------|-----------|
| 1 | 0.950 | 0.975 | 0.995 |
| 2 | 0.776 | 0.842 | 0.929 |
| 3 | 0.642 | 0.708 | 0.828 |
| 4 | 0.564 | 0.624 | 0.733 |
| 5 | 0.510 | 0.565 | 0.669 |
| 6 | 0.470 | 0.521 | 0.618 |
| 7 | 0.438 | 0.486 | 0.577 |
| 8 | 0.411 | 0.457 | 0.543 |
| 9 | 0.388 | 0.432 | 0.514 |
| 10 | 0.368 | 0.410 | 0.490 |
| 11 | 0.352 | 0.391 | 0.468 |
| 12 | 0.338 | 0.375 | 0.450 |
| 13 | 0.325 | 0.361 | 0.433 |
| 14 | 0.314 | 0.349 | 0.418 |
| 15 | 0.304 | 0.338 | 0.404 |
| 16 | 0.295 | 0.328 | 0.392 |
| 17 | 0.286 | 0.318 | 0.381 |
| 18 | 0.278 | 0.309 | 0.371 |
| 19 | 0.272 | 0.301 | 0.363 |
| 20 | 0.264 | 0.294 | 0.356 |
| 25 | 0.24 | 0.27 | 0.32 |
| 30 | 0.22 | 0.24 | 0.29 |

* Adaptada de F. J. Massey, Jr., "The Kolmogorov-Smirnov test for goodness of fit," *J. Amer. Statist. Ass.*, Vol. 46 (1951), p. 70, con autorización del autor y del editor.

**ANEXO B8: VALORES CRITICOS PARA LA PRUEBA DE RACHAS DE WALD
WOLFOWITZ**

(MILLER y MILLER, 1993, pag. 202)

| <i>N</i> | <i>M</i> | A P = 0.05, el número de rachas
es significativo si es: | |
|----------|----------|--|-----------|
| | | menor que | mayor que |
| 2 | 12-20 | 3 | NA |
| 3 | 6-14 | 3 | NA |
| 3 | 15-20 | 4 | NA |
| 4 | 5-6 | 3 | 8 |
| 4 | 7 | 3 | NA |
| 4 | 8-15 | 4 | NA |
| 4 | 16-20 | 5 | NA |
| 5 | 5 | 3 | 9 |
| 5 | 6 | 4 | 9 |
| 5 | 7-8 | 4 | 10 |
| 5 | 9-12 | 4 | NA |
| 5 | 13-18 | 5 | NA |
| 6 | 6 | 4 | 10 |
| 6 | 7-8 | 4 | 11 |
| 6 | 9-12 | 5 | 12 |
| 6 | 13-18 | 6 | NA |
| 7 | 7 | 4 | 12 |
| 7 | 8 | 5 | 12 |
| 7 | 9 | 5 | 13 |
| 7 | 10-12 | 6 | 13 |
| 8 | 8 | 5 | 13 |
| 8 | 9 | 6 | 13 |
| 8 | 10-11 | 6 | 14 |
| 8 | 12-15 | 7 | 15 |

**ANEXO C8: VALORES CRITICOS DE LA PRUEBA DE RANGOS Y SIGNOS DE
WILCOXON**

(OSTLE, 1977, pag. 601)

| <i>n</i> | Nivel de significancia para prueba lateral | | |
|----------|--|-----|------|
| | .025 | .01 | .005 |
| | Nivel de significancia para prueba bilateral | | |
| | .05 | .02 | .01 |
| 6 | 0 | — | — |
| 7 | 2 | 0 | — |
| 8 | 4 | 2 | 0 |
| 9 | 6 | 3 | 2 |
| 10 | 8 | 5 | 3 |
| 11 | 11 | 7 | 5 |
| 12 | 14 | 10 | 7 |
| 13 | 17 | 13 | 10 |
| 14 | 21 | 16 | 13 |
| 15 | 25 | 20 | 16 |
| 16 | 30 | 24 | 20 |
| 17 | 35 | 28 | 23 |
| 18 | 40 | 33 | 28 |
| 19 | 46 | 38 | 32 |
| 20 | 52 | 43 | 38 |
| 21 | 59 | 49 | 43 |
| 22 | 66 | 56 | 49 |
| 23 | 73 | 62 | 55 |
| 24 | 81 | 69 | 61 |
| 25 | 89 | 77 | 68 |
| * | | | |

* Para $n > 25$, T es aproximadamente estándar normal con media $n(n + 1)/4$ y varianza $n(n + 1)(2n + 1)/24$

