

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



Universidad de El Salvador
Hacia la libertad por la cultura

LEY DE BENFORD Y SUS APLICACIONES

PRESENTADO POR:

RAÚL ANTONIO MARTÍNEZ GÁMEZ
CARLOS ERNESTO CANIZALES RIVERA

**PARA OPTAR AL GRADO DE:
LICENCIATURA EN ESTADISTICA**

ASESOR:

Dr. JOSÉ NERYS FUNES TORRES.

CIUDAD UNIVERSITARIA, SAN SALVADOR, FEBRERO 2009.

UNIVERSIDAD DE EL SALVADOR

RECTOR : ING. RUFINO ANTONIO QUEZADA SÁNCHEZ

SECRETARIO

GENERAL : LIC. DOUGLAS VLADIMIR ALFARO CHÁVEZ

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

DECANO : DR. RAFAEL ANTONIO GÓMEZ ESCOTO

SECRETARIA : MARÍA TRINIDAD TRIGUERROS DE CASTRO

ESCUELA DE MATEMÁTICA

DIRECTOR : ING. CARLOS MAURICIO CANJURA LINARES

DEDICATORIA Y AGRADECIMIENTOS

Le dedico este triunfo principalmente a Dios. Le agradezco a Dios todo poderoso por esta bendición, gracias a el pude alcanzar una de las metas mas importantes en mi vida, le agradezco a mis padres María Marta Gámez y José Martínez por su apoyo principalmente a mi madre por el amor intenso que me tiene y por su ayuda incondicional, a mi hermana Johanna Martínez y Ingrid Vilorio le doy la gracias por su apoyo y cariño, a mis amigos y mi compañero de tesis les doy las gracias por su apoyo y sus consejos. De igual forma le agradezco a mi asesor Dr. José Nerys Funes quien con sus enseñanzas contribuyo para el desarrollo de este trabajo de graduación.

Raúl Antonio Martínez Gámez.

Quiero agradecer a Dios por haber alcanzado una de mis más mayores metas y anhelos en mi vida, por haber puesto en mí el deseo de salir siempre adelante sin importar los obstáculos y dificultades que se presentan a la vez de realizar los sacrificios necesarios para superarlos. Agradezco también a mi familia por su apoyo, a mi madre, abuelos, tíos, etc. por su valioso aporte económico y por dar tan importantes consejos y enseñarme que no existen caminos fáciles en la vida; agradezco a mis ex compañeros de clase, principalmente a Azucena Jaimes por motivarme a salir siempre adelante mediante sus consejos, también a Ingrid Martínez, Carolina Linares, Raquel Alvarenga, etc. Además agradezco a Nadia Carranza por enseñar el espíritu de un verdadero estudiante, por sus consejos y ayudas a lo largo de mi carrera y por estar siempre dispuesta a aclarar cualquier interrogante. También a mi compañero de tesis por ayudarme a ver las cosas desde otra perspectiva, a mi asesor de tesis Dr. José Nerys Funes Torres por su paciencia, apoyo, consejo, enseñanzas, etc. que nos transmitió dentro y fuera del aula y contribuir en el desarrollo de nuestro trabajo de graduación.

Carlos Ernesto Canizales Rivera

TRABAJO DE GRADUACIÓN APROBADO POR:

ASESOR : Dr. JOSÉ NERYS FUNES TORRES.

ÍNDICE DE CONTENIDOS

Introducción.....	viii
-------------------	------

Capítulo I: Introducción a la ley de Benford.

1.1 Introducción.....	2
1.2 Descubrimiento de la Ley de Benford.....	2
1.3 Primeros pasos para explicar la Ley de Benford.....	7
1.4 Los dígitos equiprobables no son invariantes de escala.....	9
1.5 Invarianza de escala.....	11
1.6 Ejemplo de variables que cumplen la Ley de Benford.....	11
1.6.1 Datos que satisfacen la ley de Benford.....	13
1.6.2 Datos que no satisfacen la Ley de Benford.....	13
1.7 Aplicaciones de la Ley de Benford.....	14
1.8 Ley de distribución de los dígitos.....	17
1.8.1 Ley de distribución del primer dígito significativo.....	17
1.8.2 Distribución de los k primeros dígitos significativos.....	24
1.8.3 Ley del distribución del segundo dígito significativo.....	29
1.8.4 Ley del k-ésimo dígito significativo.....	33
1.8.5 Probabilidad condicionada de los dígitos.....	34
1.9 Generalización de la Ley de Benford, distribución mantisa.....	37

Capítulo II: Explicación de la Ley de Benford

2.1 Introducción.....	44
2.2 La función mantisa y la sigma álgebra.....	44
2.3 Lemas sobre la sigma álgebra.....	52

2.4 Ley general del dígito significativo.....	61
2.5 Escala invariante.....	63
2.6 Base invariante.....	72
2.7 Teorema principal.....	79

Capítulo III: Ley de Benford y sus aplicaciones

3.1 Introducción.....	89
3.2 Prueba de bondad de ajuste.....	90
3.3 La Ley de Benford en el recuento de votos.....	93
3.4 Aplicación de la Ley de Benford en las elecciones presidenciales de El Salvador en el 2004.....	97
3.4.1 Descripción de los datos.....	97
3.4.2 Análisis del primer dígito significativo del recuento de votos.....	100
3.4.3 Análisis del segundo dígito significativo del recuento de votos....	108
3.4.4 Conclusiones de la aplicación.....	115
Apéndice: Teoría de la Medida y Teoría Ergódica.....	117
Anexos.....	120
Bibliografía.....	128

INTRODUCCIÓN

Durante muchos años la Ley de Benford no ha sido más que una simple curiosidad estadística sin fundamento matemático alguno ni aplicaciones reales. Hoy en día, la Ley está firmemente basada en la Teoría de la Probabilidad, goza de un gran interés del público y presenta importantes aplicaciones a la vista de la estadística. La Ley se ha propuesto como un posible test de evaluación de los resultados obtenidos, ya sea por medios analíticos o de simulación, mediante modelos matemáticos en los que intervienen datos que verifiquen la distribución logarítmica, como por ejemplo: En los resultados de elecciones presidenciales y datos fiscales como la declaración de impuesto sobre la renta. En este sentido se ha utilizado para detectar posibles situaciones de fallos e irregularidades.

La Ley de Benford que también es conocida como la Ley de los números anómalos asegura que, en toda colección de números de la vida real, los dígitos iniciales de los números no son equiprobables, es decir aquellos números que empiezan por el dígito 1 ocurren con mayor frecuencia, seguidos de los que empiezan por el 2, etc. hasta el 9 que es el menos frecuente, esto indica que a medida que crece el valor del primer dígito, más improbable es que éste forme parte de un número. Este hecho, que se admite como Ley, se puede aplicar a datos relacionados con el mundo natural o con elementos sociales.

Sin duda la aplicación más llamativa e importante de la Ley, consiste en su utilización como posible *medio para ayudar a la detección de datos "erróneos" o "fraudulentos"*. Si se detecta que un gran conjunto de datos, tienen un comportamiento contrario a lo esperado, no verifica la Ley, puede ser debido a la presencia de datos inexactos o quizás inventados o "retocados" lo que posibilita aplicaciones tales como la detección de fraude fiscal y de manipulaciones contables, por mencionar algunas.

Evidentemente la no satisfactibilidad de la Ley de Benford en un conjunto determinado de datos no es prueba suficiente de la existencia de irregularidades, pero, sin embargo, constituye un buen indicio para justificar una inspección más detallada. Una aplicación real es la auditoria contable en la que se han desarrollado sistemas basados en el análisis de los dígitos como ayuda en la búsqueda de patrones repetitivos y extraños en los datos.

En muchos países, como por ejemplo Venezuela y México; se ha utilizado como principal herramienta en la detección de fraudes electorales. En las pasadas elecciones Presidenciales de México de 2006, se encontraron ciertos distritos con anomalías en el conteo de voto, es decir, que la distribución de los primeros dígitos, especialmente la del segundo, no se ajustaba satisfactoriamente a la Ley de Benford, lo cual llevó a las autoridades a un recuento de los votos para verificar los datos proporcionados por el Instituto Federal Electoral (IFE).

En este sentido, el objetivo principal del siguiente documento es presentar detalladamente la teoría matemática y estadística concerniente a la Ley de Benford; en el capítulo 1 deducimos las fórmulas de distribución de los dígitos significativos, además de dar las principales características que los datos deben de cumplir para poder someterlos a un escrutinio con la Ley de Benford; en el capítulo 2 demostramos los principales teoremas y lemas que fundamenten la ley y algunos ejemplos de ellos; en el tercer capítulo señalamos como aplicar las fórmulas y a la vez como interpretar los resultados obtenidos mediante la Ley utilizando para ello, los datos de la Elección Presidencial El Salvador 2004.

CAPITULO I
INTRODUCCIÓN A LA
LEY DE BENFORD

1.1 Introducción.

En este capítulo se presenta una breve introducción a la Ley de Benford, su descubrimiento, sus campos de aplicación, los datos en los cuales se esperan que se cumpla dicha Ley, y en cuales no, además se presenta la derivación de las principales fórmulas de distribución para los primeros dígitos significativos y la generalización de la Ley a datos continuos.

En la primera sección se presenta un breve resumen del descubrimiento de la Ley de Benford; en la sección segunda, se muestran algunos aportes de varios autores en su intento de dar una explicación matemática de la Ley; en la tercera y cuarta sección se presentan una breve noción sobre el significado de Escala Invariante mediante un ejemplo ilustrativo; en la quinta sección se dan algunos ejemplos de variables que satisfacen o no la Ley de Benford; en la sexta sección se presentan algunas campos de aplicación recientemente descubierto de para la Ley; en la séptima sección se presentan la derivación estadística de las principales fórmulas de distribución de los primeros dígitos significativos, la distribución conjuntas y condicionadas de los dígitos; en la octava y última sección se introduce la Distribución Mantisa y la generación de la Ley de Benford a datos mediante esta función.

1.2 Descubrimiento de la Ley de Benford.

El descubrimiento de la Ley de Benford se da en el año de 1881 por el matemático y astrónomo **Simon Newcomb**. Un día, él consultaba una tabla de logaritmos para realizar unos cálculos y se dio cuenta; que las primeras páginas se encontraban más viejas y desgastadas que las finales, ¿A qué se debía esto? Al parecer para realizar sus cálculos la gente utilizaba con mayor frecuencia el logaritmo de

números que suelen empezar por dígitos pequeños que los que empiezan por dígitos altos.

Este astrónomo dedujo que los dígitos iniciales de los números (al menos los utilizados en su trabajo que provenían principalmente de la observación de los astros) no son equiprobables; sino que el 1 aparece como dígito inicial más frecuente, seguido del 2, etc. hasta el 9 que es el menos probable. Mediante un breve e ingenioso razonamiento, aunque sin presentar realmente un argumento formal ni fórmula matemática, Newcomb enunció verbalmente una relación o ley logarítmica: **“la ley de probabilidad de ocurrencia de números es tal que las mantisas de sus logaritmos son equiprobables”**. Además concluyó que la probabilidad de los dos primeros dígitos significativos (base 10) satisfacen las siguientes relaciones:

$$(1) \quad \text{Pr ob}(\text{primer dígito significativo} = d_1) = \log_{10}(1 + d_1^{-1}), \quad d = 1, \dots, 9$$

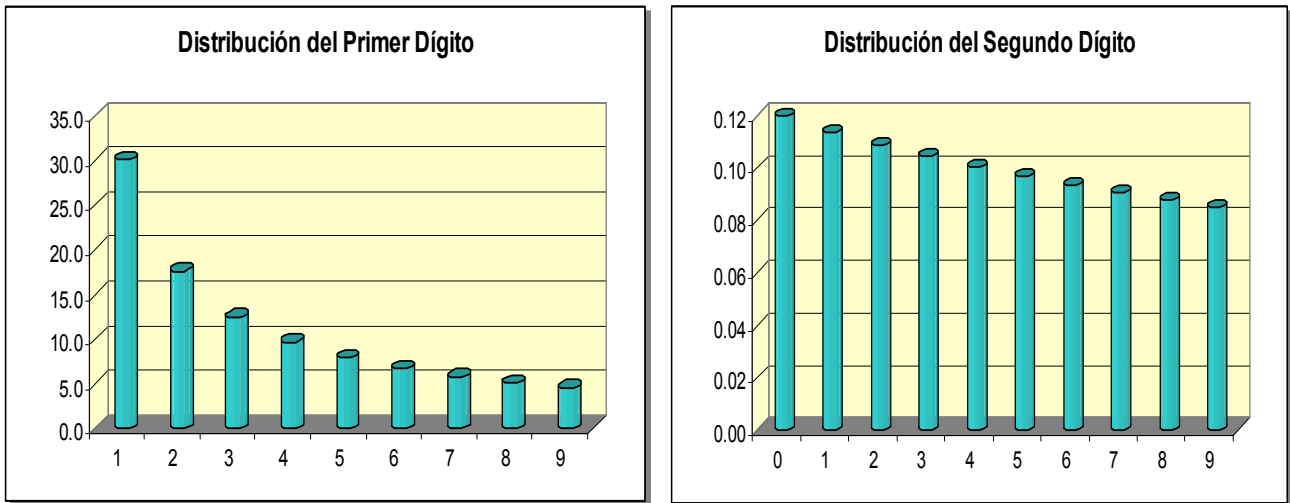
$$(2) \quad \text{Pr ob}(\text{segundo dígito significativo} = d_2) = \sum_{d_1=1}^9 \log_{10}(1 + (10d_1 + d_2)^{-1}), \quad d_1 = 0, 1, \dots, 9$$

Además él dio la siguiente tabla de probabilidad para los primeros dos dígitos:

Tabla 1.1. Frecuencia de aparición de los primeros dos dígitos significativos dado por Newcomb.

d	0	1	2	3	4	5	6	7	8	9
P(d₁=d)		0.3010	0.1761	0.1249	0.0969	0.0792	0.0669	0.0580	0.0512	0.0458
P(d₂=d)	0.1197	0.1139	0.1088	0.1043	0.1003	0.0967	0.0934	0.0904	0.0876	0.0850

Figura 1.1. Frecuencia de aparición de los primeros dos dígitos significativos dada por Newcomb.



Newcomb puntualizó que la distribución de un dígito converge a la distribución uniforme así como su posición crece en la mantisa, es decir; en el caso de la distribución del tercer dígito, la probabilidad será casi la misma para cada dígito, y para las distribuciones del cuarto y de los siguientes dígitos, las diferencias serán menos apreciables.

Aunque Newcomb no aportó ninguna explicación matemática para su descubrimiento. Lo anotó como una simple curiosidad, y frente a una falta de interés general, fue rápidamente olvidada hasta 1938, cuando el Físico Frank Benford de la compañía General Electric, se dio cuenta del mismo patrón. Entusiasmado por el descubrimiento, estudió 13,779 números provenientes de 17 muestras de todo tipo: áreas fluviales, constantes y magnitudes físicas y químicas, funciones matemáticas e incluso números de direcciones de personas y números tomados de portadas de revistas. A partir de esa basta colección de números, calculó las frecuencias de aparición del primer dígito en cada una de las muestras, y también el promedio de todas juntas. El resultado se resume en la siguiente tabla:

Tabla 1.2. Frecuencia de aparición del primer dígito significativo en los experimentos de Frank Benford.

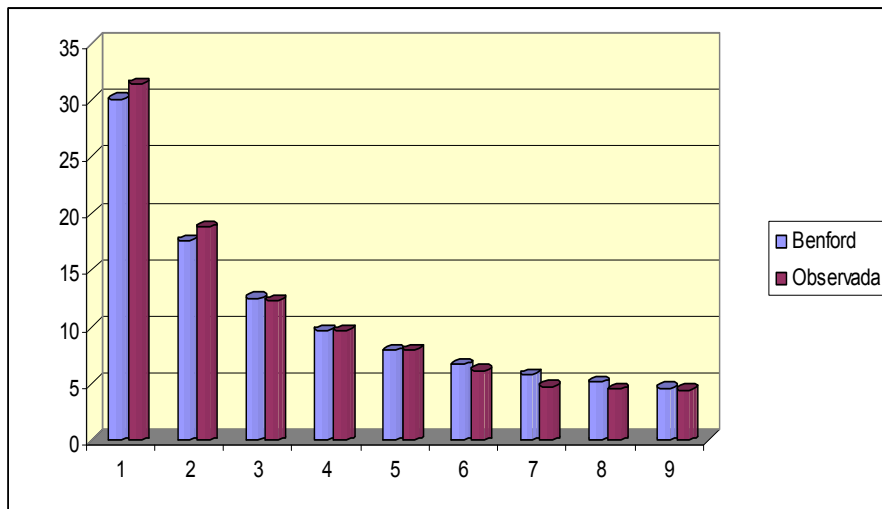
Título/Dígitos	1	2	3	4	5	6	7	8	9	Muestras
Área de Ríos	31	16.4	10.7	11.3	7.2	8.6	5.5	4.8	5.1	335
Población	33.9	20.4	14.2	8.1	7.2	6.2	4.1	3.7	2.2	3259
Constantes	41.3	14.4	4.8	8.6	10.6	5.8	1	2.9	10.6	104
Artículos de Periódicos	30	18	12	10	8	6	6	5	5	100
Golpes Espectrales.	24	18.4	16.2	14.6	10.6	4.1	3.2	4.8	4.1	1389
Presión	29.6	18.3	12.8	9.8	8.3	6.4	5.7	4.4	4.7	703
Pérdida H.P.	30	18.4	11.9	10.8	8.1	7	5.1	5.1	3.6	690
Mol. Wgt.	26.7	25.2	15.4	10.8	6.7	5.1	4.1	2.8	3.2	1800
Drenaje	27.1	23.9	13.8	12.6	8.2	5	5	2.5	1.9	159
Peso atómico	47.2	18.7	5.5	4.4	6.6	4.4	3.3	4.4	5.5	91
Digest	33.4	18.5	12.4	7.5	7.1	6.5	5.5	4.9	4.2	308
Datos de costo	32.4	18.8	10.1	10.1	9.8	5.5	4.7	5.5	3.1	741
Voltios de rayos X	27.9	17.5	14.4	9	8.1	7.4	5.1	5.8	4.8	707
Liga Americana	32.7	17.6	12.6	9.8	7.4	6.4	4.9	5.6	3	1458
Cuerpo negro	31	17.3	14.1	8.7	6.6	7	5.2	4.7	5.4	1165
Direcciones	28.9	19.2	12.6	8.8	8.5	6.4	5.6	5	5	342
Tasa de muertes	27	18.6	15.7	9.4	6.7	6.5	7.2	4.8	4.1	418
Promedio	31.4	18.8	12.3	9.7	8.0	6.1	4.8	4.5	4.4	
Error Probable¹	1.31	1.19	0.19	0.03	0.06	0.56	1.02	0.60	0.13	

1.El Error Probable se ha calculado mediante la siguiente expresión:

$$\left| \text{promedio} - 100 * \log_{10} \left(1 + \frac{1}{d} \right) \right|$$

Benford observó que incluso usando tal mezcla de datos, los números encajaban bastante bien en la ley que Newcomb había descubierto medio siglo antes. Alrededor del 30% empezaban por 1, el 18% por 2, etc. (obsérvese la figura 2). Su análisis era una prueba de la existencia de la ley, pero Benford tampoco fue capaz de explicar bien por qué era así.

