UNIVERSIDAD DE EL SALVADOR

ELEMENTOS DE ESTADISTICA MATEMATICA

(Tesis Doctoral)

ING. RODOLFO JENKINS G.

FACULTAD DE ECONOMIA

Instituto de Estudios Económicos

(1962)

P R O L O G O

Una de las labores que el Instituto de Estudios Económicos de la Facultad de Economía ha considerado de gran importancia para el buen desarrollo de la docencia en la misma, ha sido la reproducción de libros, monografías y artículos, cuyos tirajes se han agotado o son demasiado limitados para llegar a todos los alumnos.

En esta oportunidad, gracias a la autorización del Ing. Rodolfo Jenkins G., Profesor de Estadística Metodológica de la Facultad de Economía, reproducimos su tesis "ELEMENTOS DE ESTADISTICA MATEMATICA APLICADA A LA INGENIERIA CIVIL", que no dudamos será de gran utilidad para todos los alumnos, especialmente para los que están cursando Estadística Metodológica.

CAPITULO I

SERIES ESTADISTICAS Y SU CLASIFICACION

Cuando se ha compilado una serie de datos cuantitativos, para fines estadísticos, invariablemente tales datos pertenecerán a uno u otro de los siguientes grupos:

1) Datos independientes del factor tiempo

2) Datos que varían con el tiempo

Los incluidos en el primer grupo constituyen una "serie cuantitativa no-cronológica"; los incluidos en el segundo, una "serie cuantitativa cronológica" (llamada simplemente "serie cronológica").

Ejemplos:

a) Serie cuantitativa no-cronológica
Resistencia de las varillas de hierro redondo de 25 mm. de diámetro (Pruebas en laboratorios alemanes, con carga para rotura a tracción)

Varilla No.	Kg/cm^2
1	3850
2	3835
3	3857
4	3860
5	3855
6	3867
7	3850
8	3855

b) Serie cronológica
Hierro estructural importado por El Salvador (Período 1950-1954)

Año	Toneladas métricas
1950	5528
1951	5619
1952	5800
1953	5760
1954	11486

CAPITULO II

FORMA DE PRESENTAR LAS SERIES CUANTITATIVAS NO-CRONOLOGICAS

Una serie cuantitativa no-cronológica, en su forma más simple, consiste en una lista de datos dispuestos en cualquier orden y sin clasificación alguna. Generalmente es denominada "serie simple" (Cuadro 1).

CUADRO 1

APERTURA DE LAS PLICAS PRESENTADAS EN LA LICITACION No. 52, PROMOVIDA POR EL MINISTERIO DE OBRAS PUBLICAS

(Ofertas hechas por 28 constructores: Arquitectos, Ingenieros y Personas Jurídicas)

Oferente No.	Valor de la oferta (en miles de colones)	Oferente No.	Valor de la oferta (en miles de colones)
1	233	15	205
2	214	16	270
3	292	17	220
4	206	18	225
5	210	19	240
6	220	20	160
7	250	21	212
8	147	22	170
9	230	23	195
10	240	24	200
11	200	25	216
12	200	26	196
13	188	27	165
14	190	28	190

Obsérvese que una serie de este tipo tiene varios inconvenientes:

a) Ocupa mucho espacio: la serie de las lecturas de los "medidores" del consumo de agua potable en San Salvador, sería gigantesca;

b) Es molesto encontrar el término más bajo, o el más alto, con la rapidez deseada;

c) Es difícil, e imposible a veces, encontrar ciertos valores medios (muy útiles en Estadística como se verá más adelante).

Por eso es preferible (y necesario en ciertos casos) ordenar los datos en forma creciente, o decreciente, según su magnitud (Cuadro 2).

CUADRO 2

VALORES DEL CUADRO 1 ORDENADOS EN FORMA CRECIENTE

(en miles de colones)

147	195	210	230
160	196	212	233
165	200	214	240
170	200	216	240
188	200	220	250
190	205	220	270
190	206	225	292

El cuadro de los datos ordenados nos permite apreciar, sin mayor -esfuerzo, lo siguiente:

1) Los términos menor y mayor son, respectivamente, 147 y 292 -- (que en adelante serán llamados "fronteras" de la serie; y su diferencia aritmética será la "oscilación" de la serie);

2) El término más frecuente es 200, puesto que aparece repetido mayor número de veces (es el término modal o "moda");

3) Las ofertas tienden a concentrarse entre 200 y 220. En este grupo, por lo tanto, debe estar el "valor medio". (En Estadística se utilizan varios tipos de "valor medio": uno de ellos es el "promedio -- aritmético" ó "media aritmética", ampliamente conocido).

A pesar de lo anterior persiste en el Cuadro 2 un inconveniente: muchos números ocupando demasiado espacio. Es posible distribuir los términos en grupos o "clases". Por ejemplo: las ofertas 160, 165 y 170 mil colones pueden ser agrupadas en la clase "ofertas de 150 mil a menos de 175 -- mil colones". Siguiendo este criterio formaremos una "distribución de frecuencias" (llamada también "serie de clases y frecuencias") como se muestra en el Cuadro 3.

CUADRO 3

DISTRIBUCION DE FRECUENCIAS CORRESPONDIENTE A LA SERIE SIMPLE DEL CUADRO 2

Valor de la oferta (en miles de colones)	No. de ofertas
De 125 a menos de 150	1
" 150 " 175	3
" 175 " 200	5
" 200 " 225	11
" 225 " 250	5
" 250 " 275	2
" 275 " 300	1
T O T A L	28

Obsérvese que la la. clase está indicada: "de 125 a menos de 150"; y así sucesivamente. Podríamos haber escrito "de 125 a 149.99999". Ambos sistemas tienen por objeto evitar la duda que surgiría al tratar de clasificar la oferta "150", por ejemplo, en una serie cuyas clases se indicasen así: "de 125 a 150", "de 150 a 175", etc.

Lo importante, en resumen, es que el límite superior de una clase - difiera una cantidad infinitesimal, del límite inferior de la clase siguiente.

Con este agrupamiento en clases hemos sintetizado la amplia información inicial. Obviamente perdimos precisión: ¿ cuál es el monto exacto de - cada una de las cinco ofertas comprendidas en la clase de "175 a menos de -- 200"?. Pero tal dificultad, más aparente que real, queda ampliamente compensada con la facilidad que ofrece toda distribución de frecuencias, para su análisis matemático. Esta afirmación es un axioma cuando se refiere a grandes conjuntos: varios items del próximo Censo de Población de El Salvador, tendrán más de DOS MILLONES de datos o términos.

Hemos visto que en una distribución de frecuencias cada clase está comprendida entre dos valores. Estos son llamados "límites de la clase". - La amplitud entre el límite superior y el límite inferior de una clase se deno mina "intervalo de clase".

Surge una pregunta: ¿cuántas clases debe tener una distribución de frecuencias?. O dicho de otra manera: ¿qué amplitud debe tener el intervalo de clase?

Sumamente útil en este sentido es la fórmula experimental de Sturges:

$$i = \frac{F_s - F_i}{1 + 3.322 \log_{10} N}$$

en la cual:

i = intervalo de clase

F_s = frontera superior de la serie

F_i = frontera inferior de la serie

N = número de términos de la serie

Esta fórmula puede dar resultado entero o fraccionario. En el último caso conviene redondear la cifra (eliminando la fracción decimal menor de 0.5 o aumentando la fracción mayor de 0.5 hasta la unidad inmediata superior).

El intervalo usado en el Cuadro 3 (25 mil colones) fue calculado con la fórmula de Sturges:

$$i = \frac{292 - 147}{1 + 3.322 \log 28} = 24.96$$

$$i = 25$$

CAPITULO III

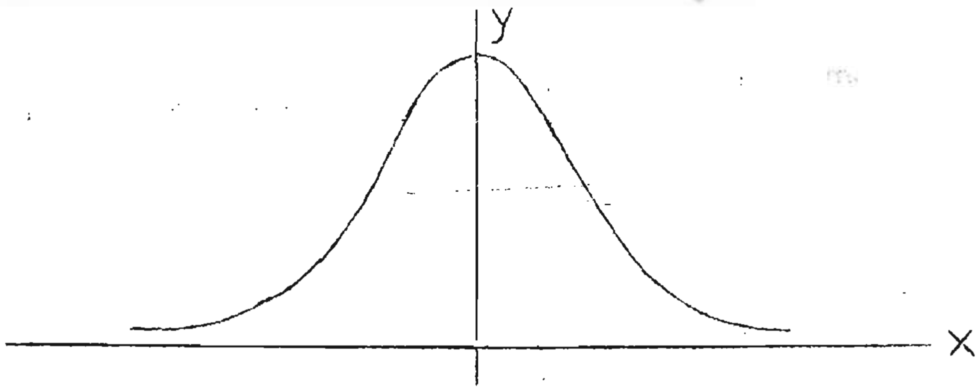
ANALISIS DE LAS SERIES CUANTITATIVAS NO-CRONOLOGICAS PROMEDIOS

Las series simples correspondientes a grandes conjuntos (lo cual implica existencia de gran número de términos) poseen varias características comunes, independientemente del fenómeno que cada serie represente.