**ANEXO D8: VALORES CRITICOS PARA LA PRUEBA DE SUMA DE RANGOS DE
WILCOXON - MANN-WITHNEY**
(JURAN y GRINA, 1993, pág. A11-29)

| n_2 | α para prueba
bilateral | α para prueba
unilateral | n_1 (muestra más pequeña) | | | | | | | | | | | | | |
|-------|-----------------------------------|------------------------------------|-----------------------------|---|----|----|----|----|----|----|----|----|-----|-----|--|--|
| | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | | |
| 3 | 0.20 | 0.10 | | 3 | 7 | | | | | | | | | | | |
| | 0.10 | 0.05 | | | 6 | | | | | | | | | | | |
| | 0.05 | 0.025 | | | | | | | | | | | | | | |
| | 0.01 | 0.005 | | | | | | | | | | | | | | |
| 4 | 0.20 | 0.10 | | 3 | 7 | 13 | | | | | | | | | | |
| | 0.10 | 0.05 | | | 6 | 11 | | | | | | | | | | |
| | 0.05 | 0.025 | | | | 10 | | | | | | | | | | |
| | 0.01 | 0.005 | | | | | | | | | | | | | | |
| 5 | 0.20 | 0.10 | | 4 | 8 | 14 | 20 | | | | | | | | | |
| | 0.10 | 0.05 | | 3 | 7 | 12 | 19 | | | | | | | | | |
| | 0.05 | 0.025 | | | 6 | 11 | 17 | | | | | | | | | |
| | 0.01 | 0.005 | | | | | 15 | | | | | | | | | |
| 6 | 0.20 | 0.10 | | 4 | 9 | 15 | 22 | 30 | | | | | | | | |
| | 0.10 | 0.05 | | 3 | 8 | 13 | 20 | 28 | | | | | | | | |
| | 0.05 | 0.025 | | | 7 | 12 | 18 | 26 | | | | | | | | |
| | 0.01 | 0.005 | | | | 10 | 16 | 23 | | | | | | | | |
| 7 | 0.20 | 0.10 | | 4 | 10 | 16 | 23 | 32 | 41 | | | | | | | |
| | 0.10 | 0.05 | | 3 | 8 | 14 | 21 | 29 | 39 | | | | | | | |
| | 0.05 | 0.025 | | | 7 | 13 | 20 | 27 | 36 | | | | | | | |
| | 0.01 | 0.005 | | | | 10 | 16 | 24 | 32 | | | | | | | |
| 8 | 0.20 | 0.10 | | 5 | 11 | 17 | 25 | 34 | 44 | 55 | | | | | | |
| | 0.10 | 0.05 | | 4 | 9 | 15 | 23 | 31 | 41 | 51 | | | | | | |
| | 0.05 | 0.025 | | 3 | 8 | 14 | 21 | 29 | 38 | 49 | | | | | | |
| | 0.01 | 0.005 | | | | 11 | 17 | 25 | 34 | 43 | | | | | | |
| 9 | 0.20 | 0.10 | 1 | 5 | 11 | 19 | 27 | 36 | 46 | 58 | 70 | | | | | |
| | 0.10 | 0.05 | | 4 | 9 | 16 | 24 | 33 | 43 | 54 | 66 | | | | | |
| | 0.05 | 0.025 | | 3 | 8 | 14 | 22 | 31 | 40 | 51 | 62 | | | | | |
| | 0.01 | 0.005 | | | 6 | 11 | 18 | 26 | 35 | 45 | 56 | | | | | |
| 10 | 0.20 | 0.10 | 1 | 6 | 12 | 20 | 28 | 38 | 49 | 60 | 73 | 87 | | | | |
| | 0.10 | 0.05 | | 4 | 10 | 17 | 26 | 35 | 45 | 56 | 69 | 82 | | | | |
| | 0.05 | 0.025 | | 3 | 9 | 15 | 23 | 32 | 42 | 53 | 65 | 78 | | | | |
| | 0.01 | 0.005 | | | 6 | 12 | 19 | 27 | 37 | 47 | 58 | 71 | | | | |
| 11 | 0.20 | 0.10 | 1 | 6 | 13 | 21 | 30 | 40 | 51 | 63 | 76 | 91 | 106 | | | |
| | 0.10 | 0.05 | | 4 | 11 | 18 | 27 | 37 | 47 | 59 | 72 | 86 | 100 | | | |
| | 0.05 | 0.025 | | 3 | 9 | 16 | 24 | 34 | 44 | 55 | 68 | 81 | 96 | | | |
| | 0.01 | 0.005 | | | 6 | 12 | 20 | 28 | 38 | 49 | 61 | 73 | 87 | | | |
| 12 | 0.20 | 0.10 | 1 | 7 | 14 | 22 | 32 | 42 | 54 | 66 | 80 | 94 | 110 | 127 | | |
| | 0.10 | 0.05 | | 5 | 11 | 19 | 28 | 38 | 49 | 62 | 75 | 89 | 104 | 120 | | |
| | 0.05 | 0.025 | | 4 | 10 | 17 | 26 | 35 | 46 | 58 | 71 | 84 | 99 | 115 | | |
| | 0.01 | 0.005 | | | 7 | 13 | 21 | 30 | 40 | 51 | 63 | 76 | 90 | 105 | | |