Figura 1.2. Frecuencia de aparición de los primeros dos dígitos significativos dada por Newcomb



Benford observó además que la Ley logarítmica era bastante ajustable en la mayoría de muestras seleccionadas por él en su experimento. A partir de los resultados empíricos Benford postuló una ley llamada "*Ley de Benford*" o "*Ley de los Números Anómalos*" para la probabilidad de que el primer dígito sea d . Según dicha ley, la probabilidad que en una serie de datos su primer dígito significativo sea igual al número 1 es alrededor del 30%, 17.6% para el 2, 12.5% para el 3 y continua decreciendo hasta llegar a 4.58% la cual corresponde al 9. El análisis de Benford era una prueba más de la existencia de la Ley.

1.3 Primeros Pasos para Explicar la Ley de Benford.

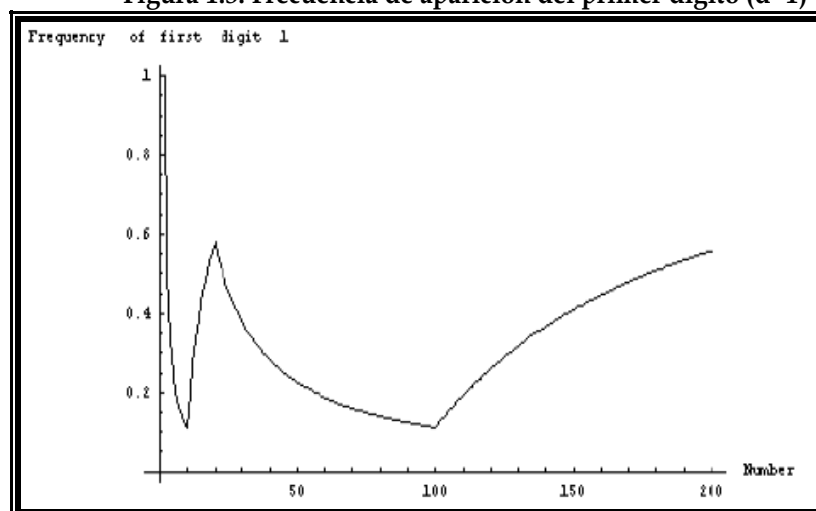
Benford mismo intentó explicar el fenómeno investigando el conjunto de los números enteros naturales, en una tentativa para probar que éstos vienen naturalmente de nuestro sistema numérico. En un inicio, trató de probar que el conjunto de los enteros que tiene al número uno como primer dígito, es decir, $\{D=1\} = \{1, 10, 11, 12, \dots, 19, 100, 101, 102, \dots, 199, 1000, 1001, \dots, 1999, \dots\}$ tienen una “probabilidad” de $\log_{10}(2)$ entre los enteros. El problema que encontró (y que muchos encontraron antes que él) es que este conjunto no tiene una frecuencia asintótica natural, es decir que el límite:

$$(3) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \text{card} \{ \{D=1\} \cap \{1, 2, 3, 4, \dots, n\} \} \text{ No Existe}$$

Dado que el límite de (3) no existe, pero a medida que crece el valor de n , está oscilando en una escala logarítmica entre dos extremos que no son siempre constantes (ver figura 3), con un valor mínimo de $\frac{1}{9} = 0.1111\dots$ y un máximo de

$$\frac{5}{9} = 0.5555\dots$$

Figura 1.3. Frecuencia de aparición del primer dígito ($d=1$)



Existen varias vías para definir un límite para (3), y varias de ellas tienden al valor deseado $\log_{10}(2)$. El problema es primeramente que los métodos de integración no son únicos y por supuesto que la mayoría de ellos no tienden a la Ley de Benford y la aproximación misma es un poco dudosa ya que ésta quiere probar que la Ley es válida para el sistema completo (es decir que es completamente universal), y por supuesto muchos conjuntos de datos incluso los naturales no verifican la Ley de Benford.

Roger Pinkham; un matemático de New Jersey en 1961, contribuyó para resolver este problema. El razonamiento de Pinkham fue el siguiente. Supongamos que realmente existe una Ley de frecuencias de dígitos. En tal caso, dicha Ley debería ser universal. Tanto si calculamos los precios en Dólares, Euros, Colones o Yenes, o si medimos la longitud en centímetros, pulgadas, metros o kilómetros, la frecuencia de aparición de los dígitos debería permanecer inalterada. Es decir, Pinkham afirmaba que la distribución de frecuencias de los dígitos debía ser *invariante frente a cambios de escala*.

Con este razonamiento, Pinkham fue el primero en observar que la Ley de Benford era invariante frente a la escala. Luego demostró que si una ley de frecuencias de dígitos era invariante frente a cambios de escala, entonces se trataba de la Ley de Benford. La prueba aportada iba confirmando que la Ley de Benford realmente existe.

Había todavía algunos problemas enormes en este razonamiento, sin embargo, la hipótesis básica fue “**existe una ley universal**”, y por supuesto que es bastante difícil o imposible de admitir. Sin embargo fue Raimi en 1969, quien demuestra finalmente la independencia de la unidad de medida de la Ley de Benford, Raimi no sólo aporta el fundamento matemático de esta independencia de escala, sino

que además en un intento divulgativo, añade explicaciones intuitivas de la invariabilidad.

El principio de los 90 parece ser un periodo de renovación para la Ley de Benford. Esto es quizá debido al descubrimiento de un nuevo campo importante de aplicación, es decir, la *detección de fraude*. Uno de los papeles claves escritos sobre el tema fue la tesis doctoral de Mark Nigrini (en 1992 en el Departamento de Contabilidad, Universidad de Cincinnati).

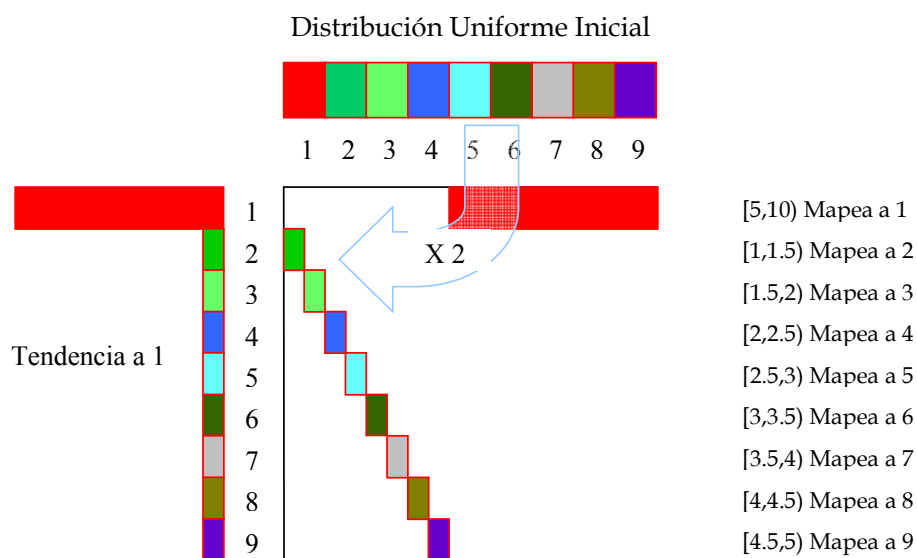
En los 90, una de las principales contribuciones al análisis de la Ley de Benford es ciertamente debido al teorema de Theodore Hill, quien fijó un marco de trabajo correcto de probabilidad para la Ley de Benford, extendió la idea de invarianza de escala a invarianza de base (¿Por qué dependería una ley universal de la base en la que los números son escritos?), e introdujo un nuevo camino para considerar la Ley de Benford. Sin embargo, esta Ley no ha podido ser probada en forma total. Una aproximación más natural y de hecho tan natural que sorprendentemente nadie la ha estudiado antes que Hill en 1995, es pensar en que los datos son una mezcla de distribuciones diferentes. Actualmente ésta aproximación parece ser relevante al experimento de Benford, donde los datos provienen de más de 20 distribuciones variadas. Hill enlazó esta idea con invarianza de escala (y de base) para hacer una explicación consistente (pero aún no completa) de la Ley de Benford.

1.4 Los dígitos equiprobables no son invariantes de escala.

La mayoría de la gente tiene la intuición de que cada uno de los dígitos del 1 al 9 tiene la misma probabilidad de aparecer como primer dígito significativo en cualquier cifra. Supongamos que este es el caso y veamos qué ocurre con un conjunto de cantidades que se convertirán de Euros a Dólares con el valor (ficticio)

de 2 Dólares por Euro. Es bastante fácil averiguar qué ocurrirá observando cada dígito uno por uno. Si el primer dígito significativo es 1, entonces al multiplicar por 2 se obtendrá un nuevo primer dígito 2 ó 3 con la misma probabilidad. Pero si el primer dígito significativo es 5, 6, 7, 8 ó 9 el nuevo primer dígito será 1. Se desprende que en el nuevo conjunto de datos, el primer dígito 1 es diez veces más probable que cualquier otro primer dígito (véase la figura 4). En el diagrama de la Figura 4 la notación $[a,b)$ indica la gama de números mayores o iguales que a pero estrictamente menores que b .

Figura 1.4. Ilustración gráfica de la invarianza de escala.



La intuición nos falló: la distribución uniforme original se orienta con fuerza hacia el dígito 1. Así que si la invarianza frente a la escala es necesaria, la distribución uniforme no es la respuesta acertada.

1.5 Invarianza de escala.

Tal y como argumentó Pinkham, el hecho de que encontremos todo tipo de datos en el mundo real que parecen encajar en la Ley de Benford, sugiere que esa ley debe ser invariante frente a la escala. ¿Por qué? Porque podemos medir nuestros datos con una gama de escalas distintas (Pies/Metros, Euros/Dólares, Galones/Mililitros, etc.). Si la ley de frecuencia de dígitos es real, debería serlo para todas (no hay razón para que sólo una escala de medidas, la que elijamos al azar, sea la correcta).

Así que si hay una ley de distribución de los primeros dígitos significativos, debería mantenerse independientemente de las unidades usadas. ¿Pero que significa realmente la invarianza frente a la escala de distribución del primer dígito significativo? Quiere decir que si multiplicamos todos los números por una constante arbitraria (como hacemos cuando cambiamos de Libras a Yenes, o Pies a Metros), la distribución de frecuencia del primer dígito debería permanecer inalterada. Como estamos interesados en la distribución de los primeros dígitos significativos, tiene sentido expresar los números en notación científica $x \times 10^n$, donde $1 \leq x < 10$ y $n \in \mathbb{Z}$ (Esto es posible con todos los números excepto el cero). El primer dígito significativo d es, sencillamente, el primer dígito de x . Podemos derivar fácilmente una distribución invariante frente a la escala para d cuando hayamos encontrado una distribución invariante frente a la escala para x .

1.6 Ejemplos de variables que satisfacen la Ley de Benford

En este punto podríamos sentirnos tentados a revisar la forma en que elegimos los números de la lotería. Adiós a las fechas de los cumpleaños y bienvenida Benford. ¿Notaremos la diferencia?

Lamentablemente, la respuesta es negativa. El resultado de la lotería es totalmente aleatorio, de forma que cada número tiene la misma probabilidad de aparecer. A largo plazo, la frecuencia del primer dígito debería permanecer, por tanto, en proporción exacta con respecto a la cantidad de números de la lotería que empezarán por ese dígito.

Por otro lado, consideremos los tiempos de los 400 metros olímpicos en segundos. ¡Ninguno comienza por 1! De la misma manera, pensemos en los mandatos, en años, de los políticos mundiales. Igualmente, muy pocos empiezan por 1. Al contrario que la lotería, estos datos *no* son aleatorios, sino que están muy condicionados. El abanico de posibilidades es muy limitado como para permitir que se cumpla una ley de frecuencia de dígito.

En otras palabras, la Ley de Benford necesita datos que no sean totalmente aleatorios ni muy condicionados, sino que estén más o menos en medio. Los datos pueden ser de una gran variedad y suelen ser el resultado típico de diversos procesos, con muchas influencias. Por ejemplo, las cifras de población de pueblos y ciudades pueden variar desde decenas o cientos a miles o millones, y les afecta un gran abanico de factores.

En general las principales características que deben de cumplir los datos numéricos en los cuales se verifica la Ley de Benford son las siguientes:

- Escala Invariante
- Base Invariante

(Posteriormente en el capítulo 2 se presentarán las demostraciones de los teoremas relacionados a las características mencionadas).

1.6.1 Datos que satisfacen la ley

Los siguientes conjuntos de datos numéricos que se presentan a continuación, según experimentos realizados por Benford satisfacen la Ley ya que cumplen las características anteriores (escala y base invariante).

- Áreas de ríos.
- Estadísticas de Baseball,
- Constantes y Magnitudes Físicas y Químicas.
- Poblaciones en Diferentes Municipios.
- Pagos de Impuestos sobre la Renta.
- La ley de Exponenciación a^x o el producto o división de un elevado número de datos aleatorios uniformemente distribuidos o sus recíprocos que en el límite presentan la distribución logarítmica.
- En el campo de la física se ha hallado la distribución logarítmica en la vida media de las desintegraciones radiactivas de partículas alfa y en las constantes físicas más usuales.
- En economía recientemente se ha observado que determinados índices bursátiles también siguen la Ley de Benford al igual que una gran cantidad de magnitudes económicas y sociales.

1.6.2 Datos que no satisfacen la ley

A continuación se presentan algunas bases de datos que según experimentos realizados por Frank Benford, no satisfacen la Ley.

- Datos procedentes de **distribuciones uniformes** (por ejemplo; números de lotería): por no cumplir el requisito de invariante de escala (véase la sección 1.3).

- Datos procedentes de distribuciones normales (por ejemplo: edades de personas).
- Datos que tienen limitado el valor del dígito inicial (por ejemplo: los precios de muchos productos suelen restringirse a unos pocos valores, muchas veces uno solo, por razones comerciales y de mercado).
- Números telefónicos (de un municipio) o documentos de identidad, pues se asignan arbitrariamente no para medir una determinada característica de un objeto sino con el propósito de identificarlo y distinguirlo de otros objetos semejantes.
- Números procedentes de evaluar funciones como x^2 , $x^{1/2}$, ó $1/x$.

1.7 Aplicaciones con la Ley de Benford.

A principios del siglo pasado la ley de de Benford fue vista como una clase de fisura matemática sin algún entorno práctico, en los años 70 con la llegada de las computadoras se pensó en algunas aplicaciones; una de ellas fue el *Diseño de computadoras* cuando se diseña una computadora o se escribe una rutina, se deben de considerar las distribuciones de los operadores. Codificar los números reales requiere por ejemplo decidir con qué número de bits se describirá la mantisa (esto es enlazado con intervalos de confianza). Un cierto patrón en la mantisa conducirá a diferentes elecciones, para optimizar los procesos de velocidad y almacenamiento.

Otra aplicación que se realizo con la Ley de Benford fue en *la Modelación matemática*, aparentemente imaginada por Varian en 1972. La cual está basada en la siguiente idea: “Si un cierto conjunto de valores siguen la Ley de Benford, entonces los modelos para los correspondientes valores predictivos también

seguirán la Ley”. Hill calificó este tipo de aplicaciones como “**Benford-in-Benford-out**”.

Pero sin duda alguna en la actualidad el campo más importante de aplicación de la ley, es la *detección de fraude*. En 1992 se descubrió que los datos financieros encajan en la Ley de Benford. Este hecho resulta tremendamente importante para detectar fraudes en las finanzas, como por ejemplo, en la declaración de impuestos.

El Dr. Mark Nigrini, un profesor de contabilidad de Dallas, lo ha utilizado con bastante éxito. Si alguien trata de falsificar, por ejemplo, su declaración de la renta, irremediablemente tendrá que inventar algún dato. Al intentarlo, la tendencia de la gente es utilizar demasiados números que comienzan por dígitos a mitad de escala como por ejemplo; 5, 6, 7, y pocos que comiencen por 1. Esta violación de la Ley de Benford hace resaltar la alarma.

El Dr. Nigrini ha creado un programa informático para detectar en qué medida algunos datos suministrados encajan con la Ley de Benford. El resultado ha sido increíblemente exitoso. Recientemente la oficina del Fiscal del Distrito de Brooklyn ha tratado siete importantes casos de fraude y el programa del Dr. Nigrini fue capaz de detectar los siete casos. También se utilizó dicho programa para analizar la declaración de la renta de Bill Clinton. Aunque, reveló que probablemente había varios redondeos en lugar de cifras exactas, no hubo indicios de fraude.

El método de Benford, tiene a su vez algunas limitaciones en la detección de fraude. De hecho, a menudo los datos pueden alejarse de la Ley de Benford por motivos perfectamente inocentes. A veces las cifras no pueden darse de forma precisa, y es entonces cuando aparece el redondeo, lo cual puede modificar al primer dígito de un número. Asimismo, sobre todo cuando se trata de precios, las

cifras 95 y 99 aparecen demasiado, debido a estrategias de marketing. En estos casos, la Ley de Benford podría indicar un fraude cuando realmente no sea el caso. El método no es infalible.

Sin embargo, el uso de esta ley no se ciñe a la caza de fraudes. Con la Ley de Benford es posible detectar un cambio significativo en las cifras reportadas por parte de empresas y/o personas en años consecutivos. Demasiado cambio indicaría que algo anda mal. Se podría ahorrar tiempo, dinero y medios si los sistemas informáticos se manejaran de forma más eficaz.

En países de Latinoamérica, específicamente en Venezuela y México; se ha utilizado como principal herramienta en la detección de fraudes electorales. En las pasadas elecciones Presidenciales de México de 2006, se encontraron ciertos distritos con anomalías en el conteo de voto, es decir, que la distribución de los primeros dígitos, especialmente la del segundo, no se ajustaba satisfactoriamente a la Ley de Benford, lo cual llevó a las autoridades a un recuento de los votos para verificar los datos proporcionados por el Instituto Federal Electoral (IFE).