Si con tales series se construyen las respectivas distribuciones de frecuencias y luego se dibujan las gráficas correspondientes, el resultado será siempre una curva que se aproxima notablemente a la curva Normal de Probabilidad (curva de De Moivre o curva de Gauss). Véase Gráfica 1.

GRAFICA 1

CURVA NORMAL DE ERROR
(DE "DE MOIVRE" O DE "GAUSS")

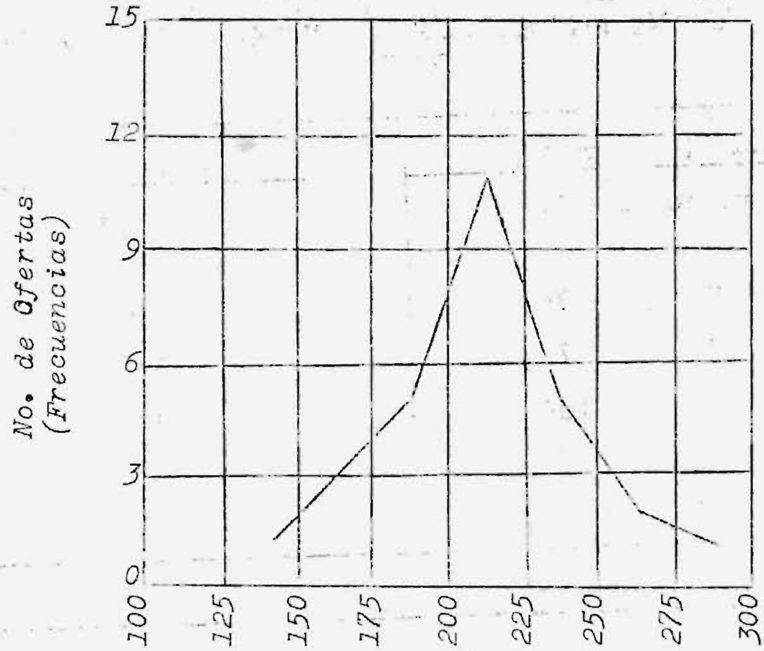


Lo anterior es debido al hecho de que los fenómenos de conjunto no - dependen de la voluntad del hombre: aunque éste tenga ingerencia en algunos de ellos, es el azar el elemento predominante en la formación y distribución de los mismos. Y el azar se rige por la ley de probabilidad, cuya expresión matemática es el desarrollo binomial. La curva normal no es más que la representación gráfica del desarrollo de la expresión $(\frac{1}{2}\, a + \frac{1}{2}\, b)^n$, cuando n se vuelve infinitamente grande (como lo demuestra el Algebra).

Cuando el número de términos de una serie es relativamente pequeño, la curva correspondiente a su distribución de frecuencias difiere de la curva de Gauss; pero en general, sigue la forma de esta última. Tal es el caso de - la distribución de frecuencias mostrada en el Cuadro 3, cuya curva aparece en la Gráfica 2.

GRAFICA 2

CURVA CORRESPONDIENTE A LA DISTRIBUCION DE FRECUENCIAS DEL CUADRO 3



Valor de la oferta
en miles de colones
(Clases)

PROMEDIOS

En cualquier serie cuantitativa no-cronológica que contenga bastantes términos, éstos tienden a concentrarse alrededor de un valor central (Véase los Cuadros 3 y 4). Lo cual significa que existen determinados valores que - se presentan a menudo, mientras que otros ocurren esporádicamente.

Por consiguiente los términos más característicos están situados en el centro de la distribución. Y entre estos valores característicos hay algunos, llamados "promedios", que son verdaderos respresentantes de la serie que los contiene.

A continuación estudiaremos los promedios más usuales en Estadística.

CUADRO 4

RESISTENCIA A LA COMPRESION DE 62 CUBOS DE HORMIGON HECHOS DE UNA MISMA MEZCLA Y CON EL MISMO TIEMPO DE FRAGUADO

Resistencia (Kg/cm^2)	No. de cubos
De 155 a menos de 160	4
" 160 " 165	9
" 165 " 170	10
" 170 " 175	24
" 175 " 180	12
" 180 " 185	2
" 185 " 190	1
T O T A L	62

Media Aritmética

Es un valor tal que si con él se substituyen los términos de una serie, se puede obtener una suma igual a la que los propios términos darían.

La designaremos "M".

Sea la serie: $X_1 + X_2 + X_3 + \ldots\ldots\ldots + X_n$ (1)

substituyendo cada término por M:

$$M + M + M + \ldots\ldots\ldots + M \quad (2)$$

igualando las sumas (1) y (2):

$$X_1 + X_2 + X_3 + \ldots\ldots\ldots + X_n = M + M + M + \ldots\ldots + M$$

$$X_1 + X_2 + X_3 + \ldots\ldots\ldots + X_n = nM$$

de donde:

$$M = \frac{\sum X}{n} \quad (3)$$

o sea que la Media Aritmética de una serie simple es igual a la suma de los términos dividida entre el número de términos.

Cuando se desea obtener la Media de una distribución de frecuencias, primeramente es necesario representar cada clase por un solo valor (lo cual equivale a substituir cada término de la clase por un término ideal): este valor es el "punto medio" de la clase, que en adelante será denominado Pm. - Conforme la ley de probabilidad cualquier otro término de la clase, distinto del Pm, tendrá menos oportunidad de representar a todos los términos de la clase, puesto que se estará acercando a uno u otro extremo (o límites) de la clase.

Si la frecuencia de la primera clase es f_1, quiere decir que el punto medio de esa clase, P_{m1}, está repetido f_1 veces. Lo mismo sucede con las demás clases.

Entonces:

$$f_1 P_{m1} + f_2 P_{m2} + f_3 P_{m3} + \ldots . f_z P_{mz} = \sum X = nM \quad (4)$$

$$M = \frac{\sum (f \ Pm)}{n} \quad (5)$$

o bien:

$$M = \frac{\sum (f \ Pm)}{\sum f} \quad (6)$$

Ejemplo:

Hallar la media aritmética para la distribución de frecuencias del Cuadro 4.

CUADRO 5

MEDIA ARITMETICA DE LA DISTRIBUCION DE FRECUENCIAS DEL CUADRO 4

Clases	f	Pm	f Pm
De 155 a menos de 160	4	157.5	630.0
" 160 " 165	9	162.5	1462.5
" 165 " 170	10	167.5	1675.0
" 170 " 175	24	172.5	4140.0
" 175 " 180	12	177.5	2130.0
" 180 " 185	2	182.5	365.0
" 185 " 190	1	187.5	187.5
TOTAL = n = $\sum f$	62	TOTAL	10590.0

$$M = \frac{\sum (f \ Pm)}{\sum f} = \frac{10590}{62} = 170.81 \quad (7)$$

Lo anterior significa que la resistencia media a la compresión, de los cubos de hormigón ensayados, es de 170.81 kilogramos por centímetro cuadrado.

Los datos originales, con los cuales fue formada la distribución de frecuencias del Cuadro anterior, son los siguientes:

155	163	167	170	172	174	177
157	164	168	171	173	174	178
157	164	168	171	173	175	178
159	164	169	171	173	175	179
160	165	169	172	173	176	179
161	166	170	172	173	176	181
162	166	170	172	173	176	184
162	166	170	172	174	176	187
162	167	170	172	174	177	

Aplicando a esta serie la fórmula (3), obtenemos su media aritmética:

$$M = \frac{10564}{62} = 170.39 \text{ Kg/cm}^2 \qquad (8)$$

Comparando los resultados (7) y (8) vemos que la pérdida de exactitud en el valor de la Media, ocasionada por el agrupamiento en clases, es insignificante. En el presente caso el porcentaje de error es:

$$e = \frac{170.81 - 170.39}{170.39} \times 100 = 0.25\%$$

Mediana

Es un valor tal que si los términos de una serie se ordenan conforme sus valores crecientes o decrecientes, ocupa el lugar central.

La designaremos "M_d".

Sea la serie: $X_1 < X_2 < X_3 < X_4 < X_5$

El término X_3 es la mediana, por ocupar el lugar central.

Según la definición anterior, el número de términos situados antes de la mediana es igual al número de términos situados después de la mediana.

Ejemplo: Hallar la mediana de la serie 7, 9, 2, 17, 5.

Colocando los términos en orden creciente:

2
5
7
9
17

Vemos que el término central es 7.