Mientras que un equipo de Friburgo está trabajando en la idea de distribuir espacio del disco duro según la Ley de Benford. Científicos belgas investigan si la Ley de Benford puede usarse para detectar irregularidades en casos clínicos. Mientras tanto, la buena correlación existente entre las estadísticas poblacionales y la Ley de Benford significa que puede usarse para verificar modelos demográficos. ¿Quién sabe en qué más casos sería útil? El Dr. Nigrini comenta: *Preveo muchas aplicaciones, pero para mí es fascinante en sí misma. Considero a Benford un gran héroe. Su ley no es mágica, pero a veces lo parece.*

1.8 Ley de distribución de los dígitos.

En esta sección se presentará la deducción formal de los primeros dígitos significativos, es decir; la ley de distribución del primer dígito significativo, la ley de distribución del segundo dígito significativo, la ley de distribución del k-ésimo dígito significativo, la ley de distribución conjunta de los dos primeros dígitos significativos, la ley de distribución conjunta de los dos k dígitos significativos, etc.

1.8.1 Ley de distribución del primer dígito significativo.

Proposición 1.8.1.1 La Distribución del primer dígito significativo d , se define de la siguiente manera:

$$P(d) = \log_{10} \left(1 + \frac{1}{d} \right) \quad d \in \{1, 2, \dots, 9\}$$

Demostración.

Todo fenómeno que sigue la Ley de Benford cumple que es *invariante frente a cambios de escala*, es decir, que la distribución de probabilidad de una variable aleatoria x es independiente de la escala de medición.

Como los valores numéricos de los datos dependen de las unidades de medición, según como Roger Pinkham estaríamos interesados en encontrar una *distribución Universal* $P(x)$ sobre tales números, tal que esta probabilidad permanezca invariante bajo cambios de escala, es decir;

$$(4) \quad P(kx) = f(k)P(x); \text{ con } k \text{ constante}$$

Si kx es el resultado de un cambio de escala de la variable x entonces la probabilidad de dicho cambio de escala es igual al producto de la $P(x)$ por una constante $f(k)$.

Para que $P(x)$ y $P(kx)$ sean funciones de distribución de probabilidad, estas tienen que cumplir que

$$\begin{aligned}
 \int_{-\infty}^{\infty} P(x)dx &= 1 \quad \text{y} \quad \int_{-\infty}^{\infty} P(kx)d(kx) = 1 \\
 \int_{-\infty}^{\infty} P(kx)d(kx) &= 1 \\
 &= \int_{-\infty}^{\infty} f(k)P(x)d(kx) \\
 &= \int_{-\infty}^{\infty} f(k)P(x)kdx \quad ; \quad \text{por ecuación (4)} \\
 &= kf(k)\int_{-\infty}^{\infty} P(x)dx \\
 &\Rightarrow kf(k)\int_{-\infty}^{\infty} P(x)dx = 1 \\
 &\Rightarrow kf(k) = 1 \quad ; \quad \text{ya que} \quad \int_{-\infty}^{\infty} P(x)dx = 1 \\
 &\Rightarrow f(k) = \frac{1}{k}
 \end{aligned}$$

Ahora si derivamos parcialmente respecto a k en la ecuación (4) tendremos que

$$\begin{aligned}
 P(kx) &= f(k)P(x) \\
 P(kx) &= \frac{1}{k}P(x) \\
 \frac{\partial}{\partial k} P(kx) &= \frac{\partial}{\partial k} \left(\frac{1}{k} P(x) \right) \\
 xP'(kx) &= \frac{-1}{k^2} P(x)
 \end{aligned}$$

$$u=kx$$

$$\frac{u}{k}P'(u)=-\frac{1}{k^2}P\left(\frac{1}{k}u\right)$$

$$\Rightarrow \frac{u}{k}P'(u)=-\frac{1}{k^2}\left(\left(\frac{1}{k}\right)P(u)\right)$$

$$\frac{u}{k}P'(u)=-\frac{k}{k^2}(P(u))$$

$$xP'(x)=-P(x)$$

Si hacemos $y=P(x)$ y $\frac{dy}{dx}=P'(x)$ y las sustituimos en la ecuación anterior, entonces

$$x\frac{\partial x}{\partial y}=-y$$

$$\Rightarrow \frac{\partial y}{y}=-\frac{\partial x}{x}$$

$$\Rightarrow \int \frac{\partial y}{y} = \int -\frac{\partial x}{x}$$

$$\Rightarrow \int \frac{\partial y}{y} = -\int \frac{\partial x}{x}$$

$$\Rightarrow \ln(y) = -\ln(x)$$

$$\Rightarrow \ln(y) = \ln\left(\frac{1}{x}\right)$$

$$\Rightarrow y = \frac{1}{x}$$

Por tanto $P(x)=\frac{1}{x}$ con $x \in [1,10)$

Puesto que ya conocemos la distribución de probabilidad que es invariante en escala tendremos que; la probabilidad de que el primer dígito significativo sea igual a d , está dada por:

$$P_d = P(X = d) = \frac{\int_1^{d+1} P(x) dx}{\int_1^{10} P(x) dx}$$

Sustituyendo $P(x)$ por $\frac{1}{x}$ en la fórmula anterior se tiene.

$$\begin{aligned} P(X = d) &= \frac{\int_1^{d+1} \frac{1}{x} dx}{\int_1^{10} \frac{1}{x} dx} \\ &= \frac{\ln(x) \Big|_1^{d+1}}{\ln(x) \Big|_1^{10}} \\ &= \frac{\ln(d+1) - \ln(1)}{\ln(10) - \ln(1)} \\ &= \frac{\ln\left(1 + \frac{1}{d}\right)}{\ln(10)} \\ &= \log_{10}\left(1 + \frac{1}{d}\right) \end{aligned}$$

Por tanto $P(X = d) = \log_{10}\left(1 + \frac{1}{d}\right)$ Con $d= 1, 2, 3\dots, 9$, como queríamos

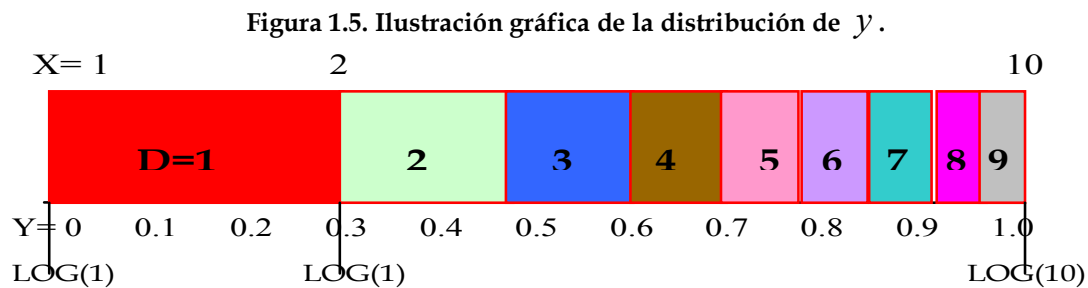
demostrar

Otra forma de probar:

Si una distribución para x es invariante frente a la escala, la distribución de $y = \log_{10}(x)$ debería permanecer inalterada cuando sumáramos un valor constante a y . ¿Por qué? Porque multiplicaríamos a x por una constante a .

$$y \text{ entonces } \log_{10}(ax) = \log_{10}(a) + \log_{10}(x) = \log_{10}(a) + y$$

Ahora bien, la única distribución de probabilidad de $y \in [0,1)$ que permanecerá invariada después de sumar una constante arbitraria a y , es la distribución uniforme. Para convencernos de esto, pensemos en la forma de la función de densidad de probabilidad para la distribución uniforme.



En la Figura 1.5, y se distribuye uniformemente entre $\log_{10}(1)=0$ y $\log_{10}(10)=1$. Si queremos encontrar la probabilidad de que d sea 1 debemos calcular.

$$\begin{aligned} P(d=1) &= P(1 \leq X < 2) \\ &= P(0 \leq y < \log_{10}(2)) \end{aligned}$$

Con lo que

$$P(d=1) = \int_0^{\log_{10}(2)} 1 dy = \log_{10}(2)$$

En general,

$$\begin{aligned}P(d=n) &= P(n \leq X < n+1) \\ &= P(\log_{10}(n) \leq y < \log_{10}(n+1))\end{aligned}$$

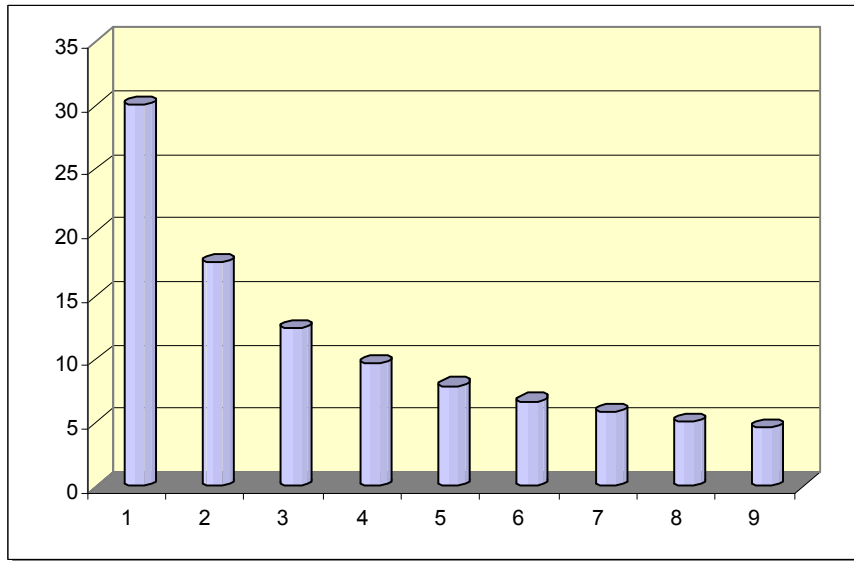
Con lo que obtenemos

$$\begin{aligned}P(d=n) &= \int_{\log_{10}(n)}^{\log_{10}(n+1)} 1 dy \\ &= \log_{10}(n+1) - \log_{10}(n) \\ &= \log_{10}\left(\frac{n+1}{n}\right)\end{aligned}$$

La fórmula $\log_{10}\left(\frac{n+1}{n}\right)$ fue precisamente la fórmula aportada por Newcomb y más tarde por Benford para la proporción de números cuyo primer dígito es n. Así que podemos demostrar que la invarianza frente a la escala para la distribución de las frecuencias del primer dígito de x implica que dicha distribución debe ser la Ley de Benford.

En el figura 1.6 se presenta la forma en que se comporta la distribución del primer dígito significativo, La Ley de Benford establece que en una lista de hechos o fenómenos de la vida real, la frecuencia que el primer dígito significativo sea igual a uno es cerca del 30%, que sea dos de aproximadamente 17%, así, los dígitos mayores tienen una menor probabilidad de ocurrencia en comparación con la de los dígitos más pequeños.

Figura 1.6. Distribución del primer dígito significativo



Probabilidad acumulada del primer dígito significativo.

Proposición 1.8.1.2. La distribución a cumulada del primer digito significativo se define de la siguiente forma:

(5) $P(x \leq d) = \log_{10}(d + 1)$; con $d = 1, 2, 3 \dots 9$

Sabemos que la distribución acumulada de una variable discreta se define como la suma de la función de densidad desde $-\infty$ hasta el valor de x :

$$P(X \leq x) = \sum_{i=-\infty}^x P(i)$$

Utilizando el concepto anterior para el caso de la distribución acumulada del primer dígito significativo, se tiene:

Demostración.

$$\begin{aligned} P(x \leq d) &= \sum_{i=1}^d P(i) = \sum_{i=1}^d \log_{10} \left(\frac{i+1}{i} \right) \\ &= \sum_{i=1}^d [\log_{10}(i+1) - \log_{10}(i)] = \sum_{i=1}^d \log_{10}(i+1) - \sum_{i=1}^d \log_{10}(i) \\ &= \sum_{i=2}^{d+1} \log_{10}(i) - \sum_{i=1}^d \log_{10}(i) = \log_{10}(d+1) + \sum_{i=2}^d \log_{10}(i) - \left[\log_{10}(1) + \sum_{i=2}^d \log_{10}(i) \right] \\ &= \log_{10}(d+1) - \log_{10}(1) = \log_{10}(d+1) \end{aligned}$$

Por tanto $P(x \leq d) = \log_{10}(d+1)$, como queríamos demostrar

1.8.2 Distribución conjunta de los primeros k dígitos significativos.

De una forma análoga con la que obtuvimos la ley de distribución del primer dígito significativo, podemos obtener una ley que proporcione la distribución conjunta de los k primeros dígitos significativos.

Proposición 1.8.2.1 La distribución logarítmica conjunta de los k primeros dígitos significativos d_1, d_2, \dots, d_k (siendo k un entero positivo) se define de la siguiente manera:

$$(6) \quad P(d_1, d_2, \dots, d_k) = \log_{10} \left(1 + \left(\sum_{i=1}^k d_i \cdot 10^{k-i} \right)^{-1} \right); \text{ para todo entero positivo } k$$

Con $d_1 = 1, 2, \dots, 9$
 $d_j = 0, 1, 2, \dots, 9; \quad j \geq 2$

Demostración.

Notemos que la sumatoria interna, es simplemente la representación en base decimal del número z cuyos primeros k dígitos significativos son respectivamente d_1, d_2, \dots, d_{k-1} y d_k , con lo que la probabilidad de ocurrencia de esos k dígitos es simplemente $\log_{10}(1 + 1/z)$.

$$\text{Donde: } z = \sum_{i=1}^k d_i \cdot 10^{k-i}$$

Para la deducción de la ecuación, hagamos el número $z = d_1 d_2 \dots d_k$.

Entonces

$z = d_1 d_2 \dots d_k$ puede escribirse de la siguiente manera:

$$z = d_1 d_2 \dots d_k = d_1 10^{k-1} + d_2 10^{k-2} + \dots + d_{k-1} 10^{k-(k-1)} + d_k$$

o

$$z = \sum_{i=1}^k d_i 10^{k-i}$$

Luego usando el mismo razonamiento empleado para la obtención de la distribución del primer dígito obtenemos lo siguiente:

$$\begin{aligned} P(d_1, \dots, d_k) &= \frac{\int_1^{1+z} P(z) d(z)}{\int_1^{10} P(z) d(z)} \\ \Rightarrow P(d_1, \dots, d_k) &= \frac{\ln(z) \Big|_1^{1+z}}{\ln(z) \Big|_1^{10}} \\ \Rightarrow P(d_1, \dots, d_k) &= \frac{\ln(1+z) - \ln(z)}{\ln(10)} \end{aligned}$$

$$\Rightarrow P(d_1, \dots, d_k) = \frac{\ln[1 + z^{-1}]}{\ln(10)}$$

$\Rightarrow P(d_1, \dots, d_k) = \log_{10}[1 + z^{-1}]$; *justo lo que queríamos demostrar*
y finalmente obtenemos que

$$P(d_1, \dots, d_k) = \log_{10} \left[1 + \left(\sum_{i=1}^k d_i 10^{k-i} \right)^{-1} \right]$$

Obsérvese que los límites de integración inferior y superior son, respectivamente, z y $z + 1$; ya que con esto estaríamos obteniendo la proporción de números en los cuales sus primeros dígitos significativos sean d_1, d_2, \dots, d_k .

Por ejemplo, la función de distribución conjunta para los primeros dos dígitos significativos es:

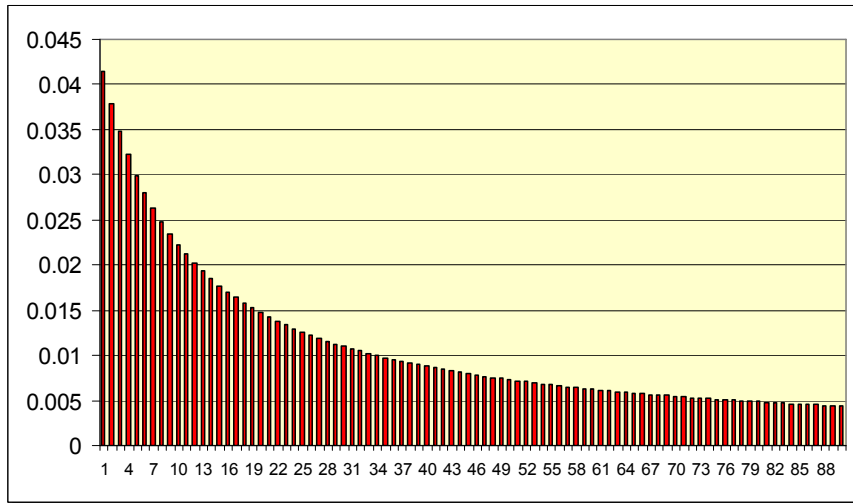
$$P(d_1, d_2) = \log_{10} \left[1 + \left(\sum_{i=1}^2 d_i 10^{2-i} \right)^{-1} \right]$$

En la ecuación anterior puede apreciarse que a medida que crecen los valores para d_1 y d_2 se tendrá una mayor probabilidad conjunta, es decir, la función de distribución es monótona decreciente, en la siguiente tabla y figura se ilustra, tal comportamiento.

Tabla 1.3. Probabilidad conjunta para los primeros dos significativos.

d_1	d_2	$P(d_1, d_2)$	d_1	d_2	$P(d_1, d_2)$	d_1	d_2	$P(d_1, d_2)$
1	0	0.041	4	0	0.011	7	0	0.006
1	1	0.038	4	1	0.01	7	1	0.006
1	2	0.035	4	2	0.01	7	2	0.006
1	3	0.032	4	3	0.01	7	3	0.006
1	4	0.03	4	4	0.01	7	4	0.006
1	5	0.028	4	5	0.01	7	5	0.006
1	6	0.026	4	6	0.009	7	6	0.006
1	7	0.025	4	7	0.009	7	7	0.006
1	8	0.023	4	8	0.009	7	8	0.006
1	9	0.022	4	9	0.009	7	9	0.005
2	0	0.021	5	0	0.009	8	0	0.005
2	1	0.02	5	1	0.008	8	1	0.005
2	2	0.019	5	2	0.008	8	2	0.005
2	3	0.018	5	3	0.008	8	3	0.005
2	4	0.018	5	4	0.008	8	4	0.005
2	5	0.017	5	5	0.008	8	5	0.005
2	6	0.016	5	6	0.008	8	6	0.005
2	7	0.016	5	7	0.008	8	7	0.005
2	8	0.015	5	8	0.007	8	8	0.005
2	9	0.015	5	9	0.007	8	9	0.005
3	0	0.014	6	0	0.007	9	0	0.005
3	1	0.014	6	1	0.007	9	1	0.005
3	2	0.013	6	2	0.007	9	2	0.005
3	3	0.013	6	3	0.007	9	3	0.005
3	4	0.013	6	4	0.007	9	4	0.005
3	5	0.012	6	5	0.007	9	5	0.005
3	6	0.012	6	6	0.007	9	6	0.005
3	7	0.012	6	7	0.006	9	7	0.004
3	8	0.011	6	8	0.006	9	8	0.004
3	9	0.011	6	9	0.006	9	9	0.004

Figura 1.7. Distribución conjunta de los primeros dos dígitos significativos.



Mientras que la función de distribución conjunta para los primeros tres dígitos significativos viene dada por:

$$P(d_1, d_2, d_3) = \log_{10} \left[1 + \left(\sum_{i=1}^3 d_i 10^{3-i} \right)^{-1} \right]$$

Es de mencionar que, si la distribución de los dígitos significativos obedece una ley logarítmica, entonces los dígitos significativos no son independiente entre si.

Ejemplo 1.8.2.2: La probabilidad de que el primer dígito significativo sea 1 y que el segundo dígito significativo sea 2 es:

Por una parte, la probabilidad conjunta del primer y segundo dígito tomando respectivamente los valores de 1 y 2.

$$p(d_1 = 1, d_2 = 2) = \log_{10} \left(\frac{13}{12} \right) = 0.035,$$

Mientras que la probabilidad de que el primer dígito significativo sea 1 por la probabilidad de que el segundo dígito sea 2 es la siguiente.

$$p(d_1=1) * p(d_2=2) = \log_{10}(2) * \left(\log_{10}\left(\frac{13}{12}\right) + \log_{10}\left(\frac{23}{22}\right) + \dots + \log_{10}\left(\frac{93}{92}\right) \right) = 0.033.$$

Se observa claramente que el valor de ambas probabilidades es ligeramente diferente.

1.8.3 Ley del segundo dígito significativo.

Proposición 1.8.3.1 La distribución de probabilidad para el segundo dígito significativo es:

$$(7) \quad p(d_2) = \sum_{d_1=1}^9 \log_{10}\left(1 + (10d_1 + d_2)^{-1}\right); \quad d_2 = 0, 1, 2, \dots, 9$$

Donde d_1 y d_2 representan el primero y segundo dígito significativo, respectivamente.

Demostración:

$$\begin{aligned} P(d_2 = d) &= \frac{\int_{10^d}^{10^{d+1}} P(x) d(x)}{\int_1^{10} P(x) d(x)} + \frac{\int_{10^{d+1}}^{10^{d+2}} P(x) d(x)}{\int_1^{10} P(x) d(x)} + \dots + \frac{\int_{10^{9+d}}^{10^{9+d+1}} P(x) d(x)}{\int_1^{10} P(x) d(x)} \\ &= \frac{\int_{10^d}^{10^{d+1}} \frac{1}{x} d(x)}{\int_1^{10} \frac{1}{x} d(x)} + \frac{\int_{10^{d+1}}^{10^{d+2}} \frac{1}{x} d(x)}{\int_1^{10} \frac{1}{x} d(x)} + \dots + \frac{\int_{10^{9+d}}^{10^{9+d+1}} \frac{1}{x} d(x)}{\int_1^{10} \frac{1}{x} d(x)} \\ &= \frac{\ln(x) \Big|_{10^d}^{10^{d+1}} + \ln(x) \Big|_{10^{d+1}}^{10^{d+2}} + \dots + \ln(x) \Big|_{10^{9+d}}^{10^{9+d+1}}}{\ln(10)} \end{aligned}$$

$$\begin{aligned}
&= \frac{\ln(11+d) - \ln(10+d)}{\ln(10)} + \frac{\ln(21+d) - \ln(20+d)}{\ln(10)} + \dots + \frac{\ln(91+d) - \ln(90+d)}{\ln(10)} \\
&= \log_{10} \left(\frac{11+d}{10+d} \right) + \log_{10} \left(\frac{21+d}{20+d} \right) + \dots + \log_{10} \left(\frac{91+d}{90+d} \right) \\
&= \log_{10} \left[1 + (10+d)^{-1} \right] + \log_{10} \left[1 + (20+d)^{-1} \right] + \dots + \log_{10} \left[1 + (90+d)^{-1} \right] \\
&= \sum_{d_1=1}^9 \log_{10} \left[1 + (10d_1 + d_2)^{-1} \right]
\end{aligned}$$

Que es precisamente lo que queríamos demostrar.

Nótese que el primer cociente de integrales es precisamente la proporción de números, cuyos primer y segundo dígitos son respectivamente 1 y d . El segundo cociente, que el primer y segundo dígito sean 2 y d , y así sucesivamente hasta llegar al noveno cociente el cual representa la proporción de números cuyo primer dígito sea 9 y su segundo dígito sea d .

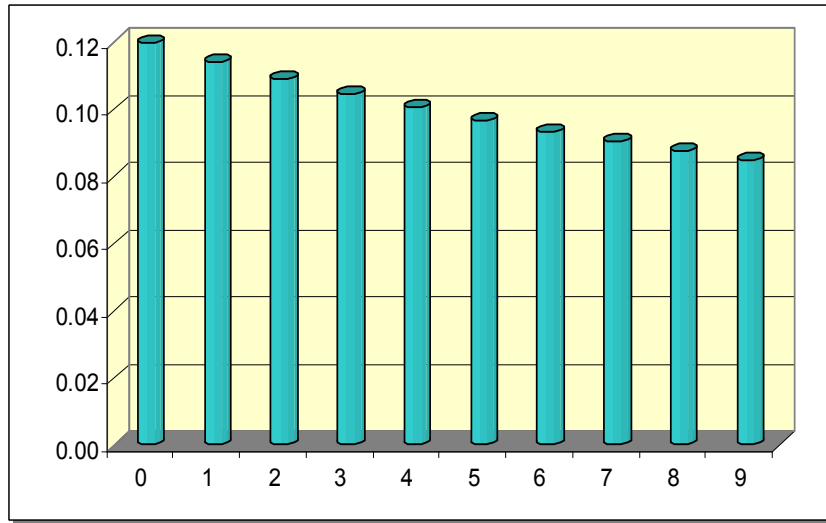
En la siguiente tabla se muestra la probabilidad de ocurrencia para el primer y segundo dígito significativo, utilizando las ecuaciones de las proposiciones 1.7.1.1 y 1.7.3.1, respectivamente.

Tabla 1.4. Probabilidad de los dos primeros dígitos significativos

d	0	1	2	3	4	5	6	7	8	9
P(d₁=d)		0.3010	0.1761	0.1249	0.0969	0.0792	0.0669	0.0580	0.0512	0.0458
P(d₂=d)	0.1197	0.1139	0.1088	0.1043	0.1003	0.0967	0.0934	0.0904	0.0876	0.0850

Observemos que la distribución de las probabilidades para el segundo dígito tiende ser un poco más uniforme que la presentada para el primer dígito. En la siguiente figura se muestra la distribución del segundo dígito.

Figura 1.8. Distribución del segundo dígito significativo.



Verificación de los axiomas de probabilidad de la distribución del segundo dígito significativo.

a) Primer axioma:

$$0 \leq P(d_2) \leq 1$$

Demostración. De la figura anterior se observa que las probabilidades para los diferentes valores que puede tomar d_2 son todas positivas y a la vez son todas menores que 1, el cual también se puede comprobar muy fácilmente de la ecuación que no importa que valor tome d_2 siempre se estaría sumando el logaritmo de valores cercanos pero mayores a uno, los cuales como ya sabemos son siempre mayores que cero y que a su vez éstos son menores a uno.

b) Segundo axioma:

$$\sum_{d_2=0}^9 P(d_2) = 1$$

Demostración. Tenemos que

$$\begin{aligned}
 \sum_{d_2=0}^9 P(d_2) &= \sum_{d_2=0}^9 \sum_{d_1=1}^9 \log_{10} (1 + (10d_1 + d_2)^{-1}) \\
 \Rightarrow \sum_{d_2=0}^9 P(d_2) &= \sum_{d_2=0}^9 \sum_{d_1=1}^9 \log_{10} \left(\frac{10d_1 + d_2 + 1}{10d_1 + d_2} \right) \\
 \Rightarrow \sum_{d_2=0}^9 P(d_2) &= \sum_{d_2=0}^9 \log_{10} \left[\prod_{d_1=1}^9 \left(\frac{10d_1 + d_2 + 1}{10d_1 + d_2} \right) \right] \\
 \Rightarrow \sum_{d_2=0}^9 P(d_2) &= \log_{10} \left[\prod_{d_2=0}^9 \prod_{d_1=1}^9 \left(\frac{10d_1 + d_2 + 1}{10d_1 + d_2} \right) \right]
 \end{aligned}$$

Y ahora el trabajo se reduce a demostrar que los dos productorios involucrados en la última expresión son iguales a 10.

Se obtiene que:

$$\begin{aligned}
 \prod_{d_2=0}^9 \prod_{d_1=1}^9 \left(\frac{10d_1 + d_2 + 1}{10d_1 + d_2} \right) &= \prod_{d_2=0}^9 \frac{(11+d_2)(21+d_2)\cdots(91+d_2)}{(10+d_2)(20+d_2)\cdots(90+d_2)} \\
 \Rightarrow \prod_{d_2=0}^9 \prod_{d_1=1}^9 \left(\frac{10d_1 + d_2 + 1}{10d_1 + d_2} \right) &= \left[\frac{11\cdots 91}{10\cdots 90} \right] \left[\frac{12\cdots 92}{11\cdots 91} \right] \cdots \left[\frac{20\cdots 100}{19\cdots 99} \right]
 \end{aligned}$$

Al simplificar obtenemos que:

$$\begin{aligned}
 \prod_{d_2=0}^9 \prod_{d_1=1}^9 \left(\frac{10d_1 + d_2 + 1}{10d_1 + d_2} \right) &= \left[\frac{20 \times 30 \times 40 \cdots 90 \times 100}{10 \times 20 \times 30 \cdots 80 \times 90} \right] \\
 \prod_{d_2=0}^9 \prod_{d_1=1}^9 \left(\frac{10d_1 + d_2 + 1}{10d_1 + d_2} \right) &= \frac{100}{10} = 10
 \end{aligned}$$

justo lo que se quería demostrar.

Con lo cual, se comprueba que $\sum_{d_2=0}^9 P(d_2) = 1$

Y que la distribución logarítmica del segundo dígito significativo, efectivamente es una distribución de probabilidad.

Distribución acumulada del segundo dígito.

Proposición 1.8.3.2: La distribución a acumulada del segundo dígito significativo se define de la siguiente forma:

$$(8) \quad \sum_{d_2=0}^k P(d_2 \leq k) = \sum_{d_2=0}^k \sum_{d_1=1}^9 \log_{10} \left(1 + (10d_1 + d_2)^{-1} \right)$$

De manera similar a la demostración del segundo axioma de probabilidad, podemos determinar la distribución acumulada para el segundo dígito significativo, es decir, determinar la probabilidad de que el valor del segundo dígito significativo sea menor o igual a un k , esto es.

$$\sum_{d_2=0}^k P(d_2 \leq k) = \sum_{d_2=0}^k \sum_{d_1=1}^9 \log_{10} \left(1 + (10d_1 + d_2)^{-1} \right)$$

La cual se reduce a:

$$\sum_{d_2=0}^k P(d_2 \leq k) = \log_{10} \left[\frac{(11+k) \cdots (91+k)}{10 \times 20 \times \cdots \times 90} \right]$$

1.8.4 Ley del k -ésimo dígito significativo.

Proposición 1.8.4.1: La función de distribución del k -ésimo dígito significativo, viene dada por la siguiente fórmula:

$$(9) \quad P(d_k) = \sum_{d_{k-1}=0}^9 \cdots \sum_{d_2=0}^9 \sum_{d_1=1}^9 \log_{10} \left[1 + \left(\sum_{i=1}^k d_i 10^{k-i} \right)^{-1} \right]$$

Con un razonamiento muy similar al utilizando en la proposición 1.7.3.1, podemos obtener la función de distribución para el k -ésimo dígito significativo.

1.8.5 Probabilidades condicionadas de los dígitos.

1.8.5.1 Probabilidad condicionada del segundo dígito. Al igual que en toda distribución de probabilidad conjunta podemos, en este caso, determinar la probabilidad de que el segundo dígito significativo tome el valor d_2 dado que el primero ya ha tomado el valor de d_1 .

Para eso usamos simplemente la definición de probabilidad condicionada para el caso de dos variables, la cual es la siguiente:

$$P(y/x) = \frac{P(y,x)}{P(x)} ; \text{ Probabilidad condicionada de } y \text{ dado el valor de } x$$

Donde:

$P(y, x)$ es la probabilidad conjunta de las variables x e y

y $P(x)$ es la distribución marginal de la variable x

Así, se tiene:

$$P(d_2 / d_1) = \frac{P(d_1, d_2)}{P(d_1)}$$

Como

$$\begin{aligned}
P(d_1, d_2) &= \log_{10} \left[1 + \left(\sum_{i=1}^2 d_i * 10^{2-i} \right)^{-1} \right] \\
&= \log_{10} \left[1 + (d_1 * 10 + d_2)^{-1} \right] \\
&= \log_{10} \left[1 + \frac{1}{d_1 * 10 + d_1} \right] \\
&= \log_{10} \left[\frac{d_1 * 10 + d_1 + 1}{d_1 * 10 + d_1} \right]
\end{aligned}$$

Y

$$\begin{aligned}
P(d_1) &= \log_{10} (1 + d_1^{-1}) \\
&= \log_{10} \left(\frac{1 + d_1}{d_1} \right)
\end{aligned}$$

Obtenemos finalmente que

$$P(d_2 / d_1) = \frac{\log_{10} \left(\frac{10 \cdot d_1 + d_2 + 1}{10 \cdot d_1 + d_2} \right)}{\log_{10} \left(\frac{d_1 + 1}{d_1} \right)}$$

En la siguiente tabla se muestra las probabilidades correspondientes.

Tabla 1.5. Probabilidad condicionada del segundo dígito.

PRIMER DÍGITO	SEGUNDO DÍGITO									
	0	1	2	3	4	5	6	7	8	9
1	0.1375	0.12553	0.11548	0.10692	0.09954	0.09311	0.08746	0.08246	0.078	0.074
2	0.12033	0.11473	0.10963	0.10496	0.10068	0.09673	0.09308	0.08969	0.08655	0.08361
3	0.11398	0.11036	0.10696	0.10377	0.10076	0.09792	0.09524	0.0927	0.09029	0.08801
4	0.11066	0.10799	0.10545	0.10303	0.10071	0.0985	0.09638	0.09435	0.0924	0.09054
5	0.10861	0.1065	0.10448	0.10252	0.10064	0.09883	0.09708	0.09539	0.09376	0.09218
6	0.10723	0.10548	0.1038	0.10216	0.10058	0.09904	0.09755	0.09611	0.0947	0.09334
7	0.10623	0.10474	0.1033	0.10189	0.10052	0.09919	0.0979	0.09663	0.0954	0.0942
8	0.10547	0.10418	0.10291	0.10168	0.10048	0.0993	0.09815	0.09703	0.09594	0.09486
9	0.10488	0.10373	0.10261	0.10151	0.10044	0.09939	0.09836	0.09735	0.09636	0.09539

1.8.5.2 Probabilidad condicionada del k -ésimo dígito. De una manera muy similar a la utilidad para la obtención de la probabilidad condicionada del segundo dígito podemos determinar la probabilidad que el k -ésimo dígito significativo tome el valor de d_k dado que los primeros $(k-1)$ dígitos han tomado los valores d_1, d_2, \dots, d_{k-1} .

Para obtener dicha probabilidad hacemos uso de la probabilidad condicionada generalizada, la cual es bastante similar para el caso de dos variables:

De esta manera tendremos que

$$P(d_k / d_1, \dots, d_j) = \frac{P(d_1, d_2, \dots, d_k)}{P(d_1, d_2, \dots, d_j)}; \quad 1 \leq j \leq k-1$$

De los resultados anteriores sabemos que

$$\begin{aligned} P(d_1, d_2, \dots, d_k) &= \log_{10} \left[1 + \frac{1}{\sum_{i=1}^k d_i * 10^{k-i}} \right] \\ &= \log_{10} \left[\frac{1 + \sum_{i=1}^k d_i * 10^{k-i}}{\sum_{i=1}^k d_i * 10^{k-i}} \right] \end{aligned}$$

Y también que

$$P(d_1, d_2, \dots, d_{k-1}) = \log_{10} \left[1 + \frac{1}{\sum_{i=1}^{k-1} d_i * 10^{(k-1)-i}} \right]$$

$$= \log_{10} \left[\frac{1 + \sum_{i=1}^{k-1} d_i 10^{(k-1)-i}}{\sum_{i=1}^{k-1} d_i 10^{(k-1)-i}} \right]$$

Con lo que obtenemos finalmente que

$$P(d_k / d_1, \dots, d_{k-1}) = \frac{\log_{10} \left[\frac{1 + \sum_{i=1}^k d_i 10^{k-i}}{\sum_{i=1}^k d_i 10^{k-i}} \right]}{\log_{10} \left[\frac{1 + \sum_{i=1}^{k-1} d_i 10^{(k-1)-i}}{\sum_{i=1}^{k-1} d_i 10^{(k-1)-i}} \right]}$$

1.9 Generalización de la Ley de Benford: la distribución mantisa.

Ya hemos probado que la distribución logarítmica conjunta de los primeros k dígitos significativos d_1, d_2, \dots, d_k ($\forall k \in \mathbb{N}$) se define por:

$$P(d_1, \dots, d_k) = \log_{10} \left[1 + \left(\sum_{i=1}^k d_i \cdot 10^{k-i} \right)^{-1} \right] \quad (6)$$

para todo entero positivo k , todo $d_1 \in \{1, \dots, 9\}$, y todo $d_j \in \{0, 1, \dots, 9\}$, $j = 2, \dots, k$.

Sin embargo, en su forma más general, la Ley de Benford es una declaración acerca de las distribuciones de las mantisas; la cual, evidentemente, pertenece al campo de los números reales:

$$(10) \quad \Pr(\text{mantisa} \leq x) = \log_{10} x; \quad x \in [1, 10)$$

Donde la *mantisa* de un número real x es el número obtenido de x cambiando el punto decimal al lugar inmediato siguiente al primer dígito significativo (diferente de cero). Ejemplos:

$$M(9) = M(0.09) = M(90) = 9$$

$$M(3.1) = M(0.0031) = M(3100) = 3.1$$

$$M(1.1) = M(11) = M(11000) = 1.1$$

Definición 1.9.1. La *función mantisa* M es la función $M : \mathbb{R}^+ \rightarrow [1,10)$ tal que $M(x) = r$, donde r es el único número en $[1,10)$ que cumple $x = r \cdot 10^n$ para algún $n \in \mathbb{Z}$.