Luego:

$$Md = 7$$

En el sencillo ejemplo anterior la mediana es el tercer término de la serie, ó sea que su número de orden es 3. En general, el número de orden correspondiente a la mediana puede ser hallado como sigue:

Sea N_o el número de orden

Número de términos situados antes de la mediana $= N_o - 1$
Número de términos situados después de la mediana $= N_o - 1$

Entonces, el número total de términos de la serie será:

$$(N_o - 1) + 1 + (N_o - 1) = n$$

de donde:

$$N_o = \frac{n + 1}{2} \qquad (9)$$

Cuando $\underline{n}$ es impar el cociente anterior es entero; por consiguiente, en este caso la mediana es uno de los términos de la serie.

Cuando $\underline{n}$ es par el cociente es fraccionario; entonces la mediana no es uno de los términos de la serie. En este caso la mediana es igual a la semisuma de los términos adyacentes al número de orden.

Para obtener la Mediana de una distribución de frecuencias, es necesario sumar a la frecuencia de cada clase las frecuencias de las clases anteriores (formando así una frecuencia acumulada). Esto permite apreciar rápidamente en que clase se halla comprendida la mediana.

Ejemplo: hallar la mediana de la distribución de frecuencias del Cuadro 3.

Clases	f	frecuencias acumuladas
De 125 a menos de 150	1	1
" 150 " 175	3	4
" 175 " 200	5	9
" 200 " 225	11	20
" 225 " 250	5	25
" 250 " 275	2	27
" 275 " 300	1	28
TOTAL	28	

$$N_o = \frac{28 + 1}{2}$$

$$N_o = 14.5$$

N_o 14.5 significa que hay catorce términos antes de la mediana. Según muestra la columna "frecuencias acumuladas", en las tres primeras clases hay sólo 9 términos. Por consiguiente la clase "de 200 a menos de 225" es la que contiene a la mediana (puesto que en las cuatro primeras clases hay más de catorce términos).

La mediana es el término número 14.5 de la serie; pero dentro de la clase que la contiene es el término número 5.5 (puesto que los otros 9 términos están en las tres primeras clases).

Basándonos nuevamente en la ley de probabilidad, supondremos que las frecuencias se distribuyen uniformemente dentro de cada clase. Entonces:

$$\frac{11}{25} = \frac{5.5}{x}$$

$$x = 12.5$$

Luego, la mediana es:

$$Md = 200 + 12.5$$
$$Md = 212.5$$

La fórmula que se ofrece a continuación está basada en el razonamiento anterior:

$$M_d = l_i + \frac{i}{f_m} (N_o - f_{aa}) \qquad (10)$$

en la cual:

l_i = límite inferior de la clase que contiene a la mediana

i = intervalo de clase

f_m = frecuencia de la clase que contiene a la mediana

N_o = número de orden de la mediana

f_{aa} = frecuencia acumulada correspondiente a la clase anterior a aquella que contiene a la mediana.

Moda

La moda (o modo) es el valor que más veces se halla repetido en una serie.

La designaremos "M_o"

En la serie simple:

2	7	8	11
4	7	9	12
4	8	10	13
5	8	10	13
6	8	10	14

tenemos:

$$M_o = 8$$

Hay series en las que dos términos diferentes se repiten igual número de veces (siendo este número de veces mayor que las veces que se repite cualquier otro término). Tales series reciben el nombre de "bimodales". Cuando una serie tiene tres o más modas, es denominada "polimodal". Casi siempre estas dos últimas condiciones son consecuencia de factores extraños (casuales o heterogéneos) introducidos en la serie.

La mayoría de los métodos usuales para calcular la moda de una distribución de frecuencias, dan un valor aproximado para este promedio; pero que resulta suficiente para usos prácticos.

El "método de la diferencia" da la siguiente fórmula:

$$M_o = l_i + \frac{\Delta_1}{\Delta_1 + \Delta_2}\, i \qquad (11)$$

en la cual:

l_i = límite inferior de la clase modal

Δ_1 = diferencia entre la frecuencia de la clase modal y la frecuencia de la clase anterior a la modal.

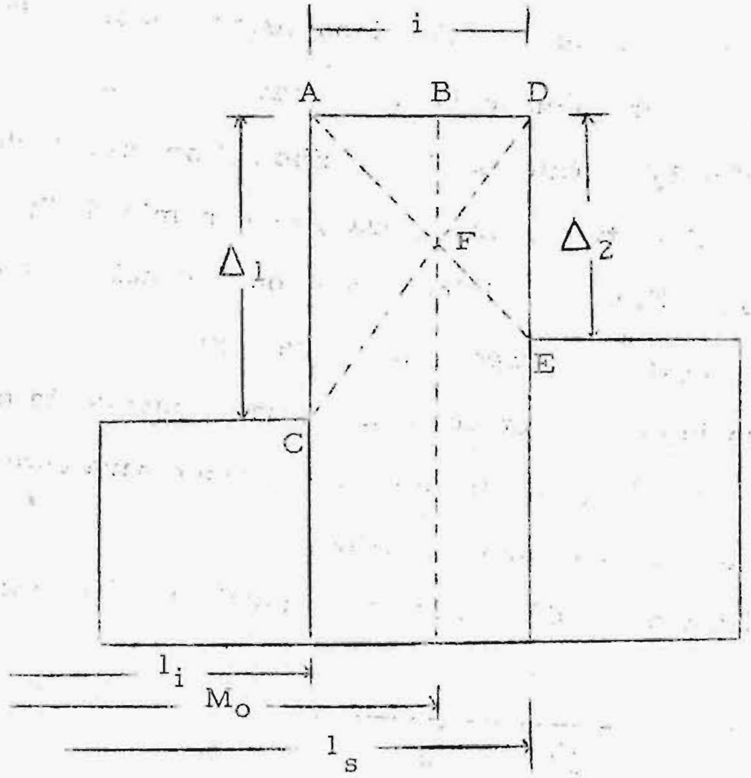
Δ_2 = diferencia entre la frecuencia de la clase posterior a la modal y la frecuencia de la clase modal.

i = intervalo de clase.

La clase modal (correspondiente a la frecuencia mayor) se determina por simple inspección.

La fórmula anterior se deduce partiendo de la hipótesis siguiente: la posición de la moda dentro de la clase modal es influenciada por las diferencias aritméticas (Δ_1 y Δ_2) entre la frecuencia de la clase modal y las frecuencias de las clases adyacentes a la modal.

Gráfica 3



En la gráfica 3, por semejanza de triángulos:

$$\frac{BD}{AD} = \frac{BF}{AC} \qquad (a)$$

$$\frac{BF}{AB} = \frac{DE}{AD} \qquad (b)$$

Combinando (a) y (b):

$$BD = \frac{DE \times AB}{AC}$$

o sea:

$$l_s - M_o = \frac{\Delta_2 (M_o - l_i)}{\Delta_1}$$

$$M_o = \frac{\Delta_1 \quad l_s + \Delta_2 \quad l_i}{\Delta_1 + \Delta_2}$$

pero:

$$l_s = l_i + i$$

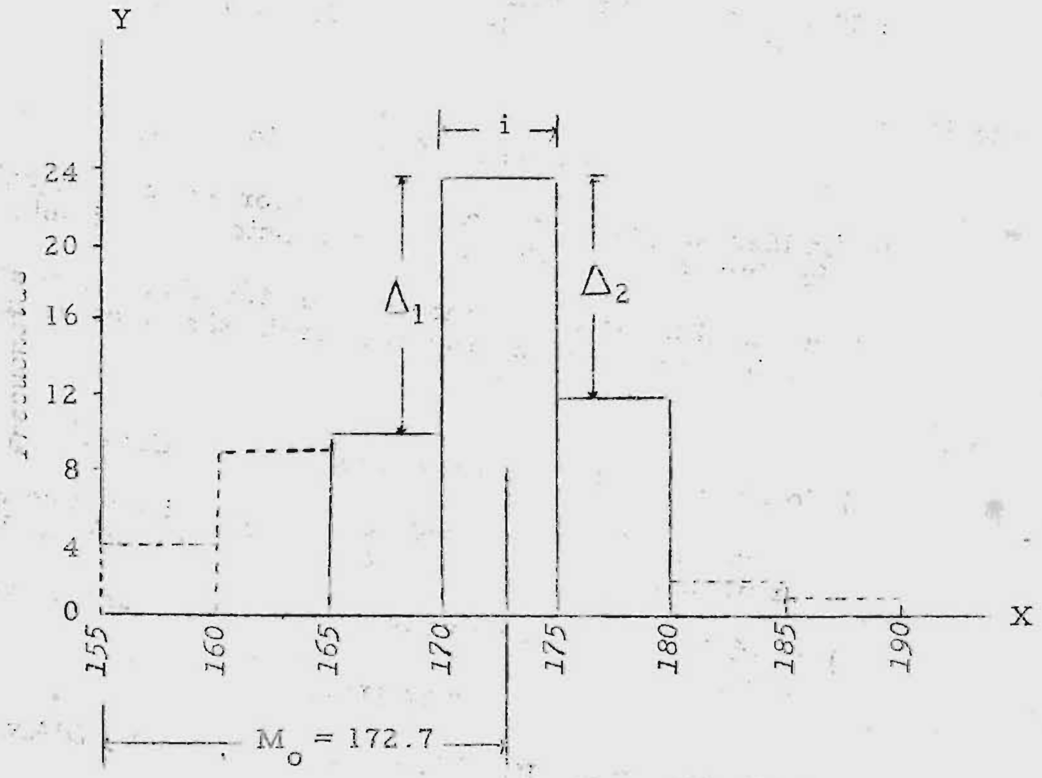
$$M_o = \frac{\Delta_1 \quad l_i + \Delta_1 i + \Delta_2 \quad l_i}{\Delta_1 + \Delta_2}$$