Es fácil visualizar la equivalencia de las ecuaciones (10) y (6).

Proposición 1.9.2. La distribución logarítmica conjunta de los primeros k dígitos significativos d_1, d_2, \dots, d_k ($\forall k \in \mathbb{N}$) se puede generalizar al caso continuo mediante la distribución de probabilidad de la *mantisa* M de la siguiente forma:

$$P(M(x) \leq r) = \log_{10}(r), \text{ donde } r \in [1,10) \quad (10)$$

Demostración:

Consiste en probar la equivalencia de las ecuaciones (6) y (10). Así:

1. (6) \Rightarrow (10). Probaremos que (10) es válida; asumiendo que (6) se cumple.

i) Supongamos que r tiene solamente un dígito significativo, es decir $r = d_1$:

$$P(M \leq r) = \begin{cases} 0; & \text{si } d_1 = 1 \\ \log_{10}(d_1); & \text{si } 1 < d_1 \leq 10 \end{cases}$$

Así: Si $d_1 = 1$ entonces $P(M \leq r) = P(M \leq 1) = P(M = 1) = 0$, porque la distribución de probabilidad de M es una función de densidad. La probabilidad de que una variable continua tome un valor en particular es cero.

Si $1 < d_1 \leq 10$ entonces $P(M \leq r) = P(D_1 \leq d_1 - 1) = \log_{10}(d_1) = \log_{10}(r)$, utilizando la fórmula de la función acumulada para el primer dígito significativo.

ii) Supongamos que r tiene más de un dígito significativo. Así: $r = d_1.d_2d_3\dots d_k$

En notación decimal:

$$r = d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}d_k = \sum_{i=1}^k 10^{-(i-1)}d_i, \text{ con } d_1 > 1, d_2 > 0, \dots, d_k > 0$$

Luego,

$$\begin{aligned} P(M \leq r) &= P(M \leq d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}d_k) \\ &= P(D_1 \leq d_1 - 1) \\ &\quad + P(D_1 = d_1, D_2 \leq d_2 - 1) \\ &\quad + P(D_1 = d_1, D_2 = d_2, D_3 \leq d_3 - 1) \\ &\quad \dots \\ &\quad + P(D_1 = d_1, D_2 = d_2, \dots, D_k \leq d_k - 1) \end{aligned}$$

$$\begin{aligned} P(M \leq r) &= P(D_1 \leq d_1 - 1) \\ &\quad + \sum_{0 \leq d'_2 \leq d_2 - 1} P(D_1 = d_1, D_2 = d'_2) \\ &\quad + \sum_{0 \leq d'_3 \leq d_3 - 1} P(D_1 = d_1, D_2 = d_2, D_3 = d'_3) \\ &\quad \dots \\ &\quad + \sum_{0 \leq d'_k \leq d_k - 1} P(D_1 = d_1, D_2 = d_2, \dots, D_k = d'_k) \end{aligned}$$

Según la hipótesis de la distribución conjunta de los k dígitos tenemos:

$$\begin{aligned}
P(M \leq r) &= \log_{10}(d_1) \\
&+ \sum_{0 \leq d'_2 \leq d_2-1} \log_{10} \left(1 + \frac{1}{10d_1 + d'_2} \right) \\
&+ \sum_{0 \leq d'_3 \leq d_3-1} \log_{10} \left(1 + \frac{1}{10^2 d_1 + d_2 + d'_3} \right) \\
&\dots \\
&+ \sum_{0 \leq d'_k \leq d_k-1} \log_{10} \left(1 + \left(\sum_{i=1}^{k-1} 10^{k-i} d_i + d'_k \right)^{-1} \right)
\end{aligned}$$

Analicemos separadamente cada sumatorio de la ecuación anterior, así:

$$\begin{aligned}
\sum_{0 \leq d'_2 \leq d_2-1} \log_{10} \left(1 + \frac{1}{10d_1 + d'_2} \right) &= \sum_{0 \leq d'_2 \leq d_2-1} \log_{10} \left(\frac{10d_1 + d'_2 + 1}{10d_1 + d'_2} \right) \\
&= \log_{10} \left[\left(\frac{10d_1 + 1}{10d_1} \right) \cdot \left(\frac{10d_1 + 2}{10d_1 + 1} \right) \cdot \dots \cdot \left(\frac{10d_1 + d_2}{10d_1 + d_2 - 1} \right) \right] \\
&= \log_{10} \left(\frac{10d_1 + d_2}{10d_1} \right)
\end{aligned}$$

Ahora, siguiendo el mismo razonamiento con los sumandos de la expresión anterior, se llega al resultado esperado. Por ejemplo, para el último sumando de $P(M \leq r)$,

$$\sum_{0 \leq d'_k \leq d_k-1} \log_{10} \left(1 + \left(\sum_{i=1}^{k-1} 10^{k-i} d_i + d'_k \right)^{-1} \right) = \log_{10} \left(\frac{\sum_{i=1}^k 10^{k-i} d_i}{\sum_{i=1}^{k-1} 10^{k-i} d_i} \right)$$

Por lo que:

$$\begin{aligned}
P(M \leq r) &= \log_{10}(d_1) + \log_{10}\left(\frac{10d_1 + d_2}{10d_1}\right) + \log_{10}\left(\frac{10^2 d_1 + 10d_2 + d_3}{10^2 d_1 + 10d_2}\right) + \dots + \log_{10}\left(\frac{\sum_{i=1}^k 10^{k-i} d_i}{\sum_{i=1}^{k-1} 10^{k-i} d_i}\right) \\
&= \log_{10}\left[(d_1) \cdot \left(\frac{10d_1 + d_2}{10d_1}\right) \cdot \left(\frac{10^2 d_1 + 10d_2 + d_3}{10^2 d_1 + 10d_2}\right) \dots \left(\frac{10^{k-1} d_1 + 10^{k-2} d_2 + \dots + 10d_{k-1} + d_k}{10^{k-1} d_1 + 10^{k-2} d_2 + \dots + 10d_{k-1}}\right) \right] \\
&= \log_{10}\left[(d_1) \cdot \left(\frac{10d_1 + d_2}{10(d_1)}\right) \cdot \left(\frac{10^2 d_1 + 10d_2 + d_3}{10(10d_1 + d_2)}\right) \dots \left(\frac{10^{k-1} d_1 + 10^{k-2} d_2 + \dots + 10d_{k-1} + d_k}{10(10^{k-2} d_1 + 10^{k-3} d_2 + \dots + d_{k-1})}\right) \right] \\
&= \log_{10}\left(\frac{10^{k-1} d_1 + 10^{k-2} d_2 + \dots + 10d_{k-1} + d_k}{10^{k-1}}\right) \\
&= \log_{10}(d_1 + 10^{-1} d_2 + \dots + 10^{-k+2} d_{k-1} + 10^{-k+1} d_k) \\
&= \log_{10}\left(\sum_{i=1}^k 10^{-(i-1)} d_i\right) \\
&= \log_{10}(r)
\end{aligned}$$

Sin embargo, si uno o varios de los d_j ($j > 1$) son nulos o si $d_1 = 1$ el resultado sigue siendo válido.

2. (10) \Rightarrow (6) . Probaremos que (6) es válida; asumiendo que (10) se cumple.

Es decir, podemos deducir la fórmula de la distribución logarítmica conjunta de los primeros k dígitos significativos d_1, d_2, \dots, d_k ($\forall k \in \mathbb{N}$) a partir de la distribución de probabilidad de la función mantisa. Así

$$\begin{aligned}
P(D_1 = d_1, D_2 = d_2, \dots, D_k = d_k) &= P(d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}d_k \leq M < \\
&\quad d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}(d_k + 1)) \\
&= \log_{10} (d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}(d_k + 1)) \\
&\quad - \log_{10} (d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}d_k) \\
&= \log_{10} \left(\frac{d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}(d_k + 1)}{d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}d_k} \right) \\
&= \log_{10} \left(1 + \frac{10^{-k+1}}{d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}d_k} \right) \\
&= \log_{10} \left(1 + \frac{10^{-(k-1)}}{d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}d_k} \right) \\
&= \log_{10} \left(1 + \frac{1}{10^{k-1}} \cdot \frac{1}{(d_1 + 10^{-1}d_2 + 10^{-2}d_3 + \dots + 10^{-k+1}d_k)_k} \right) \\
&= \log_{10} \left(1 + \frac{1}{(10^{k-1}d_1 + 10^{k-2}d_2 + \dots + 10d_{k-1} + d_k)} \right) \\
&= \log_{10} \left(1 + \left(\sum_{i=1}^k 10^{k-i} d_i \right)^{-1} \right)
\end{aligned}$$

Como queríamos demostrar.

Finalmente, la distribución logarítmica conjunta de los primeros k dígitos significativos d_1, d_2, \dots, d_k se puede generalizar al caso continuo mediante la distribución de probabilidad de la *mantisa* M .

CAPÍTULO II
EXPLICACIÓN DE
LA LEY DE BENFORD

2.1 Introducción

En este capítulo se explica el fundamento matemático y estadístico de la Ley de Benford, principalmente el de definir el espacio propio de probabilidad, ya que es una de las claves principales para entender la Ley, también, estudiaremos los aportes desarrollado por algunos autores en su intento de explicarla, tales como; Theodore Hill, Roger Pinkhan, Mark Nigrini, entre otros.

En la primera sección se define y se dan algunos ejemplos de la Función Mantisa y la Sigma Álgebra, además de presentar una breve relación de éstas; en la sección segunda, se estudian y se demuestran algunos lemas de interés sobre la Sigma Álgebra los cuales estarán ligados a las secciones siguientes; en la tercera sección se presentan las funciones de distribución para los primeros dígitos en cualquier base; en la cuarta sección se presentan la Escala Invariante, algunos ejemplos de estas distribuciones y se finaliza con un teorema muy importante que trata sobre la unicidad de la Escala Invariante; en la quinta sección se presenta la Base invariante, algunos ejemplos y definiciones claves; en la sexta y última sección se presentan dos teoremas muy fuertes sobre la Base Invariante y algunas definiciones y proposiciones necesarias para la demostración de éstos.

2.2 La Función Mantisa y la Sigma Álgebra

Antes de pasar a definir formalmente la función mantisa y la sigma álgebra presentamos la siguiente definición.

Definición 2.2.1: El trío $(\Omega, \mathfrak{M}, P)$ recibe el nombre de **espacio probabilístico o espacio de probabilidad**, donde Ω es un espacio muestral, \mathfrak{M} es una σ -álgebra de sucesos sobre Ω y P es la medida de probabilidad sobre \mathfrak{M} .

Siguiendo la definición de espacio probabilístico en este punto es importante definir el conjunto \mathfrak{M} (σ -álgebra), ya que se tiene una noción sobre el resto de los elementos (Ω, P) del espacio de probabilidad para todo fenómeno que sigue la Ley de Benford, el espacio muestral, generado por un fenómeno de Benford es \mathbb{R}^+ y la medida de incertidumbre del espacio probabilístico, es la función de distribución del primer, segundo, ..., k -ésimo dígito significativo, de un número determinado.

Definición 2.2.2: Para cada entero $b > 1$, la función mantisa (base b) M_b , es la función $M_b: \mathbb{R}^+ \rightarrow [1, b)$ de tal manera que $M_b(x) = r$, donde r es un único número que está en el intervalo $[1, b)$ con $x = rb^n$ para algún $n \in \mathbb{Z}$.

Nota: La restricción del dominio de M_b a \mathbb{R}^+ , es sólo por conveniencia y este puede extenderse a todos los reales por medio de $M_b(0) = 0$ y $M_b(-x) = M_b(x)$, pero este caso no se abordará ya que hemos visto con anterioridad que si algún fenómeno sigue la Ley de Benford es imposible que se observen valores negativos, pero esto dependerá que tipo de interpretaciones les demos a éstos (una ganancia de -3 también puede interpretarse como una pérdida de +3), por el momento trabajaremos únicamente con valores positivos. Obsérvese además que la definición 1.9.1 es un caso particular cuando $b = 10$.

Definición 2.2.3: Para $E \subset [1, b)$, sea

$$\langle E \rangle_b = M_b^{-1}(E) = \biguplus_{n \in \mathbb{Z}} b^n E \subset \mathbb{R}^+$$

Siendo esta la mantisa σ -álgebra y comúnmente denotada por \mathfrak{M}_b y es la σ -álgebra sobre los \mathbb{R}^+ generada por M_b .

Donde \uplus implicará la unión finita (infinita) de elementos disjuntos dos a dos.

Ejemplos 2.2.4 (Función Mantisa)

1) $M_{10}(9) = 9 = M_{100}(9)$, $9 \in [1, 10) \subset [1, 100)$; dado que $9 = 9 \cdot 10^0 = 9 \cdot 100^0$, $0 \in \mathbb{Z}$

2) $M_2(9) = 1.001$, $1.001 \in [1, 2)$ dado que $9 = \frac{9}{8} \cdot 2^3$, $3 \in \mathbb{Z}$

3) $M_{10}(21.14) = 2.114$, dado que $2.114 = 2.114 \cdot 10^1$, $1 \in \mathbb{Z}$

4) si $E = [1, 5) \subset [1, 10)$, $b = 10$, $n \in \mathbb{Z} : (-\infty, \infty)$, entonces,

$$\begin{aligned} \langle E \rangle_{10} &= M_{10}^{-1}(E) = \uplus_{n \in \mathbb{Z}} 10^n [1, 5) \\ &= \dots \uplus [0.01, 0.05) \uplus [0.1, 0.5) \uplus [1, 5) \uplus [10, 50) \uplus [100, 500) \uplus \dots \end{aligned}$$

5) $\langle \{1\} \rangle_{10} = \{10^n : n \in \mathbb{Z}\}$, porque $\langle \{1\} \rangle_{10} = M_{10}^{-1}(\{1\}) = \uplus_{n \in \mathbb{Z}} 10^n \{1\} = \{10^n\} \subset \mathbb{R}^+$

6) $\langle [1, b) \rangle_b = \mathbb{R}^+$, porque $\langle [1, b) \rangle_b = M_b^{-1}([1, b)) = \uplus_{n \in \mathbb{Z}} b^n [1, b) = \mathbb{R}^+$, $\forall b > 1$, $b \in \mathbb{Z}$;

La interpretación que podemos darle a la definición 2.2.3 podría ser la siguiente: la σ -álgebra mantisa es en realidad el conjunto de todos los números reales cuyo primer dígito significativo se encuentra en el intervalo E , o más bien el conjunto de todos los números reales tales que su mantisa se encuentra en E .

Observación: se verifica que la σ -álgebra mantisa es en realidad una sigma álgebra, para ello debemos probar las siguientes propiedades.

i) $\Omega \subset \mathfrak{M}_b$

ii) $\forall S \in \mathfrak{M}_b$ es $S^c \in \mathfrak{M}_b$

iii) $\forall \{S_k\}_{k \in \mathbb{N}} \in \mathfrak{M}_b$ es $\bigcup_{k \in \mathbb{N}} S_k \in \mathfrak{M}_b$

Donde $\mathfrak{M}_b = \langle E \rangle_b = \left\{ \uplus_{n \in \mathbb{Z}} E b^n : E \subset [1, b) \right\}$

Verificación

i) ¿ $\Omega \subset \mathfrak{M}_b$?

Notemos primeramente que $\Omega = \mathbb{R}^+$, esto es debido a la definición 2.1.3

$$\Rightarrow \mathbb{R}^+ \subset [0, +\infty) = [0, 1) \uplus [1, +\infty)$$

Pero

$$[0, 1) = \dots \uplus [1, b) b^{-3} \uplus [1, b) b^{-2} \uplus [1, b) b^{-1} = \uplus_{n < 0} [1, b) b^n$$

$$[1, +\infty) = [1, b) b^0 \uplus [1, b) b^1 \uplus [1, b) b^2 \uplus \dots = \uplus_{n \geq 0} [1, b) b^n$$

$$\begin{aligned} \Rightarrow [0, 1) \cup [1, +\infty) &= \left(\uplus_{n < 0} [1, b) b^n \right) \uplus \left(\uplus_{n \geq 0} [1, b) b^n \right) \\ &= \dots \uplus [1, b) b^{-2} \uplus [1, b) b^{-1} \uplus [1, b) b^0 \uplus [1, b) b^1 \uplus [1, b) b^2 \uplus \dots \\ &= \uplus_{n \in \mathbb{Z}} [1, b) b^n \end{aligned}$$

$$\Rightarrow \mathbb{R}^+ \subset \uplus_{n \in \mathbb{Z}} b^n [1, b) = \mathfrak{M}_b$$

$$\therefore \mathbb{R}^+ \subset \mathfrak{M}_b$$

Que es justo lo que queríamos demostrar.

ii) ¿ $\forall S \in \mathfrak{M}_b$ es $S^c \in \mathfrak{M}_b$?

Sea $A = [a_1, b_1)$; con $1 \leq a_1 < b_1 \leq b$

Por $A \in \mathfrak{M}_b$ se tiene que

$$\begin{aligned}\mathfrak{M}_b(A) &= \mathfrak{M}_b([a_1, b_1]) \\ &= \biguplus_{n \in \mathbb{Z}} [a_1, b_1] b^n \subset \mathbb{R}^+\end{aligned}$$

Entonces $A^c = [1, a_1) \uplus [b_1, b)$ por otra parte $A \uplus A^c = [1, b)$ y puesto que $\mathfrak{M}_b([1, b)) = \mathbb{R}^+$, obtendremos lo siguiente:

$$\begin{aligned}\Rightarrow \mathfrak{M}_b([1, b)) &= \biguplus_{n \in \mathbb{Z}} [1, b) b^n \\ &= \biguplus_{n \in \mathbb{Z}} \{[1, a_1) \uplus [a_1, b_1) \uplus [b_1, b)\} b^n \\ &= \left[\biguplus_{n \in \mathbb{Z}} [1, a_1) b^n \right] \uplus \left[\biguplus_{n \in \mathbb{Z}} [a_1, b_1) b^n \right] \uplus \left[\biguplus_{n \in \mathbb{Z}} [b_1, b) b^n \right] \\ &= \left[\biguplus_{n \in \mathbb{Z}} [1, a_1) b^n \right] \uplus \left[\biguplus_{n \in \mathbb{Z}} [b_1, b) b^n \right] \uplus \left[\biguplus_{n \in \mathbb{Z}} [a_1, b_1) b^n \right] \\ &= \left[\biguplus_{n \in \mathbb{Z}} \{[1, a_1) \uplus [b_1, b)\} b^n \right] \uplus \left[\biguplus_{n \in \mathbb{Z}} [a_1, b_1) b^n \right] \\ &= \left[\biguplus_{n \in \mathbb{Z}} A^c b^n \right] \uplus \left[\biguplus_{n \in \mathbb{Z}} A b^n \right] \\ &= \left[\biguplus_{n \in \mathbb{Z}} A^c b^n \right] \uplus \mathfrak{M}_b(A) \\ &= \mathbb{R}^+\end{aligned}$$

Esto se cumplirá siempre y cuando $\mathfrak{M}_b(A^c) = \biguplus_{n \in \mathbb{Z}} A^c b^n$, lo cual implica que

$$A^c \in \mathfrak{M}_b$$

Que es justo lo que queríamos demostrar.

iii) ¿ $\forall \{S_k\}_{k \in \mathbb{N}} \in \mathfrak{M}_b$ es $\bigcup_{k \in \mathbb{N}} S_k \in \mathfrak{M}_b$?

Tomemos $E \subset [1, b)$ y hagamos una serie de particiones infinitas y disjuntas del intervalo E , es decir $(E_1, E_2, E_3, \dots, E_K \subset E \subset [1, b))$ tal que $E_1 \cup E_2 \cup E_3 \cup \dots \cup E_K = E$, para $k \in \mathbb{N}$

$$\Rightarrow S_1 = \biguplus_{n \in \mathbb{Z}} b^n E_1, S_2 = \biguplus_{n \in \mathbb{Z}} b^n E_2, S_3 = \biguplus_{n \in \mathbb{Z}} b^n E_3 \dots S_k = \biguplus_{n \in \mathbb{Z}} b^n E_k \subset \biguplus_{n \in \mathbb{Z}} b^n E$$

Es decir que $\{S_k\}_{k \in \mathbb{N}} \in \mathfrak{M}_b$

Al realizar las uniones de los S_k tendremos:

$$\begin{aligned} \bigcup_{k \in \mathbb{N}} S_k &= S_1 \cup S_2 \cup S_3 \dots \cup S_k \\ &= \biguplus_{n \in \mathbb{Z}} b^n E_1 \cup \biguplus_{n \in \mathbb{Z}} b^n E_2 \cup \biguplus_{n \in \mathbb{Z}} b^n E_3 \cup \dots \cup \biguplus_{n \in \mathbb{Z}} b^n E_k \\ &= \biguplus_{n \in \mathbb{Z}} (E_1 \cup E_2 \cup E_3 \cup \dots \cup E_k) \\ &= \biguplus_{n \in \mathbb{Z}} b^n E \in \mathfrak{M}_b \\ \therefore \bigcup_{k \in \mathbb{N}} S_k &\subset \mathfrak{M}_b \end{aligned}$$

Que es justo lo que queríamos demostrar.

Por tanto \mathfrak{M}_b es una σ -álgebra en \mathbb{R}^+ .

Nota Aclaratoria: el significado de $S \in \mathfrak{M}_b$ es que $\mathfrak{M}_b(S) = \biguplus_{k \in \mathbb{Z}} b^k S$

Observación: es fácil verificar que \mathfrak{M}_b es la σ -álgebra generada por M_b .

i) Primero verifiquemos que $\mathfrak{M}_b \subset \sigma(M_b)$

Si $E = [1, b)$ entonces

$$\begin{aligned}\mathfrak{M}_b(E) &= \mathfrak{M}_b([1, b)) \\ &= \bigsqcup_{n \in \mathbb{Z}} [1, b) b^n \\ &= \bigsqcup_{n \in \mathbb{Z}} [b^n, b^{n+1}) \\ &= \mathbb{R}^+\end{aligned}$$

La cual también es equivalente a la siguiente expresión.

$$\mathfrak{M}_b(E) = \bigsqcup_{i \in \mathbb{N}} r_i; \quad r_1 < r_2 < r_3 < r_4 < \cdots < r_n < r_{n+1} < \cdots$$

Usando la notación científica podemos escribir cada r_i de la siguiente forma $r_i = r_i^+ b^n$ para algún $n \in \mathbb{Z}$ y $r_i^+ \in [1, b)$, ahora si nos damos cuenta la notación científica concuerda exactamente con la definición de la función mantisa y por tanto podemos asegurar que

$$\begin{aligned}\mathfrak{M}_b(E) &= \bigsqcup_{i \in \mathbb{N}} M_b(r_i) \\ &= \bigsqcup_{i \in \mathbb{N}} r_i^+ b^n\end{aligned}$$

Es decir, que $\mathfrak{M}_b \subset \sigma(M_b)$

Si $E \subset [1, b)$ entonces

$\mathfrak{M}_b(E) = \biguplus_{n \in \mathbb{Z}} E b^n \subset \mathbb{R}^+$, el procedimiento sería el mismo al del caso anterior, teniendo en cuenta únicamente que en esta ocasión los $r_i \in E b^n$ y reescribiendo los $r_i = r_i^+ b^n$.

ii) Ahora verifiquemos que $M_b \subset \sigma(\mathfrak{M}_b)$

Sea $X \in \mathbb{R}^+$ con lo que $M(X) = r b^n$ para algún $n \in \mathbb{Z}$ y $r \in [1, b)$

$$\begin{aligned} \Rightarrow M_b(X) &= r \\ &= \bigcap_{n \in \mathbb{N}} \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \\ &= \bigcap_{n \in \mathbb{N}} \bigcap_{k \in \mathbb{N}} \left\langle \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \right\rangle_{b^k} \end{aligned}$$

La última de las igualdades es cierta debido a que si nos ponemos a realizar la intersección de σ -álgebra mantisa tales que contengan el mismo intervalo inicial E pero con diferentes bases (potencias de b) al final llegaremos a obtener el mismo intervalo de partida E, es decir;

$$\begin{aligned} \bigcap_{k \in \mathbb{N}} \left\langle \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \right\rangle_{b^k} &= \left\langle \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \right\rangle_b \bigcap \left\langle \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \right\rangle_{b^2} \bigcap \cdots \bigcap \left\langle \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \right\rangle_{b^k} \\ &= \left\{ \biguplus_{n \in \mathbb{Z}} b^n \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \right\} \bigcap \left\{ \biguplus_{n \in \mathbb{Z}} (b^2)^n \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \right\} \bigcap \cdots \bigcap \left\{ \biguplus_{n \in \mathbb{Z}} (b^k)^n \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \right\} \end{aligned}$$

$$\begin{aligned}
&= \left\{ \dots b^{-1} \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \uplus b^0 \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \uplus b^1 \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \uplus \dots \right\} \\
&\cap \left\{ \dots b^{-2} \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \uplus b^0 \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \uplus b^2 \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \uplus \dots \right\} \\
&\quad \vdots \\
&\cap \left\{ \dots b^{-k} \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \uplus b^0 \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \uplus b^k \left(r - \frac{1}{n}, r + \frac{1}{n} \right) \uplus \dots \right\} \\
&= \left(r - \frac{1}{n}, r + \frac{1}{n} \right)
\end{aligned}$$

En la próxima sección presentamos algunas propiedades muy interesantes que la σ -álgebra mantisa cumple, propiedades que están estrechamente relacionadas a la escala invariante y a la base invariante.

2.3 Lemas sobre la Sigma Álgebra

En el siguiente lema se presentan cuatro propiedades importantes que tiene que cumplir la σ -álgebra mantisa (\mathfrak{M}_b).

Lema 2.3.1: Para todo $b > 1$.

- i) $\langle E \rangle_b = \bigoplus_{k=0}^{n-1} \langle b^k E \rangle_{b^n}$
- ii) $\mathfrak{M}_b = \{ \langle E \rangle_b : E \in B[1, b] \}$ donde B es un conjunto de Borel
- iii) $\mathfrak{M}_b \subset \mathfrak{M}_{b^n} \subset B$ para todo $n \in \mathbb{N}$
- iv) \mathfrak{M}_b es cerrada bajo la multiplicación escalar, si $S \in \mathfrak{M}_b, a > 0$ entonces $aS \in \mathfrak{M}_b$

Prueba de la propiedad i)

Para todo $n \in \mathbb{N}$ y $E \subset [1, b)$, $\langle E \rangle_b = \bigoplus_{k=0}^{n-1} \langle b^k E \rangle_{b^n}$

Sabemos por la definición 2.1.3 que $\langle E \rangle_b = \bigoplus_{n \in \mathbb{Z}} b^n E$, por lo tanto

$$\begin{aligned} \bigoplus_{k=0}^{n-1} \langle b^k E \rangle_{b^n} &= \bigoplus_{k=0}^{n-1} \bigoplus_{j \in \mathbb{Z}} (b^n)^j (b^k E) \\ &= \left(\bigoplus_{j \in \mathbb{Z}} b^{nj} E \right) \uplus \left(\bigoplus_{j \in \mathbb{Z}} b^{nj+1} E \right) \uplus \left(\bigoplus_{j \in \mathbb{Z}} b^{nj+2} E \right) \uplus \dots \\ &\quad \uplus \left(\bigoplus_{j \in \mathbb{Z}} b^{nj+(n-2)} E \right) \uplus \left(\bigoplus_{j \in \mathbb{Z}} b^{nj+(n-1)} E \right) \\ &= \bigoplus_{j \in \mathbb{Z}} b^j E \\ &= \langle E \rangle_b \end{aligned}$$

La igualdad se cumple debido a que si nos ponemos a pensar detenidamente (observando exclusivamente los exponentes de b) en la primera unión se encuentran todos los múltiplos de n , en la segunda los múltiplos de n sumándoles 1, y así sucesivamente, hasta que en la última se encuentran los múltiplos de n restándoles 1, por lo que al unir las se tendrán todos los exponentes enteros, es decir, la definición original.

Con lo que $\langle E \rangle_b = \bigoplus_{k=0}^{n-1} \langle b^k E \rangle_{b^n}$

Justo lo que queríamos demostrar.

Prueba de la propiedad ii)

$\mathfrak{M}_b = \{ \langle E \rangle_b : E \in B[1, b) \}$ donde B es un conjunto de Borel.

Sabemos por los resultados anteriores que $\langle E \rangle_b = \bigoplus_{n \in \mathbb{Z}} b^n E$ es una sigma álgebra, por lo que basta probar que $E \in B[1, b)$.

Donde $B[1, b)$ no es más que la σ -álgebra de Borel engendrada por la clase de intervalo $[1, b)$.

Observemos además que $[1, b)$ también está en el conjunto de Borel $(B[1, b))$, es decir que: $[1, b) \in B[1, b)$, esto debido a lo siguiente

$$[1, b) = \bigcup_{k \in \mathbb{N}} \left[1, b - \frac{1}{k} \right) = \bigcap_{k \in \mathbb{N}} \left[1, b + \frac{1}{k} \right)$$

Pero sabemos que $E \subset [1, b)$, por la definición de la σ -álgebra mantisa entonces tendremos que:

$$E \subset [1, b) \in B[1, b)$$

$$\Rightarrow E \in B[1, b)$$

Con lo que $\mathfrak{M}_b = \{ \langle E \rangle_b : E \in B[1, b) \}$

Justo lo que queríamos demostrar.

Prueba de la propiedad iii)

$$\mathfrak{M}_b \subset \mathfrak{M}_{b^n} \subset B \text{ Para todo } n \in \mathbb{N};$$

Donde B es la σ -álgebra de Borel sobre \mathbb{R}^+

$$\mathfrak{M}_b = \langle E \rangle_b = \biguplus_{m \in \mathbb{Z}} b^m E; \quad E \subset [1, b)$$

$$\mathfrak{M}_{b^n} = \langle E^* \rangle_{b^n} = \biguplus_{m \in \mathbb{Z}} b^m E^*; \quad E^* \subset [1, b^n)$$

$$E \subset [1, b) \subset E^* \subset [1, b^n); \quad \forall n \in \mathbb{N}$$

Multiplicando por b^m a la anterior expresión y luego aplicando unión numerable de disjuntos, obtenemos los siguientes resultados

$$\Rightarrow \biguplus_{m \in \mathbb{Z}} b^m E \subset \biguplus_{m \in \mathbb{Z}} b^m E^*$$

$$\Rightarrow \mathfrak{M}_b \subset \mathfrak{M}_{b^n}$$

Con lo que $\mathfrak{M}_b \subset \mathfrak{M}_{b^n} \subset B$

Justo lo que queríamos demostrar.

Prueba de la propiedad iv)

$$S \in \mathfrak{M}_b, \quad a > 0 \text{ entonces, } aS \in \mathfrak{M}_b.$$

Primero veamos un caso particular para notar lo que sucede.

Sea $E = [1, 5)$, $b = 10$ y $a = 0.3$, entonces

$$\mathfrak{M}_{10}([1, 5)) = \biguplus_{n \in \mathbb{Z}} 10^n [1, 5), \text{ así}$$

$$\begin{aligned}
0.3 * \mathfrak{M}_{10}([1,5)) &= 0.3 * \biguplus_{n \in \mathbb{Z}} 10^n [1,5) \\
&= \biguplus_{n \in \mathbb{Z}} 10^n 0.3 * [1,5) \\
&= \biguplus_{n \in \mathbb{Z}} 10^n [0.3,1.5) \\
&= \biguplus_{n \in \mathbb{Z}} 10^n \{ [0.3,1) \uplus [1,1.5) \} \\
&= \left\{ \biguplus_{n \in \mathbb{Z}} 10^n [3,10) \right\} \uplus \left\{ \biguplus_{n \in \mathbb{Z}} 10^n [1,1.5) \right\} \\
&= \mathfrak{M}_{10}([3,10)) \uplus \mathfrak{M}_{10}([1,1.5))
\end{aligned}$$

Ahora tomemos la misma base y el mismo intervalo pero con $a = 3$, tendremos que

$$\begin{aligned}
3 * \mathfrak{M}_{10}([1,5)) &= 3 * \biguplus_{n \in \mathbb{Z}} 10^n [1,5) \\
&= \biguplus_{n \in \mathbb{Z}} 10^n 3 * [1,5) \\
&= \biguplus_{n \in \mathbb{Z}} 10^n [3,15) \\
&= \biguplus_{n \in \mathbb{Z}} 10^n \{ [3,10) \uplus [10,15) \} \\
&= \left\{ \biguplus_{n \in \mathbb{Z}} 10^n [3,10) \right\} \uplus \left\{ \biguplus_{n \in \mathbb{Z}} 10^n [1,1.5) \right\} \\
&= \mathfrak{M}_{10}([3,10)) \uplus \mathfrak{M}_{10}([1,1.5))
\end{aligned}$$

De la anterior expresión podemos darnos cuenta que al multiplicar una constante por un elemento de \mathfrak{M}_b se obtiene la unión de dos elementos siempre de \mathfrak{M}_b , observemos que la restricción de $\alpha S \in \mathfrak{M}_b$ no restringe a tener el mismo intervalo inicial, si no más bien se restringe a la unión de dos intervalos reescritos de tal forma que se encuentre en el intervalo deseado ($[1,10)$ para nuestro caso),

además observemos que cuando tengamos distintas constantes pero que difieren únicamente en una potencia de la base se obtendrá el mismo resultado.

Con esto en mente la prueba toma el siguiente rumbo.

$$\text{Sea } S = \biguplus_{m \in \mathbb{Z}} b^m E \in \mathfrak{M}_b; \quad E \subset [1, b)$$

$$aS = \biguplus_{n \in \mathbb{Z}} ab^n E$$

Supongamos que $E = [1, \theta)$ con $\theta \leq b$

$$aS = \biguplus_{m \in \mathbb{Z}} ab^m E = \biguplus_{m \in \mathbb{Z}} b^m aE$$

$$\text{Sea } E^* = aE = [a, a\theta)$$

$$aS = \biguplus_{n \in \mathbb{Z}} b^n E^*$$

Ahora como hicimos en el caso particular partamos E^* en dos subintervalos y llamémosle E_1 y E_2 , los cuales tendrán la siguiente forma:

$$E_1 = [ab^k, b), \text{ para algún } k \in \mathbb{Z} \text{ y}$$

$$E_2 = [1, a\theta b^m), \text{ para algún } m \in \mathbb{Z}$$

Notemos que para el caso particular

$$E_1 = [0.3 * 10^1, 10), \text{ para algún } k = 1 \in \mathbb{Z} \text{ y}$$

$$E_2 = [1, 0.3 * 5 * 10^0), \text{ para algún } m = 0 \in \mathbb{Z}$$

Con lo que obtenemos que

$$\begin{aligned}
aS &= \uplus_{n \in \mathbb{Z}} b^n E^* \\
&= \uplus_{n \in \mathbb{Z}} b^n [a, a\theta) \\
&= \uplus_{n \in \mathbb{Z}} b^n \left\{ [ab^k, b) \uplus [1, a\theta b^m) \right\} \\
&= \left\{ \uplus_{n \in \mathbb{Z}} b^n [ab^k, b) \right\} \uplus \left\{ \uplus_{n \in \mathbb{Z}} b^n [1, a\theta b^m) \right\} \\
&= \mathfrak{M}_b \left([ab^k, b) \right) \uplus \mathfrak{M}_b \left([1, a\theta b^m) \right)
\end{aligned}$$

$$\therefore aS \in \mathfrak{M}_b$$

Justo lo que queríamos demostrar.

Definición 2.3.2. Para $b \in \mathbb{N} / \{1\}$ y $i \in \mathbb{N}$, $D_b^{(i)}(x)$ es el i -ésimo dígito significativo del número x (en la base b). Donde $\mathbb{N} / \{1\}$ representa el conjunto de los números naturales a excepción del número 1.