$$M_o = \frac{(\Delta_1 + \Delta_2)\ l_i}{\Delta_1 + \Delta_2} + \frac{\Delta_1 i}{\Delta_1 + \Delta_2}$$

$$M_o = l_i + \frac{\Delta_1}{\Delta_1 + \Delta_2}\ i$$

GRAFICA 4

MODA CORRESPONDIENTE A LA DISTRIBUCION DE FRECUENCIAS DEL CUADRO 4



Fórmula de Pearson

Conociendo previamente la Media y la Mediana de una distribución de frecuencias (obtenida de un elevado número de términos) es posible calcular la Moda, aplicando la fórmula empírica de Karl Pearson:

$$M_o = M - 3\,(M - M_d) \qquad (12)$$

Media Irracional

Este promedio (del cual la Media Aritmética es un caso particular) -- tiene una aplicación muy amplia. Será estudiado en el CAPITULO IV.

Caracteres de los principales promedios

Media Aritmética:

a) Es un valor matemáticamente exacto, susceptible de operaciones algebraicas.

b) Es afectada excesivamente por cualquier término cuyo valor sea desproporcionado con respecto a los demás términos de la serie.

c) Para calcularla en las series simples, no es necesario ordenar previamente los términos.

d) Tiene extensa aplicación en el análisis de las distribuciones de frecuencias que se aproximan a la Curva de Gauss.

Mediana:

a) No se presta a toda clase de operaciones algebraicas.

b) La idea de que la mediana es un valor que tiene igual número de términos a cada lado, es de concepción muy simple.

c) Su magnitud no es afectada por los términos cuyo valor es desproporcionado con respecto a los demás términos de la serie.

Moda:

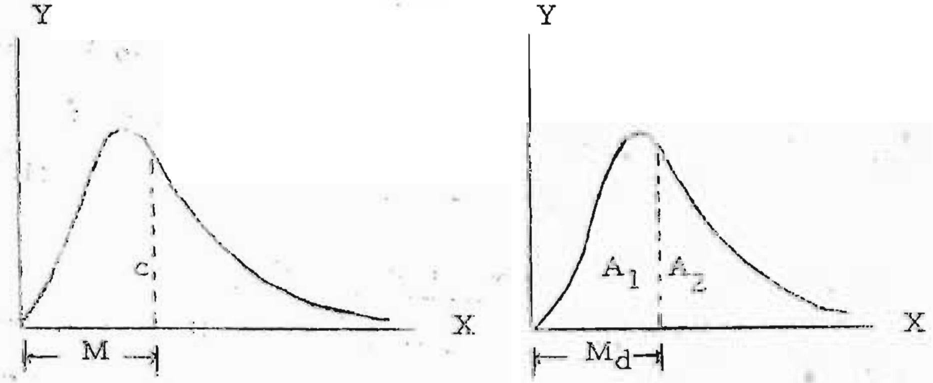
a) No es susceptible de operaciones algebraicas.

b) Su magnitud no es afectada por los términos cuyo valor es desproporcionado con respecto a los demás términos de la serie.

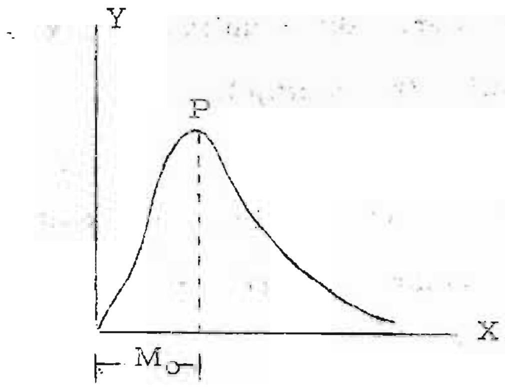
c) Es fácil encontrar su valor aproximado; pero es muy prolijo hallar su valor exacto.

GRAFICA 5

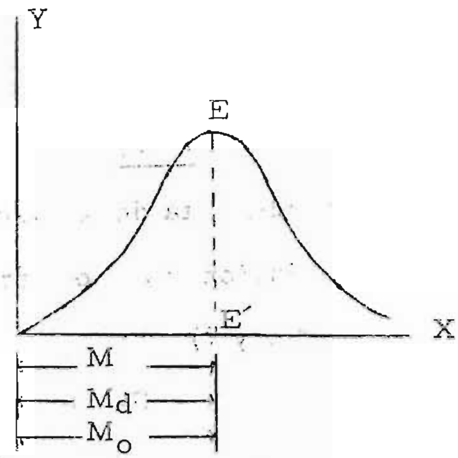
REPRESENTACION GRAFICA DE LA MEDIA, LA MEDIANA Y LA MODA.



c = Centro de gravedad   Areas iguales $A_1 = A_2$



P = punto de inflexión
(elevación máxima)



EE´ = eje de simetría
(caso de una distribución normal)

CAPITULO IV

ANALISIS DE LAS SERIES CUANTITATIVAS NO-CRONOLOGICAS DISPERSION

En el capítulo anterior dijimos que un promedio es un valor capaz de representar a la serie de la cual fue obtenido. Esta afirmación es verdadera hasta cierto punto, puesto que dos o más series diferentes pueden tener igual promedio.

Las series escritas a continuación tienen todas igual Media Aritmética:

A	B	C	D
4	1	3	1
5	1	4	1
6	4	4	1
7	4	9	2
8	20	10	25
M = 6	M = 6	M = 6	M = 6

Existe, sinembargo, una diferencia. En unas series (la "A" por ejemplo) cada término difiere muy poco de la Media; en otras (la "D" por ejemplo) cada término difiere bastante de la Media.

Esta desviación de los términos de una serie con respecto a su promedio se denomina "variación" ó "dispersión" y puede ser medida en forma exacta.

Cuando el grado o magnitud de la dispersión es pequeño, el promedio es verdaderamente representativo de la serie. Cuando tal magnitud es muy -

grande, el promedio tiene escasa significación.

La magnitud de la dispersión puede ser expresada en unidades absolutas (colones, kilogramos, etc.) o en valor relativo (porcentaje).

Desviación Media

Es una medida absoluta de la dispersión. Su valor es igual a la Media Aritmética de las desviaciones de los términos de una serie con respecto a su promedio. (Cuadros 6 y 7).

CUADRO 6

DESVIACION MEDIA DE UNA SERIE SIMPLE
(SALARIOS DE SIETE TRACTORISTAS)

Salario por día (¢)	d
7	2
8	1
8	1
9	0
10	1
12	3
12	3
$\sum d$ =	11

Mediana = 9
d = desviaciones de cada término -- respecto al promedio.
D.M. = desviación media.

$$D.M. = \frac{\sum d}{n} = \frac{11}{7} = 1.57 \text{ colones.}$$

CUADRO 7

DESVIACION MEDIA DE LA SERIE DE FRECUENCIAS DEL CUADRO 4

Clases	f	Pm	d $(P_m - M)$	fd
De 155 a menos de 160	4	157.50	13.31	53.24
" 160 " 165	9	162.50	8.31	74.79
" 165 " 170	10	167.50	3.31	33.10
" 170 " 175	24	172.50	1.69	40.50
" 175 " 180	12	177.50	6.70	80.40
" 180 " 185	2	182.50	11.69	23.38
" 185 " 190	1	187,50	16.69	16.69
TOTALES	62	1207.5		322.10

M = 170.81
d = desviaciones de los puntos medios respecto al promedio.

$$D.M. = \frac{\sum fd}{n}$$

$$D.M. = \frac{322.10}{62} = 5.20\ Kg/cm^2$$

Desviación Típica o "Standard"

Es una medida absoluta de la desviación. Su valor es igual a la raíz cuadrada de la Media Aritmética de los cuadrados de las desviaciones. Se designa comunmente por la letra griega sigma:

$$\sigma = \sqrt{\frac{\sum d^2}{n}} \qquad (15)$$

Algebraicamente, la Desviación Típica es la Media Cuadrática de las desviaciones. Por consiguiente, la fórmula (15) puede ser deducida partiendo de la Media Irracional.