Ejemplos 2.3.3

$$\begin{aligned}
D_{10}^{(1)}(\pi) &= 3, \quad D_{10}^{(2)}(\pi) = 1, \quad D_{10}^{(3)}(\pi) = 4 \\
D_2^{(1)}(\pi) &= 1, \quad D_2^{(2)}(\pi) = 1, \quad D_2^{(3)}(\pi) = 0
\end{aligned}$$

Más formalmente $D_b^{(i)} : \mathbb{R}^+ \rightarrow \{0, 1, \dots, b-1\}$ y se define por $D_b^{(i)} = (g_b^{-1})_i \circ M_b$

donde M_b es la función mantisa y g_b es la función $g_b : \mathbb{R}^+ \rightarrow [1, b)$ tal que:

$$g_b(i_0, i_1, i_2, \dots) = \sum_{k=0}^{\infty} \frac{i_k}{b^k} \text{ para } i_0 \in \{1, 2, \dots, b-1\} \text{ e } i_k \in \{0, 1, 2, \dots, b-1\} \text{ para } k \in \mathbb{N}$$

Ejemplo 2.3.4

Calcular el primer dígito significativo del número π en base 10 ($D_{10}^{(1)}(\pi)$)

$$\begin{aligned}g_{10}(3,1,4,1,5,1,6\dots) &= \sum_{k=0}^{\infty} \frac{i_k}{10^k} \\&= \frac{i_0}{10^0} + \frac{i_1}{10^1} + \frac{i_2}{10^2} + \frac{i_3}{10^3} + \frac{i_4}{10^4} + \frac{i_5}{10^5} + \frac{i_6}{10^6} + \dots \\&= \frac{3}{10^0} + \frac{1}{10^1} + \frac{4}{10^2} + \frac{1}{10^3} + \frac{5}{10^4} + \frac{1}{10^5} + \frac{6}{10^6} + \dots \\&= 3 + 0.1 + 0.04 + 0.001 + 0.0005 + 0.00001 + 0.000006 + \dots \\&= \pi\end{aligned}$$

La inversa de la función g_{10} es:

$$g_{10}^{-1}(\pi) = (3,1,4,1,5,1,6,\dots)$$

Usando la definición 2.3.2 tendremos que:

$$\begin{aligned}D_{10}^{(1)}(\pi) &= \left(g_{10}^{-1}(M_{10}(\pi))\right)_1 \\&= \left(g_{10}^{-1}(\pi)\right)_1 \\&= 3\end{aligned}$$

Lema 2.3.5. Para cada entero $b > 1$, M_b es la σ -álgebra generada por

$$\{D_b^{(i)} : i \in \mathbb{N}\}$$

Prueba del lema

Para ello veamos si M_b puede escribirse en términos de la función $\{D_b^{(i)}\}$ y viceversa.

Sabemos que $M_b(x) = r$ por la definición de la función mantisa con $r \in [1, b)$

$M_b(x) = r$ (r puede expresarse como un sumatorio de términos)

$$= \sum_{k=0}^{\infty} \frac{i_k}{b^k} \text{ para todo } i_{k+1} = 0$$

$= g_b(i_0, i_1, i_2, \dots, i_{k-1}, i_k, \dots)$ por la definición de función g_b

$$= g_b(\{D_b^{(i)}\}) \text{ para } i \in \mathbb{N}$$

$$\Rightarrow M_b(x) = g_b(\{D_b^{(i)}\})$$

M_b puede escribirse en términos de la función $\{D_b^{(i)}\}$

Ahora veamos si $\{D_b^{(i)}\}$ puede escribirse en términos de M_b , utilizando la definición del i -ésimo dígito significativo para un número x en la base b tendremos que:

$\{D_b^{(i)}\} = \{(g_b^{-1})_i \circ M_b, i \in \mathbb{N}\}$; es el conjunto de todos los dígitos significativos de un número x en la base b .

Vemos que los $\{D_b^{(i)}\}$ puede expresarse en términos de los M_b , por tanto M_b es la σ -álgebra generada por $\{D_b^{(i)} : i \in \mathbb{N}\}$

2.4 Ley General del Dígito Significativo.

En el primer capítulo deducimos las leyes de distribución de los dígitos significativos en la base $b = 10$, ahora, en esta sección procedemos a generalizar las expresiones anteriormente encontrados a cualquier base distinta de 10, la cual es conocida como la Ley General del Dígito Significativo que está únicamente en Función de los $D_b^{(i)}$ (el i -ésimo dígito significativo en la base b).

Para la deducción de las fórmulas se procede igual que para el caso de la base $b = 10$, sólo que esta vez estamos en una base general b ($b > 1$) siguiendo los mismos pasos realizados anteriormente se llega a las siguientes expresiones.

Definición 2.4.1. La probabilidad de que el primer dígito significativo en la base b tome un valor igual a " d " viene dada por la siguiente expresión matemática.

$$P(D_b^{(1)} = d) = \log_b(1 + d^{-1})$$

Definición 2.4.2. La probabilidad de que los primeros k dígitos significativos en la base b tome respectivamente los valores de $d_1, d_2, d_3, \dots, d_k$ viene dada por la siguiente expresión matemática.

$$P\left(\bigcap_{i=1}^k [D_b^{(i)} = d_i]\right) = \log_b \left(1 + \left(\sum_{i=1}^k d_i b^{k-i} \right)^{-1} \right)$$

$$k \in \mathbb{N}$$

$$\text{Para todo } d_1 = 1, 2, \dots, 9$$

$$d_j = 0, 1, 2, \dots, 9 ; j \geq 2$$

Definición 2.4.3. La probabilidad de que el segundo dígito significativo en la base b tome un valor igual a " d_2 " viene dada por la siguiente expresión matemática.

$$P(d_2) = \sum_{d_1=1}^{b-1} \log_b(1 + (bd_1 + d_2)^{-1})$$

Notemos que cada una de las correspondientes expresiones obtenidas en la base 10, son casos especiales de las correspondientes que acabamos de encontrar.

Para facilitar un poco la notación, muy a menudo es conveniente identificar una medida de probabilidad sobre \mathfrak{M}_b (la sigma mantisa) con su representación canónica como una probabilidad de un $B[1, b)$.

Para tal efecto mostramos a continuación el siguiente lema.

Lema 2.4.4. La siguiente relación

$$P(\langle E \rangle_b) = \hat{P}(E); \quad E \in B[1, b)$$

Define una correspondencia entre la medida de probabilidad P en (R^+, \mathfrak{M}_b) y una medida de probabilidad \hat{P} en $[1, b)$

Antes de realizar la demostración es necesario dar la siguiente definición

Definición 2.4.5. Un Isomorfismo es una correspondencia biunívoca entre dos conjuntos con la misma estructura. Cada elemento del primer conjunto se puede poner en relación con un elemento del segundo mediante una operación. Por ejemplo, la multiplicación por una constante entera relaciona un conjunto de enteros con otro conjunto de enteros. Una relación que no se ciñe a esto, por ejemplo, el conjunto de números obtenidos a partir de la extracción de la raíz cuadrada de otro conjunto no forman un isomorfismo.

Prueba del Lema 2.4.4

Verifiquemos que efectivamente $\langle E \rangle_b$ y E forman un Isomorfismo.

Es trivial, pues una vez que se fija E , se obtiene un único elemento $\langle E \rangle_b$, para ser más específicos se obtendrá la unión numerable de Isomorfismos $(b^n E)$.

Ahora por ser $\langle E \rangle_b$ y E σ -álgebras, existirá una medida de probabilidad definida en ellas, sean P y \hat{P} , respectivamente. Por otra parte $\langle E \rangle_b$ y E forman un isomorfismo, con lo cual se completa la prueba.

2.5 Escala Invariante

La *Escala Invariante* es una de las asunciones más fuertes de la Ley de Benford. Este aporte fue gracias a Roger Pinkham, su idea fue la siguiente; de existir una ley de frecuencia de los dígitos esta ley debería de ser independiente del sistema de unidades, es decir, que multiplicando todos los números por una constante no afectaría la medida de probabilidad. Esta propiedad es llamada *Escala Invariante* (algunos autores suelen llamarla también como Invarianza de Escala), es decir, si $S \in \mathfrak{M}_b$ y recordando el lema 2.2.1 tenemos que $\alpha S \in \mathfrak{M}_b$ para todo $\alpha > 0$, dado que S y αS son subconjuntos \mathfrak{M}_b , el espacio de medida para ambos es $(\mathbb{R}^+, \mathfrak{M}_b)$ esto nos lleva a que la medida de probabilidad sobre S y αS deberían de ser igual, es decir, que $P(S) = P(\alpha S)$.

Definición 2.5.1. Una medida de probabilidad P en $(\mathbb{R}^+, \mathfrak{M}_b)$ es escala invariante para $\alpha > 0$ si $P(S) = P(\alpha S)$ para todo $S \in \mathfrak{M}_b$, siempre que α no sea una potencia racional de b .

Ejemplo 2.5.2: Es fácil verificar que $P_b(\langle [1, r] \rangle_b) = \log_b(r)$ satisface la definición de escala invariante.

Solución:

$$\text{Sea } \hat{P}_b([1, r]) = P_b(\langle [1, r] \rangle_b) = \log_b(r)$$

Hagamos $b^a = r$

Y además notemos que $P_b(\langle [1, r] \rangle_b) = \log_b(r) - \log_b(1)$, con lo cual tiene sentido plantear que $P_b(\langle [a, r] \rangle_b) = \log_b(r) - \log_b(a)$, para todo $a < r$

Por tanto,

$$\begin{aligned} P_b(\alpha \langle [1, r] \rangle_b) &= P_b(\alpha \langle [1, r] \rangle_b) \\ &= P_b\left(\alpha \uplus_{n \in \mathbb{Z}} b^n [1, r]\right) \\ &= P_b\left(\uplus_{n \in \mathbb{Z}} b^n [\alpha, \alpha r]\right) \\ &= P_b\left(\uplus_{n \in \mathbb{Z}} b^n \left\{[\alpha, b^k] \uplus [b^k, \alpha r]\right\}\right) \\ &= P_b\left(\left\{\uplus_{n \in \mathbb{Z}} b^n [\alpha, b^k]\right\} \uplus \left\{\uplus_{n \in \mathbb{Z}} b^n [b^k, \alpha r]\right\}\right) \end{aligned}$$

Reescribiremos cada subintervalo de tal forma que se obtenga siempre un intervalo en $[1, b)$.

Así

$$\begin{aligned} P_b(\alpha \langle [1, r] \rangle_b) &= P_b\left(\left\{ \biguplus_{n \in \mathbb{Z}} b^n [\alpha b^\theta, b^\theta] \right\} \uplus \left\{ \biguplus_{n \in \mathbb{Z}} b^n [b^\omega, \alpha r b^\omega] \right\}\right) \\ &= P_b\left(\langle [\alpha b^\theta, b^\theta] \rangle_b\right) + P_b\left(\langle [b^\omega, \alpha r b^\omega] \rangle_b\right) \end{aligned}$$

Donde $[\alpha b^\theta, b^\theta)$ y $[b^\omega, \alpha r b^\omega) \subset [1, b)$

Con lo que finalmente, se tiene.

$$\begin{aligned} P_b(\alpha \langle [1, r] \rangle_b) &= \log_b(b^\theta) - \log_b(\alpha b^\theta) + \log_b(\alpha r b^\omega) - \log_b(b^\omega) \\ &= \log_b\left(\frac{b^\theta}{\alpha b^\theta}\right) + \log_b\left(\frac{\alpha r b^\omega}{b^\omega}\right) \\ &= \log_b\left(\frac{b^\theta}{\alpha b^\theta} * \frac{\alpha r b^\omega}{b^\omega}\right) \\ &= \log_b(r) \end{aligned}$$

Que es justo lo que queríamos verificar.

Ejemplo 2.5.3. Es fácil verificar que $P_*(\langle E \rangle_b) = \begin{cases} 1; & \text{si } 1 \in E \\ 0; & \text{en otro caso} \end{cases}$

no satisface la definición de escala invariante.

Solución:

Tomemos $E = [1, 2)$ con lo que $P_*(\langle E \rangle_b) = 1$

Hagamos $\alpha = 2$

$$\begin{aligned}
P_*(2\langle E \rangle_b) &= P_*(2\langle [1, 2] \rangle_b) \\
&= P_*\left(2\bigoplus_{n \in \mathbb{Z}} b^n [1, 2]\right) \\
&= P_*\left(\bigoplus_{n \in \mathbb{Z}} b^n [2, 4]\right) \\
&= P_*(\langle [2, 4] \rangle_b) = 0
\end{aligned}$$

Con lo que se comprueba que no satisface la definición de escala invariante.

Ejemplo 2.5.4. Es fácil verificar que $Q_b(\langle [1, r] \rangle_b) = \frac{r-1}{b-1}$ no satisface la definición de escala invariante.

Solución:

Es de notar lo siguiente $Q_b(\langle [a, r] \rangle_b) = \frac{r-a}{b-a}$, para toda $a < r$, con lo que la prueba toma el siguiente rumbo y además tomando en cuenta las consideraciones realizadas en la prueba de la propiedad **iv** del Lema **2.2.1** (que siempre es posible reescribir el intervalo tal que éste se encuentre entre 1 y b);

$$\begin{aligned}
Q_b(\alpha \langle [1, r] \rangle_b) &= Q_b\left(\alpha \bigoplus_{n \in \mathbb{Z}} b^n [1, r]\right) \\
&= Q_b\left(\bigoplus_{n \in \mathbb{Z}} b^n [\alpha, \alpha r]\right)
\end{aligned}$$

$$\begin{aligned}
&= Q_b \left(\bigoplus_{n \in \mathbb{Z}} b^n \left\{ [\alpha, b^k] \uplus [b^k, \alpha r] \right\} \right) \\
&= Q_b \left(\bigoplus_{n \in \mathbb{Z}} b^n \left\{ [\alpha b^\theta, b^\theta] \uplus [b^\omega, \alpha r b^\omega] \right\} \right) \\
&= Q_b \left(\left\{ \bigoplus_{n \in \mathbb{Z}} b^n [\alpha b^\theta, b^\theta] \right\} \uplus \left\{ \bigoplus_{n \in \mathbb{Z}} b^n [b^\omega, \alpha r b^\omega] \right\} \right) \\
&= Q_b \left(\left\langle [\alpha b^\theta, b^\theta] \right\rangle_b \right) + Q_b \left(\left\langle [b^\omega, \alpha r b^\omega] \right\rangle_b \right) \\
&= \frac{b^\theta - \alpha b^\theta}{b - \alpha b^\theta} + \frac{\alpha r b^\omega - b^\omega}{b - b^\omega} \\
&= \frac{b^\theta - \alpha b^\theta}{b - \alpha b^\theta} + \frac{\alpha r b^\omega - b^\omega}{b - b^\omega}
\end{aligned}$$

Y se observa claramente que no se obtiene el resultado deseado, por lo tanto $Q_b \left(\left\langle [1, r] \right\rangle_b \right)$ no es de escala invariante.

Teorema 2.5.5. P_b es la única medida de probabilidad de escala invariante en $(\mathbb{R}^+, \mathfrak{M}_b)$

Demostración:

Del resultado del ejemplo 2.5.2 sabemos que P_b es de escala invariante, ahora verifiquemos que efectivamente es la única.

En esta prueba usaremos la siguiente notación $x \in [a, b)$ módulo 1, lo cual significa que $0 < a \leq b \leq 1$, es decir, el máximo valor que puede tomar el límite derecho es 1, o más bien $|x| \leq 1$.

Supongamos para esto que existe una medida de probabilidad P en $(\mathbb{R}^+, \mathfrak{M}_b)$ la cual es de escala invariante para algún $\alpha \in \mathbb{R}$, con lo que también la correspondiente medida \hat{P} lo será, ahora tomemos \bar{P} como la medida de probabilidad en la escala logarítmica $B[0,1)$, es decir,

$$\begin{aligned} P(\langle [1, b^a] \rangle_b) &= \hat{P}([1, b^a]) \\ &= \bar{P}([\log_b(1), \log_b(b^a)]) \quad \mathbf{2.5.5.1} \\ &= \bar{P}([0, a]); a \in [0, 1] \end{aligned}$$

En este punto tiene sentido plantear lo siguiente $P(\langle [1, b^a] \rangle) = \bar{P}([0, a]) = a$, por el hecho de que si nos damos cuenta cuando $a=1$, tendremos que $P(\langle [1, b] \rangle_b) = \bar{P}([0, 1]) = 1$ (por ser $\langle [1, b] \rangle_b$ una sigma álgebra).

Para comprender mejor la demostración, tomemos el siguiente caso particular.

Sea $\alpha = 3$, $b = 10$ y tomemos el intervalo $[1, 5)$, entonces

$$\begin{aligned} P(3\langle [1, 5] \rangle_{10}) &= P(\langle [3, 15] \rangle_b) \\ &= P(\langle \{[3, 10) \uplus [10, 15]\} \rangle_b) \\ &= P(\langle [3 * 10^0, 10^1 * 10^0] \rangle_b) + P(\langle [10^1 * 10^{-1}, 3 * 5 * 10^{-1}] \rangle_b) \\ &= \hat{P}([3 * 10^0, 10 * 10^0]) + \hat{P}([10 * 10^{-1}, 3 * 5 * 10^{-1}]) \end{aligned}$$

$$\begin{aligned}
&= \bar{P}\left(\left[\log_b(3 \cdot 10^0), \log_b(10^1 \cdot 10^0)\right]\right) + \bar{P}\left(\left[\log_b(10^1 \cdot 10^{-1}), \log_b(3 \cdot 5 \cdot 10^{-1})\right]\right) \\
&= \bar{P}\left(\left[\log_b(3) + \log_b(10^0), \log_b(10^1) + \log_b(10^0)\right]\right) \\
&\quad + \bar{P}\left(\left[\log_b(10^1) + \log_b(10^{-1}), \log_b(3) + \log_b(5) + \log_b(10^{-1})\right]\right) \\
&= \bar{P}\left(\left[\log_b(3), 1\right]\right) + \bar{P}\left(\left[1 + (-1), \log_b(3) + \log_b(5) - 1\right]\right) \\
&= \bar{P}\left(\left[\log_b(3), \log_b(3) + \log_b(5)\right]\right)
\end{aligned}$$

Ahora tomemos el mismo intervalo pero con $\alpha = 30$, (obsérvese que $30 = 3 \cdot 10^1$), con lo que tendremos:

$$\begin{aligned}
P\left(30 \langle [1, 5] \rangle_{10}\right) &= P\left(\langle [30, 150] \rangle_b\right) \\
&= P\left(\langle \{[30, 100] \uplus [100, 150]\} \rangle_b\right) \\
&= P\left(\langle [30 \cdot 10^{-1}, 100 \cdot 10^{-1}] \rangle_b\right) + P\left(\langle [100 \cdot 10^{-2}, 150 \cdot 10^{-2}] \rangle_b\right) \\
&= \hat{P}\left([30 \cdot 10^{-1}, 100 \cdot 10^{-1}]\right) + \hat{P}\left([100 \cdot 10^{-2}, 150 \cdot 10^{-2}]\right) \\
&= \bar{P}\left(\left[\log_{10}(30 \cdot 10^{-1}), \log_{10}(100 \cdot 10^{-1})\right]\right) + \bar{P}\left(\left[\log_{10}(100 \cdot 10^{-2}), \log_{10}(150 \cdot 10^{-2})\right]\right) \\
&= \bar{P}\left(\left[\log_{10}(30) + \log_{10}(10^{-1}), \log_{10}(100) + \log_{10}(10^{-1})\right]\right) \\
&\quad + \bar{P}\left(\left[\log_{10}(100) + \log_{10}(10^{-2}), \log_{10}(150) + \log_{10}(10^{-2})\right]\right) \\
&= \bar{P}\left(\left[\log_b(30) + (-1), 2 + (-1)\right]\right) + \bar{P}\left(\left[2 + (-2), \log_b(150) - 2\right]\right) \\
&= \bar{P}\left(\left[\log_b(30) + (-1), \log_b(150) - 1\right]\right)
\end{aligned}$$