La Media Irracional es un valor que si elevado a una potencia determinada substituye a los términos de una serie elevados a la misma potencia, da una suma igual a la que darían los términos así elevados.

La designaremos "M_i"

Sea la serie $X_1, X_2, X_3, \ldots\ldots X_n$

Elevando a m todos los términos:

$$X_1^m,\ X_2^m,\ X_3^m,\ \ldots\ldots X_n^m \qquad (A)$$

Substituyendo cada término por M_i^m

$$M_i^m,\ M_i^m,\ M_i^m,\ \ldots\ldots M_i^m \qquad (B)$$

Por definición:

$$M_i^m + M_i^m + M_i^m + \ldots\ldots M_i^m = X_1^m + X_2^m + X_3^m + \ldots\ldots X_n^m$$

de donde:

$$nM_i^m = X_1^m + X_2^m + X_3^m + \ldots\ldots X_n^m$$

$$M_i = \sqrt[m]{\frac{X_1^m + X_2^m + X_3^m + \ldots\ldots X_n^m}{n}} \qquad (16)$$

Si en la fórmula (16) de la Media Irracional hacemos $m = 1$:

$$M_i = \frac{\sum X}{n} = \text{Media Aritmética}$$

Si hacemos $m = 2$:

$$M_i = \sqrt{\frac{\sum x^2}{n}} = \text{Media Cuadrática} \qquad (17)$$

Si hacemos $m = 3$:

$$M_i = \sqrt[3]{\frac{\sum x^3}{n}} = \text{Media Cúbica}$$

Aplicando la fórmula (17) a las desviaciones:

$$\sigma = \sqrt{\frac{\sum d^2}{n}}$$

CUADRO 8

DESVIACION TIPICA DE UNA SERIE SIMPLE (PESO DEL HORMIGON ARMADO, EN KILOGRAMOS POR METRO CUBICO, OBTENIDO DE NUEVE MIEMBROS ESTRUCTURALES PREFABRICADOS).

Peso	d	d^2
2385	- 12	144
2385	- 12	144
2386	- 11	121
2390	- 7	49
2390	- 7	49
2400	3	9
2400	3	9
2415	18	324
2422	25	625
TOTAL	0	1474

$M = 2397$ Kg/m^3

$n = 9$

$$\sigma = \sqrt{\frac{\sum d^2}{n}}$$

$$\sigma = \sqrt{\frac{1474}{9}} = 12.8 \text{ Kg/m}^3$$

Obsérvese en el cuadro anterior que la suma algebraica de las desviaciones de los términos de la serie respecto a su Media Aritmética, es igual a cero. Esta es una de las propiedades de dicho valor medio.

En una distribución de frecuencias, la Desviación Típica se calcula con la fórmula:

$$\sigma = \sqrt{\frac{\sum fd^2}{n}} \qquad (18)$$

en la cual, como anteriormente:

$$n = \sum f$$

CUADRO 9

DESVIACION TIPICA EN UNA SERIE DE FRECUENCIAS
(COSTO DEL METRO CUADRADO DE AREA EDIFICADA, OBTENIDO EN LA CONSTRUCCION DE 97 VIVIENDAS TIPO MEDIO, SISTEMA "MIXTO" DE UN SOLO PISO)

M = 128

Costo en ₡	f	Pm	d	d^2	fd^2
De 100 a menos de 110	3	105	− 23	529	1587
" 110 " 120	13	115	− 13	169	2197
" 120 " 130	40	125	− 3	9	360
" 130 " 140	34	135	7	49	1666
" 140 " 150	7	145	17	289	2023
	97				7833

$$\sigma = \sqrt{\frac{7833}{97}} = 9 \text{ colones.}$$

Características de la Desviación Típica

a) Es un valor susceptible de operaciones algebraicas (lo cual no sucede con la D.M.)

b) Se halla influída por la magnitud de cada uno de los términos de la serie (en lo cual se asemeja a la D.M.).

c) Es de gran utilidad en el estudio de la Cruva Normal.

d) Como todo valor absoluto (incluso la D.M.) se presta poco para comparar varias series entre sí. Pero relacionada con la Media es imprescindible en tal trabajo.

Coeficiente de dispersión

Es una medida relativa de la dispersión. Su valor se expresa en tanto por ciento de la Media Aritmética:

$$C = \frac{\sigma}{M} \times 100$$

Es la más importante entre las medidas de la dispersión, por las siguientes razones:

a) Es susceptible de operaciones algebraicas.

b) Puesto que es un valor abstracto, permite comparar entre sí dos o más series expresadas en unidades de distintas clases.

c) Permite establecer el grado en que la Media Aritmética es representativa de una serie determinada.

Respecto a la cualidad "c", es muy útil la tabla siguiente:

Valor del coeficiente de dispersión	Grado en que la Media representa a la serie
De 0 a menos de 10%	Media altamente representativa
" 10 " 20%	Media bastante representativa
" 20 " 30%	Media poco representativa
" 30 " 40%	Media cuya representación es dudosa
" 40% o más	Media carente de significación

Finalmente, conviene saber que las características de una distribución de frecuencias pueden ser resumidas en cuatro elementos:

1) Un promedio
2) Una medida de dispersión
3) Una medida de asimetría
4) Una medida de kurtosis.

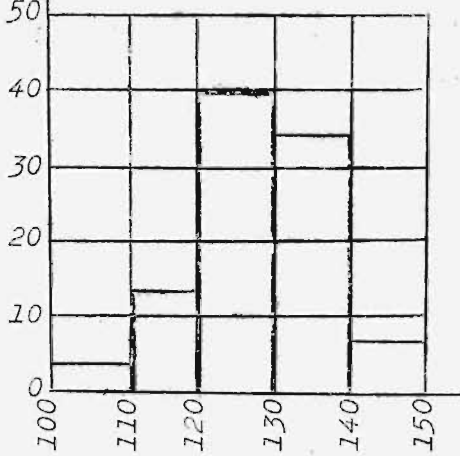
La asimetría (disposición simétrica o asimétrica de los términos situados a uno y otro lado del valor medio) y la kurtosis (grado de achatamiento -- de una curva de frecuencias en relación a la Curva Normal), están fuera del alcance de este estudio.

CAPITULO V

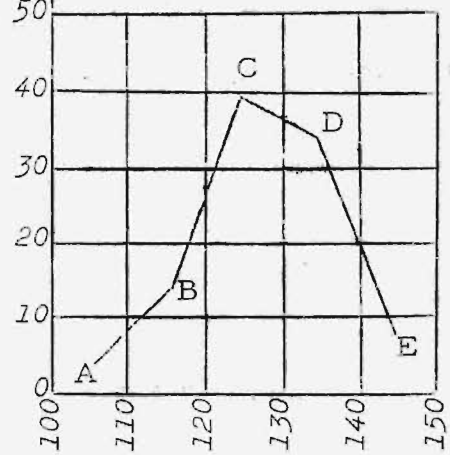
ANALISIS DE LAS SERIES CUANTITATIVAS NO-CRONOLOGICAS
AJUSTE DE UNA CURVA A UNA DISTRIBUCION DE FRECUENCIAS

Hemos visto como, en una distribución de frecuencias, cada clase puede ser representada por su "punto medio". La Gráfica 6 corresponde a la -- distribución de frecuencias del Cuadro 9. La Gráfica 7 corresponde al mismo Cuadro, con la única diferencia de que las clases han sido substituídas - por sus "puntos medios".

GRAFICA 6



GRAFICA 7



A medida que en una distribución de frecuencias se incluyen más datos o términos, debe aumentar el número de clases. Esto puede ser apreciado en la fórmula de Sturges (Página 5): al aumentar "N" disminuye "i" (cuanto menor es el intervalo en una distribución determinada, mayor es el número de clases).

Y al ser cada vez mayor el número de términos y el número de clases, obviamente la línea quebrada A, B, C, D, E, (Gráfica 7) irá suavizándose hasta convertirse en una curva (que será perfectamente contínua cuando N sea infinitamente grande).

Salvo algunas excepciones, dicha curva tendrá la forma aproximada de una campana, siendo por tanto semejante a la Curva Normal o de Gauss.

Las excepciones están constituídas por distribuciones de frecuencias cuya respectiva curva-límite es una parábola, una hipérbola, etc.

En la práctica, N nunca será lo suficiente grande para que podamos obtener una curva perfecta que represente una determinada distribución. Es posible, sinembargo, adaptar o ajustar una curva que nos describa la forma general (o tendencia) que sigue dicha distribución. En otras palabras: una curva que nos muestre como sería tal distribución si N fuese infinitamente grande.

Es lógico que a una distribución de frecuencias cuya gráfica tenga la forma aproximada de una campana, le adaptaremos una Curva Normal (y no -

una hipérbola, por ejemplo).

La ecuación de la Curva Normal (en función de la Desviación Típica) cuando el eje YY´coincide con la Media (Gráfica 1) es:

$$y = y_m e^{-\frac{x^2}{2\sigma^2}} \qquad (19)$$

en la cual:

y_m = ordenada máxima

e = base de los logaritmos neperianos = 2.7183

σ = desviación típica.

La ordenada máxima es:

$$y_m = \frac{Ni}{\sigma\sqrt{2\pi}} \qquad (20)$$

de manera que la ecuación (19) puede ser escrita así:

$$y = \frac{N_i}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \qquad (21)$$

Ejemplo: Ajustar una Curva Normal a la distribución de frecuencias del Cuadro 9.

Tenemos:

$N = 97$

$i = 10$

$\sigma = 9$

La ordenada máxima (ecuación 20) es:

$$y_m = \frac{97 \times 10}{9\sqrt{6.2832}} = 42.94$$

Para poder dibujar la curva ajustada, necesitamos calcular varias ordenadas a ambos lados de la ordenada máxima. Por su cálculo fácil, escogeremos las ordenadas correspondientes a las siguientes abscisas:

$$x = \pm \sigma$$

$$x = \pm 2\sigma$$

$$x = \pm 3\sigma$$

Para $x \pm \sigma$, en la ecuación (19):

$$y = 42.94 \; e^{-\frac{\sigma^2}{2\sigma^2}}$$

$$y = 42.94 \; e^{-\frac{1}{2}} = 26.06$$

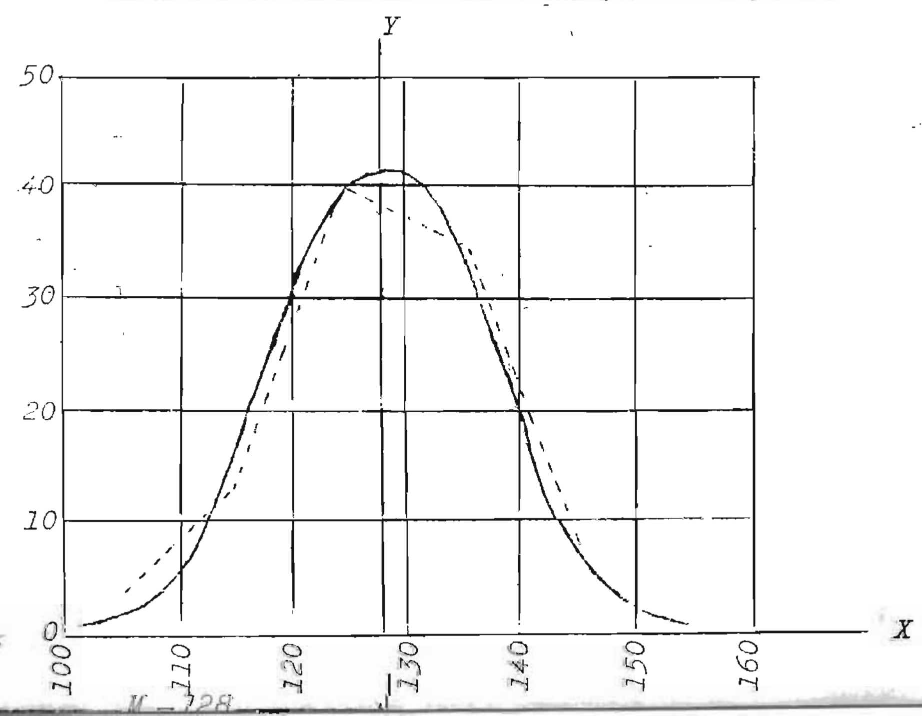
en la misma forma:

Para x	y
$\pm 2\sigma$	5.81
$\pm 3\sigma$	0.48
± 1	42.68
± 3	40.66
± 6	34.31
± 12	17.73
± 15	10.65
± 21	2.84
± 24	1.22

La curva ajustada aparece en la Gráfica 8.

GRAFICA 8

CURVA NORMAL AJUSTADA A LA DISTRIBUCION DE FRECUENCIAS DEL CUADRO 9



El cálculo de las ordenadas de la Curva Normal puede simplificarse usando el Cuadro 10, en el cual se han tabulado los valores de:

$$e^{-\frac{x^2}{2\sigma^2}}$$

en función del cociente $\frac{x}{\sigma}$. Puesto que la ordenada máxima es un factor constante en la ecuación (19); para calcular una ordenada cuya abscisa sea x basta multiplicar dicha constante por el valor tabulado correspondiente a $\frac{x}{\sigma}$

Tablas más detalladas se encuentran en diversos textos sobre esta materia.

CUADRO 10

ORDENADAS DE LA CURVA NORMAL

(EXPRESADAS EN FRACCION DECIMAL DE LA ORDENADA MAXIMA)

ABSCISA = x

$\frac{x}{\sigma}$		$\frac{x}{\sigma}$		$\frac{x}{\sigma}$		$\frac{x}{\sigma}$	
0.00	1.0000	0.85	0.6968	1.70	0.2358	2.55	0.0387
0.05	0.9988	0.90	0.6669	1.75	0.2163	2.60	0.0341
0.10	0.9950	0.95	0.6368	1.80	0.1979	2.65	0.0299
0.15	0.9888	1.00	0.6065	1.85	0.1806	2.70	0.0261
0.20	0.9802	1.05	0.5762	1.90	0.1645	2.75	0.0228
0.25	0.9692	1.10	0.5461	1.95	0.1494	2.80	0.0198
0.30	0.9560	1.15	0.5162	2.00	0.1353	2.85	0.0172
0..35	0.9406	1.20	0.4868	2.05	0.1223	2.90	0.0149
0..40	0.9231	1.25	0.4578	2.10	0.1103	2.95	0.0129
0.45	0.9037	1.30	0.4296	2.15	0.0991	3.00	0.0111
0.50	0.8825	1.35	0.4020	2.20	0.0889	3.20	0.0060
0.55	0.8596	1.40	0.3753	2.25	0.0796	3.40	0.0031
0.60	0.8353	1.45	0.3495	2.30	0.0710	3.60	0.0015
0.65	0.8096	1.50	0.3247	2.35	0.0632	3.80	0.0007
0.70	0.7827	1.55	0.3008	2.40	0.0561	4.00	0.0003
0.75	0.7548	1.60	0.2780	2.45	0.0497	4.50	0.0001
0.80	0.7262	1.65	0.2563	2.50	0.0439	5.00	0.0000

<u>Zona Normal</u>

En las series cuya distribución de frecuencias adopta la forma de l

Curva de Gauss, los términos pueden ser clasificados en 4 grupos:

A) Términos normales.- Son aquellos cuyo valor está comprendido entre M - σ y M + σ

B) Términos infranormales.- Cuyo valor está comprendido entre M - 2 σ y M - σ

C) Términos supranormales.- Cuyo valor esta comprendido entre M + σ y M + 2 σ

D) Términos excesivamente anormales.- Cuyo valor es inferior a M - 2 σ o superior a M + 2 σ

En cada uno de los grupos quedará comprendido el siguiente porcentaje de términos:

Grupo A	68.27%
Grupo B	13.59%
Grupo C	13.59%
Grupo D	4.55%
TOTAL	100.00%

Estos porcentajes corresponden a las áreas limitadas por la Curva - Normal y las ordenadas levantadas en las abscisas que determinan cada grupo.

El área limitada por dos ordenadas cualesquiera puede ser computada por Cálculo Infinitesimal. Pero en los tratados de Estadística aparecen tabuladas las áreas de uso más frecuente, expresadas en tanto por ciento de la superficie total bajo la curva.

Clasificando los costos por metro cuadrado de construcción (usados - en el Cuadro 9) dentro de los cuatro grupos citados, tendremos:

Costos normales: entre 119 y 137 colones
Costos infranormales: entre 110 y 119 colones
Costos supranormales: entre 137 y 146 colones
Costos excesivamente anormales: menos de 110 y más de 146 colones.

CAPITULO VI

ANALISIS DE LAS SERIES CRONOLOGICAS

CALCULO DE LA TENDENCIA SECULAR

Cuando un fenómeno varía en relación al tiempo, la serie que lo describe recibe el nombre de cronológica.

Dicha variación puede presentar tres formas:

a) Serie estática: los valores de sus términos difieren muy poco entre sí, siendo sus variaciones meramente casuales. Ejemplo: la temperatura del cuerpo humano saludable es casi constante todo el tiempo (36.7°); las pequeñas variaciones que presenta son debidas a hechos casuales: ejercicio físico, reposo, etc.

b) Serie dinámica evolutiva: los valores de sus términos difieren bastante entre sí, Sus variaciones generalmente obedecen a causas económico-sociales. Puede ser creciente (Producción anual de cemento en El Salvador) o decreciente (Producción anual de velas de estearina en El Salvador).

c) Serie dinámica no evolutiva: aún cuando sus términos pueden diferir bastante entre sí, no tiende a crecer o decrecer a través de períodos largos de tiempo. Si presenta máximos y mínimos en forma periódica, se denomina "cíclica" (Consumo horario de agua en una ciudad cuya población se mantenga más o menos constante). Si su fluctuación no presenta períodos regulares, se denomina "errática".

Por la índole de este trabajo, estudiaremos únicamente las series dinámicas evolutivas (Cuadro 11).

CUADRO 11

CONSUMO DE ENERGIA ELECTRICA EN EL SALVADOR (PERIODO 1946 - 1956)

Año	Millones de KWH
1946	28.0
1947	33.5
1948	37.3
1949	44.3
1950	51.3
1951	60.6
1952	67.6
1953	74.0
1954	82.0
1955	92.7
1956	102.0

Tendencia Secular

La mayoría de las series dinámicas evolutivas acusan una tendencia definida. Cuando esta tendencia se mantiene durante un período de tiempo relativamente largo, recibe el nombre de "tendencia secular".

Dicho de otro modo: la tendencia secular es la forma general del movimiento suave que presenta una serie evolutiva, en un lapso relativamente largo.

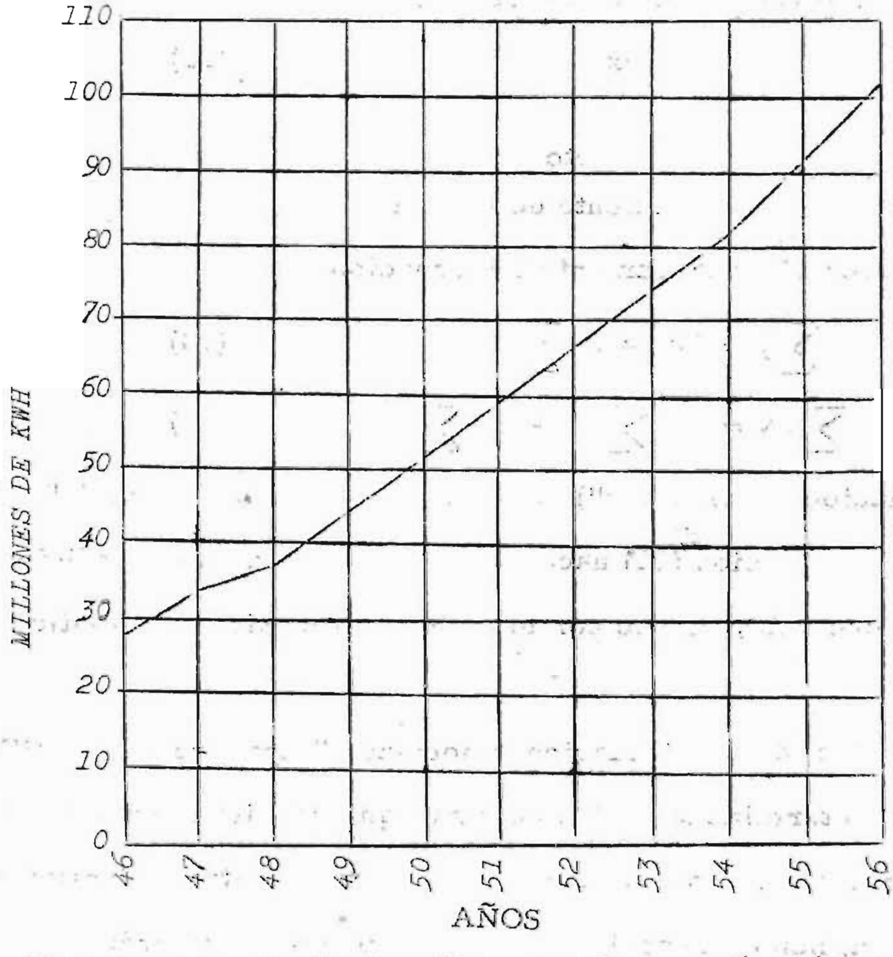
Así como es posible adaptar a cada distribución de frecuencias, una -

curva que indique su forma general; es factible adaptar a cada serie cronológica una curva que represente su tendencia.

El primer paso en el proceso para calcular la tendencia, consiste en dibujar a escala los valores de la serie (Gráfica 9)

GRAFICA 9

CONSUMO DE ENERGIA ELECTRICA EN EL SALVADOR



En este caso particular (Consumo de energía eléctrica en el País) la poligonal correspondiente a la serie dada se aproxima a una recta. En otros fenómenos la poligonal se aproxima a una parábola, a una hipérbola, etc.

Adaptaremos, pues, una línea recta. Esto se deduce de la Gráfica 9. Es conveniente hacer notar, sinembargo, que no siempre es perceptible a -- simple vista la forma general (o tendencia) de una poligonal. En tales casos, la experiencia del analista es un valioso auxiliar.

Para que la recta sea representativa de la tendencia (o mejor dicho: para que exprese matemáticamente la tendencia) debe cumplir dos condiciones:

1) La suma algebraica de las desviaciones debe ser igual a cero (Desviación=ordenada de un punto de la serie menos ordenada de la recta correspondiente a la abscisa de ese punto).

2) La suma de los cuadrados de las desviaciones debe ser un mínimo.

Conforme la Teoría de los Mínimos Cuadrados, al cumplirse el segundo requisito, el primero queda satisfecho automáticamente.

La ecuación general de una recta es:

$$y = a + bx \qquad (22)$$

donde:

a = intercepto
b = pendiente de la recta.

Al resolver simultáneamente las ecuaciones:

$$\sum y = Na + b \sum x \qquad (23)$$

$$\sum xy = a \sum x + b \sum x^2 \qquad (24)$$

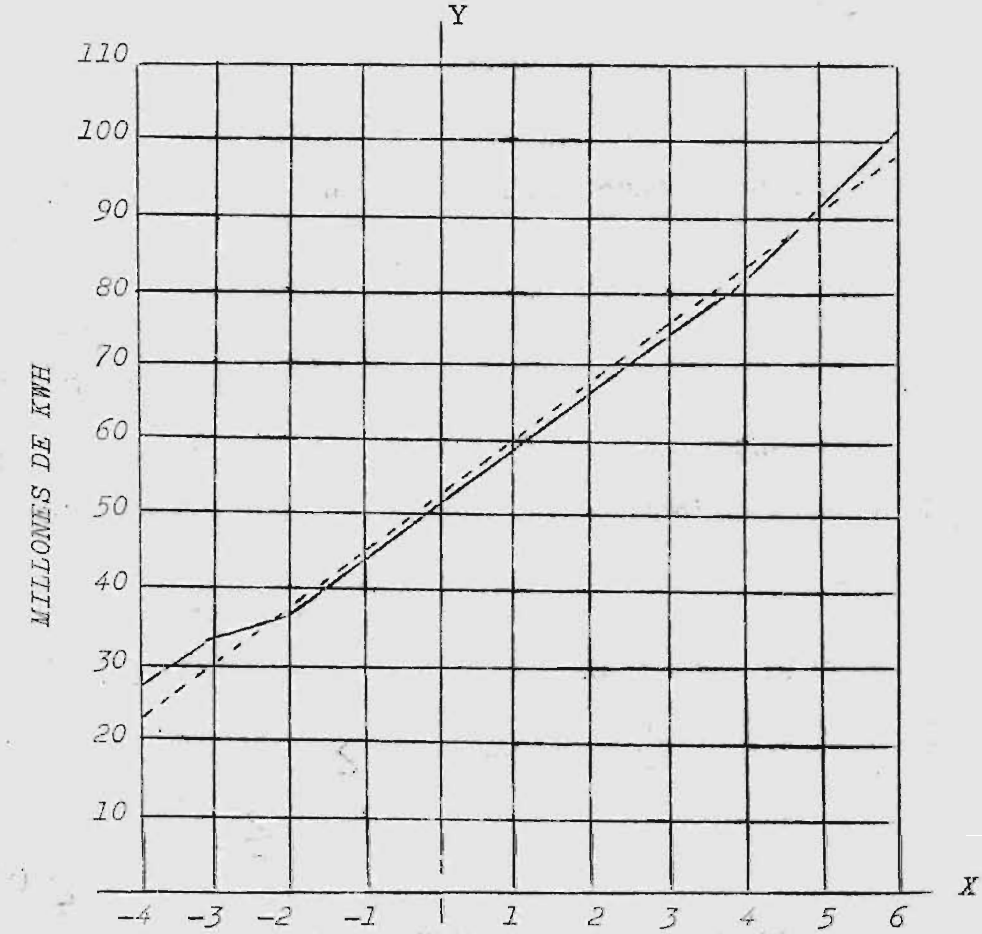
(llamadas "ecuaciones normales") se obtienen los valores para "a" y "b" que substituidos en la ecuación (22) hacen que ésta llene las dos condiciones mencionadas anteriormente, siendo por lo tanto la expresión matemática de la tendencia.

La dedución de las "ecuaciones normales" aparece en el APENDICE.

Una vez determinado el tipo de curva que conviene adaptar a la serie, se escoge un sistema de coordenadas. En el ejemplo citado (Consumo de energía eléctrica) hemos escogido los ejes indicados en la gráfica 10, para facilitar las operaciones algebraicas.

GRAFICA 10

TENDENCIA DE CONSUMO DE ENERGIA ELECTRICA
EN EL SALVADOR



Obsérvese que los años (abscisas) deben ser substituidos por núme-- ros correlativos. Al final del cálculo se harán las equivalencias del caso.

Los valores necesarios para formar las ecuaciones normales han sido tabulados en el Cuadro 12

CUADRO 12

	x	y	xy	x^2
	- 4	28.0	- 112.0	16
	- 3	33.5	- 100.5	9
	- 2	37.3	- 74.6	4
	- 1	44.3	- 44.3	1
	0	51.3	0.0	0
	1	60.6	60.6	1
	2	67.6	135.2	4
	3	74.0	222.0	9
	4	82.0	328.0	16
	5	92.7	463.5	25
	6	102.0	612.0	36
$\sum =$	11	673.3	1489.9	121

N = N° de términos
= 11

Substituyendo estos valores en las ecuaciones (23) y (24):

$$673.3 = 11a + 11b$$
$$1489.9 = 11a + 121b$$

y resolviendo simultáneamente estas últimas:

$$a = 53.79$$
$$b = 7.42$$

Reemplazando estos valores en la ecuación (22), obtenemos:

$$y = 53.79 + 7.42\,x \qquad (25)$$

que es la recta correspondiente a la tendencia secular de la serie dada. Esta recta aparece, con trazos cortos, en la Gráfica 10.

Cuando la representación gráfica de una serie indique que su tendencia no es rectilínea, puede ser adaptada una curva a dicha serie. La ecuación general será en este caso:

$$y = a + bx + cx^2$$

Y las correspondientes ecuaciones normales:

$$\sum(y) = na + b\sum(x) + c\sum(x^2)$$
$$\sum(xy) = a\sum(x) + b\sum(x^2) + c\sum(x^3)$$
$$\sum(x^2y) = a\sum(x^2) + b\sum(x^3) + c\sum(x^4)$$

El procedimiento de cálculo es análogo al de la tendencia rectilínea.

Obsérvese que cuando la mayor potencia de "x" en una ecuación general es 1, las ecuaciones normales son dos; cuando la mayor potencia de "x" es 2, las ecuaciones normales son tres; etc. Puesto que el cálculo para resolver cuatro ó más ecuaciones simultáneas es prolijo, no conviene adaptar curvas cuya ecuación general sea:

$$y = a + bx + cx^2 + dx^3 + \ldots\ldots + \alpha x^n$$

Pero la razón principal para evitar el uso de curvas en las que el exponente de "x" es 4 u otro número mayor, es que una curva de este tipo se adapta a todas las variaciones de los datos y por lo mismo no es capaz de representar la tendencia secular de la serie. Dicho en otra forma: el problema consiste en hallar una curva que indique la tendencia general del fenómeno, y no una curva que pase cerca de todos los puntos de la serie o que una tales puntos.

Una vez que la tendencia ha sido calculada, es posible predecir valores futuros. Haciendo $x = 15$ en la ecuación (25) obtendremos el consumo probable de energía eléctrica para el año 1965:

$$y = 53.79 + 7.42 \times 15$$
$$y = 165.09 \text{ millones de KWH}$$

Es lógico que esta cifra estimada carecerá de valor, si se presenta un suceso imprevisto que interfiera la tendencia (guerra, crisis, desarrollo de grandes planes de industrialización, etc.)

A P E N D I C E

I) METODO BREVE PARA CALCULAR LA MEDIA

$$M = A + \frac{\sum \left[f(Pm-A) \right]}{\sum f} \qquad \text{(a)}$$

donde:

A = valor arbitrario

El cálculo puede ser abreviado aun más, haciendo "A" igual al punto medio de una clase cualquiera. En tal caso, la fórmula (a) se transforma en:

$$M = A + \frac{\left[\sum (fh) \right] \; i}{\sum f} \qquad \text{(b)}$$

donde:

$h = 0$ en la clase donde se hizo $A = Pm$

$h = -1$, $h = -2$, $h = -3$, etc., en las clases sucesivas anteriores a la clase donde se hizo $A = Pm$

$h = 1$, $h = 2$, $h = 3$, etc., en las clases sucesivas posteriores a la clase donde se hizo $A = Pm$

II) METODO BREVE PARA CALCULAR LA DESVIACION TIPICA

En una serie simple:

$$\sigma = \sqrt{\frac{\sum X^2}{N} - \left(\frac{\sum X}{N} \right)^2} \qquad \text{(c)}$$

donde:

X = valor de cada término

N = número total de términos

En una distribución de frecuencias:

$$\sigma = i \sqrt{\frac{\sum (fh^2)}{\sum f} - \left[\frac{\sum (fh)}{\sum f} \right]^2} \qquad \text{(d)}$$

donde:

$h = 0$ en una clase cualquiera

$h = -1$, $h = -2$, $h = -3$, etc., en las clases sucesivas anteriores a la clase donde $h = 0$

$h = 1$, $h = 2$, $h = 3$, etc., en las clases sucesivas posteriores a la clase donde $h = 0$

III) DEDUCCION DE LAS ECUACIONES NORMALES PARA UNA TENDENCIA RECTILINEA

Siendo: y = ordenada de un valor dado, cuya abscisa es "x"

y_c = ordenada de la tendencia (valor calculado) correspondiente a la misma abscisa "x"

la desviación será entonces: $y - y_c$

y su cuadrado: $(y - y_c)^2$

pero $y_c = a + bx$

luego: $(y - y_c)^2 = \left[y - (a + bx) \right]^2 = (y - a - bx)^2$

$$(y - y_c)^2 = y^2 - 2ay - 2bxy + a^2 + 2abx + b^2x^2$$

La suma de los cuadrados de las desviaciones debe ser un mínimo -- (véase Capítulo VI). Esta condición se cumple cuando la primera derivación es igual a cero.

Haciendo la suma de los cuadrados de las desviaciones, derivando parcialmente respecto a "a" e igualando a cero:

$$\sum (y - y_c)^2 = \sum y^2 - 2a \sum y - 2b \sum xy + Na^2 + 2ab \sum x + b^2 \sum x^2$$

$$\frac{d \sum (y - y_c)^2}{da} = -2 \sum y + 2Na + 2b \sum x$$

$$-2 \sum y + 2Na + 2b \sum x = 0$$

$$\sum y = Na + b \sum x$$

que es una de las ecuaciones normales. La otra se obtiene en forma análoga, pero derivando parcialmente con respecto a "b".

IV) UNIDADES SIGMATICAS

Para comparar dos series cronológicas cuyas ordenadas se hayan expresadas en unidades diferentes, es necesario reducir dichas ordenadas a -- una unidad común, denominada "unidad sigmática".

Su expresión es:

$$U = \frac{y - M}{\sigma} \qquad (e)$$

en la cual:

U = unidad sigmática

y = ordenada original (real)

M = media de las ordenadas originales de la serie

σ = desviación típica de las ordenadas originales de la serie, con respecto a la Media.

Una vez reducidas a unidades sigmáticas las ordenadas de ambas series, éstas pueden ser representadas gráficamente, referidas a un eje común de coordenadas.

V) FORMULAS PARA CALCULAR LOS CAMBIOS DE POBLACION

$$P_b = C_a + N + I - D - E \qquad (f)$$

en la cual:

P_b = población calculada para la fecha "b"

C_a = población censada en la fecha "a" (anterior a la fecha "b")

N = nacimientos en el período "b-a"

I = inmigrantes en el período "b-a"

D = defunciones en el período "b-a"

E = emigrantes en el período "b-a"

$$P_b = C_a (1 + k)^n \qquad (g)$$

en la cual:

k = coeficiente de crecimiento

n = período "b-a"

B I B L I O G R A F I A

"METODOS ESTADISTICOS APLICADOS A LA ECONOMIA Y LOS NEGOCIOS".
F. Cecil Mills.

"ESTADISTICA GENERAL APLICADA"
F. Croxton y D. Cowden.

"PRINCIPIOS DE METODOLOGIA ESTADISTICA"
G. Irving Gavett.

"METODOS ESTADISTICOS"
Andrés García Pérez.

En la reproducción de este trabajo participaron, en la labor de mecanografía, revisión y compaginado, la Sta. Carmen G. Escobar, y el Sr. Carlos Alberto Pocasangre López, ambos colaboradores del Instituto de Estudios Económicos de la Facultad de Economía.