Sin embargo, recordemos que $30 = 3 \cdot 10^1$ y que $150 = 3 \cdot 5 \cdot 10^1$,

$$\begin{aligned}
P\left(30 \langle [1, 5] \rangle_{10}\right) &= \bar{P}\left(\left[\log_{10}(3) + \log_{10}(10) + (-1), \log_{10}(3) + \log_{10}(5) + \log_{10}(10) - 1\right]\right) \\
&= \bar{P}\left(\left[\log_{10}(3), \log_{10}(3) + \log_{10}(5)\right]\right)
\end{aligned}$$

Es decir, obtenemos el mismo resultado

Con esto en mente la demostración toma el siguiente rumbo, multiplicando por α a la expresión 2.5.5.1 tendremos lo siguiente.

$$\begin{aligned}
P\left(\alpha \left\langle [1, b^a] \right\rangle_b\right) &= P\left(\left\langle [\alpha, \alpha b^a] \right\rangle_b\right) \\
&= P\left(\left\langle \left\{ [\alpha b^\theta, b^\theta] \uplus [b^{\theta-1}, \alpha b^a b^{\theta-1}] \right\} \right\rangle_b\right) \\
&= P\left(\left\langle [\alpha b^\theta, b^\theta] \right\rangle_b\right) + P\left(\left\langle [b^{\theta-1}, \alpha b^a b^{\theta-1}] \right\rangle_b\right) \\
&= \hat{P}\left([\alpha b^\theta, b^\theta]\right) + \hat{P}\left([b^{\theta-1}, \alpha b^a b^{\theta-1}]\right) \\
&= \bar{P}\left([\log_b(\alpha b^\theta), \log_b(b^\theta)]\right) + \bar{P}\left([\log_b(b^{\theta-1}), \log_b(\alpha b^a b^{\theta-1})]\right) \\
&= \bar{P}\left([\log_b(\alpha) + \log_b(b^\theta), \log_b(b^\theta)]\right) \\
&\quad + \bar{P}\left([\log_b(b^{\theta-1}), \log_b(\alpha) + \log_b(b^a) + \log_b(b^{\theta-1})]\right) \\
&= \bar{P}\left([\log_b(\alpha) + \theta, \theta]\right) \\
&\quad + \bar{P}\left([\theta - 1, \log_b(\alpha) + a + \theta - 1]\right)
\end{aligned}$$

Escribiendo todo en un solo intervalo, obtenemos que

$$P\left(\alpha \left\langle [1, b^a] \right\rangle_b\right) = \bar{P}\left([\log_b(\alpha) + \theta + \theta - 1, \theta + \log_b(\alpha) + a + \theta - 1]\right)$$

Hagamos $\alpha' = \log_b(\alpha) + \theta + \theta - 1$, con base a los resultados de los dos casos particulares vistos anteriormente, es fácil notar que $\alpha' = \log_b(\xi)$; para $\xi \in [1, b)$

y $k \in \mathbb{Z}$, con lo que $0 \leq \alpha' \leq 1$ y $P\left(\alpha \left\langle [1, b^a] \right\rangle_b\right) = \bar{P}\left([\alpha', \alpha' + a]\right)$ módulo 1.

Por otra parte observemos que podemos hacer lo siguiente

$$\begin{aligned}\bar{P}(R_{\alpha'}[0, a)) &= \bar{P}([\alpha', \alpha' + a)) \\ &= \bar{P}([0, a)) = a\end{aligned}$$

Lo cual nos lleva a decir con toda confianza que el hecho de multiplicar por la constante α en P equivale vía logaritmo $\log_b(\bullet)$ a una rotación irracional en el círculo módulo 1 (esto debido a que α' difícilmente será un número racional).

En otras palabras, tomemos un punto $x \in [0, a)$ cuya medida de probabilidad en \bar{P} es a , $\begin{pmatrix} x \\ a \end{pmatrix} \in \mathbb{R}^2$, y mediante el proceso descrito anteriormente obtenemos un nuevo punto $x \in [\alpha, a + \alpha)$ módulo 1 el cual también tiene la misma medida a , $\begin{pmatrix} x + \alpha' \\ a \end{pmatrix} \in \mathbb{R}^2$, lo cual significa que equivale a rotar el punto original $\begin{pmatrix} x \\ a \end{pmatrix}$ en un ángulo de $\alpha' \in [0, 1]$, recuérdese que un punto en una circunferencia es determinado o se determina exclusivamente por su ángulo (puesto que el radio es constante).

Ahora bien la invarianza de P con respecto a la multiplicación por α , equivale a la invarianza de \bar{P} con respecto a la rotación irracional, esto como ya sabemos indica que \bar{P} debe ser la medida de Lebesgue (consúltese el Apéndice), de este modo por **2.4.5.1** se tiene

$$\begin{aligned}
\bar{P}([0, a)) &= a - 0 \\
&= \log_b(b^a) - \log_b(1) \\
&= \hat{P}([1, b^a))
\end{aligned}$$

Por el lema 2.3.4 tendremos que

$$\begin{aligned}
\hat{P}([1, b^a)) &= P(\langle [1, b^a) \rangle_b) \\
&= \log_b(b^a) - \log_b(1)
\end{aligned}$$

Que es justo lo que queríamos demostrar

2.6 Base invariante

Antes de pasar a formalizar la definición de Base Invariante, veamos un pequeño ejemplo ilustrativo para tener mayor claridad de lo que queremos dar a conocer.

Consideremos el conjunto de todos los números reales positivos cuyo primer dígito significativo es menor a 5 y llamemos a dicho conjunto S , es decir, tendremos que S es.

$$S = \dots \uplus [0.01, 0.05) \uplus [0.1, 0.5) \uplus [1, 5) \uplus [10, 50) \uplus \dots$$

Que con ayuda de la primera propiedad del lema 2.3.1, se puede escribir como sigue.

$$\begin{aligned}
S &= \langle [1, 5] \rangle_{10} \\
&= \langle [1, 5] \rangle_{100} \uplus \langle [10, 50] \rangle_{100} \\
&= \langle [1, 5] \rangle_{1000} \uplus \langle [10, 50] \rangle_{1000} \uplus \langle [100, 500] \rangle_{1000}
\end{aligned}$$

Sin embargo, es fácil verificar que si hacemos $a = \log_{10}(5)$, S puede ser representado de cualquiera de las siguientes formas.

$$\begin{aligned}
S &= \langle [1, b^a] \rangle_b && ; b = 10 \\
&= \langle [1, b^{a/2}] \rangle_b \uplus \langle [b^{1/2}, b^{(1+a)/2}] \rangle_b && ; b = 100 \\
&= \langle [1, b^{a/3}] \rangle_b \uplus \langle [b^{1/3}, b^{(1+a)/3}] \rangle_b \uplus \langle [b^{2/3}, b^{(2+a)/3}] \rangle_b && ; b = 1000
\end{aligned}$$

Es de fácil comprobación que al sustituir el valor de $a = \log_{10}(5)$ en cada una de las expresiones anteriores con el correspondiente valor de b se obtiene el mismo resultado que en la expresión anterior de S.

Así, si la medida de probabilidad P en (R^+, \mathfrak{M}_b) es base invariante, las tres medidas de las correspondientes representaciones de S subconjuntos en $[1, b)$ deberían ser las mismas, es decir,

$$\begin{aligned}
\hat{P}[1, b^a] &= \hat{P}[1, b^{a/2}] + \hat{P}[b^{1/2}, b^{(1+a)/2}] \\
&= \hat{P}[1, b^{a/3}] + \hat{P}[b^{1/3}, b^{(1+a)/3}] + \hat{P}[b^{2/3}, b^{(2+a)/3}]
\end{aligned}$$

Y de una manera similar para potencias mayores de base b^n , donde recordemos que \hat{P} es la restricción de P hacia una $B[1, b)$ dada en la expresión del lema 2.3.1.

Ahora estamos en condiciones de dar la siguiente definición clave.

Definición 2.6.1 Una medida de probabilidad P en (R^+, \mathfrak{M}_b) es base invariante si la correspondiente medida de probabilidad \hat{P} en $[1, b)$ satisface la siguiente relación.

$$\hat{P}[1, b^a) = \sum_{k=0}^{n-1} \left(\hat{P} \left[b^{k/n}, b^{(k+a)/n} \right) \right) \text{ para toda } n \in N \text{ y } a \in (0, 1)$$

Comentario: obsérvese que al utilizar la propiedad 1 del lema 2.3.1 se tiene:

$$\begin{aligned} \hat{P}([1, b^a)) &= P(\langle [1, b^a) \rangle_b) \\ &= P\left(\bigoplus_{k=0}^{n-1} \langle b^k [1, b^a) \rangle_{b^n}\right) \\ &= \sum_{k=0}^{n-1} P(\langle b^k [1, b^a) \rangle_{b^n}) \\ &= \sum_{k=0}^{n-1} P(\langle [b^k, b^{(k+a)}) \rangle_{b^n}) \end{aligned}$$

Ahora al dividir por b^n cada uno de los $[b^k, b^{(k+a)})$ se obtienen nuevas sigmas mantisas correspondientes a una única base b con lo que:

$$\begin{aligned} \hat{P}([1, b^a)) &= \sum_{k=0}^{n-1} P(\langle [b^{k/n}, b^{(k+a)/n}) \rangle_b) \\ &= \sum_{k=0}^{n-1} \hat{P}([b^{k/n}, b^{(k+a)/n})) \end{aligned}$$

El significado de Base Invariante es el siguiente: si los primeros dígitos significativos de un conjunto de datos tienen una distribución de probabilidad particular, ésta debería permanecer inalterada independiente de la base numérica en la cual estos datos han sido escritos, por ejemplo; si los dígitos más pequeños son más (menos) comunes que los altos, entonces este mismo comportamiento debería seguirse presentando al cambiar de un sistema decimal a uno hexadecimal.

La definición de base invariante es mucho más general, ya que involucra bases enteras que no corresponda exclusivamente a potencias de b . Por otra parte, la definición es bastante fuerte como para garantizar la invarianza con respecto a cualquier base simple b .

Ejemplo 2.6.2 Sea P_b la medida de probabilidad en $(\mathbb{R}^+, \mathfrak{M}_b)$ definida por

$$P_b(\langle [1, r] \rangle_b) = \log_b(r); \text{ para } r \in [1, b)$$

Es fácil verificar que satisface la definición de base invariante.

Solución:

$$\text{Sea } \hat{P}_b([1, r]) = P_b(\langle [1, r] \rangle_b) = \log_b(r)$$

Hagamos $b^a = r$

Y además notemos que $P_b(\langle [1, r] \rangle_b) = \log_b(r) - \log_b(1)$, con lo cual tiene sentido

plantear que $P_b(\langle [a, r] \rangle_b) = \log_b(r) - \log_b(a)$, para todo $a < r$

Lo que realmente queremos demostrar o a lo que pretendemos llegar es a la siguiente expresión.

$$\begin{aligned} \hat{P}_b([1, r]) &= \sum_{k=0}^{n-1} [\log_b(b^{(k+a)/n}) - \log_b(b^{k/n})] \\ &= \sum_{k=0}^{n-1} \hat{P}_b([b^{k/n}, b^{(k+a)/n}]) \end{aligned}$$

Con esto en mente y utilizando la primera propiedad del lema 2.2.1 la prueba toma el siguiente rumbo.

$$\begin{aligned}
P_b(\langle [1, r] \rangle_b) &= P_b(\langle [b^0, b^a] \rangle_b) \\
&= P_b\left(\bigoplus_{k=0}^{n-1} \langle b^k [b^0, b^a] \rangle_{b^n}\right) \\
&= P_b\left(\bigoplus_{k=0}^{n-1} \langle [b^{k+0}, b^{k+a}] \rangle_{b^n}\right) \\
&= \sum_{k=0}^{n-1} P_b(\langle [b^{k+0}, b^{k+a}] \rangle_{b^n}) \\
&= \sum_{k=0}^{n-1} \left[\log_{b^n}(b^k) - \log_{b^n}(b^{(k+a)}) \right] \\
&= \sum_{k=0}^{n-1} \left[\frac{\log_b(b^k)}{\log_b(b^n)} - \frac{\log_b(b^{(k+a)})}{\log_b(b^n)} \right] \\
&= \sum_{k=0}^{n-1} \left[\frac{\log_b(b^k)}{n} - \frac{\log_b(b^{(k+a)})}{n} \right] \\
&= \sum_{k=0}^{n-1} \left[\frac{1}{n} \log_b(b^k) - \frac{1}{n} \log_b(b^{(k+a)}) \right] \\
&= \sum_{k=0}^{n-1} \left[\log_b(b^{k/n}) - \log_b(b^{(k+a)/n}) \right] \\
&= \sum_{k=0}^{n-1} \hat{P}_b([b^{k/n}, b^{(k+a)/n}]) \\
&= \hat{P}_b([b^0, b^a])
\end{aligned}$$

Que es justo lo que queríamos verificar.

Ejemplo 2.6.3 Sea P_* la medida de probabilidad en $(\mathbb{R}^+, \mathfrak{M}_b)$ definida por

$$P_*(\langle E \rangle_b) = \begin{cases} 1; & \text{si } 1 \in E \\ 0; & \text{en otro caso} \end{cases} \quad \text{Para todo } E \in B[1, b)$$

Entonces la correspondiente medida de probabilidad de Borel \widehat{P}_* en $[1, b)$ es la medida dirac-delta δ_1 , y precisamente P_* es base invariante.

Solución:

Utilicemos un procedimiento bastante similar al aplicado en el ejemplo anterior, es decir

$$\text{Sea } P_*\left(\langle E \rangle_b\right) = \widehat{P}(E)$$

Entonces

$$\begin{aligned} P_*\left(\langle E \rangle_b\right) &= P_*\left(\bigoplus_{k=0}^{n-1} \langle b^k E \rangle_{b^n}\right) \\ &= \sum_{k=0}^{n-1} P_*\left(\langle b^k E \rangle_{b^n}\right) \end{aligned}$$

Y ahora notemos que si $1 \in E$ entonces $1 \in b^0 E$ y que además $1 \notin b^k E$ para cualquier $k \neq 0$, con lo cual, obtendremos que $P_*\left(\langle b^k E \rangle_{b^n}\right) = \begin{cases} 1; & \text{si } k = 0 \\ 0; & \text{si } k \neq 0 \end{cases}$,

con lo que finalmente:

$$\begin{aligned} P_*\left(\langle E \rangle_b\right) &= 1 + 0 + 0 + \dots + 0 \\ &= P_*\left(\langle b^{0/n} E \rangle_b\right) + P_*\left(\langle b^{1/n} E \rangle_b\right) + \dots + P_*\left(\langle b^{(n-1)/n} E \rangle_b\right) \\ &= \sum_{k=0}^{n-1} \widehat{P}_*\left(\langle b^{k/n} E \rangle_b\right) \end{aligned}$$

Que es justo lo que queríamos verificar.

Ejemplo 2.6.4 Sea Q_b la medida de probabilidad en $(\mathbb{R}^+, \mathfrak{M}_b)$ definida por

$$Q_b(\langle [1, r] \rangle_b) = \frac{r-1}{b-1} \text{ Para todo } r \in [1, b), \text{ es la distribución "Uniforme" en}$$

\mathfrak{M}_b . Y además, es de darnos cuenta que la correspondiente medida de probabilidad \hat{Q}_b en $[1, b)$ no satisface la definición de base invariante.

$$\text{Es de notar lo siguiente } Q_b(\langle [a, r] \rangle_b) = \frac{r-a}{b-a}, \text{ para toda } a < r$$

Solución:

Lo que deseamos verificar es lo siguiente:

$$\begin{aligned} \hat{Q}_b([1, r]) &= \sum_{k=0}^{n-1} \hat{Q}_b([b^{k/n}, b^{(k+a)/n}]) \\ &= \sum_{k=0}^{n-1} \frac{b^{(k+a)/n} - b^{k/n}}{b^a - b^0} \\ &= \sum_{k=0}^{n-1} b^{k/n} \left(\frac{b^{a/n} - 1}{b^a - b^0} \right) \end{aligned}$$

Partamos de lo siguiente.

$$\begin{aligned} Q_b(\langle [1, r] \rangle_b) &= Q_b(\langle [b^0, b^a] \rangle_b) \\ &= Q_b\left(\biguplus_{k=0}^{n-1} \langle b^k [b^0, b^a] \rangle_{b^n}\right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=0}^{n-1} Q_b \left(\left\langle \left[b^k, b^{(k+a)} \right] \right\rangle_{b^n} \right) \\
&= \sum_{k=0}^{n-1} \frac{b^{(k+a)} - b^k}{b^n - b^k} \\
&= \sum_{k=0}^{n-1} b^k \frac{b^a - 1}{b^n - b^k}
\end{aligned}$$

Se observa claramente que no llegamos al resultado esperado, con lo cual Q_b no es base invariante.

2.7 Teorema Principal

En esta sección presentamos la demostración de dos teoremas muy importantes concernientes a la Ley de Benford, presentados en 1995 por Theodore Hill, los cuales relacionan la escala invariante con la base invariante, sin embargo, antes necesitamos introducir algunos conceptos nuevos de teoría ergódica.

Definición 2.7.1: Sea $T: \mathfrak{X} \rightarrow \mathfrak{X}$ una transformación medible en un espacio (X, \mathfrak{X}) . Una medida μ se dice que es invariante por T si $\mu(T^{-1}(A)) = \mu(A) \quad \forall A \in \mathfrak{X}$.

Observación: para la demostración del teorema **2.4.5** observamos que la invarianza de escala correspondía a la invarianza ante las rotaciones irracionales en el círculo $(x \rightarrow x + \alpha')$ módulo 1, en este punto nos daremos cuenta que la invarianza de base equivale a la invarianza bajo la multiplicación $(x \rightarrow nx)$ módulo 1, siendo $n \in \mathbb{N}$.

Verifiquémoslo

Sea $T_n(x) = nx$ una transformación medible tal que $|nx| \leq 1$, con lo que es fácil verificar que $T_n^{-1}(x) = \frac{1}{n}x$, ahora por la definición de base invariante sabemos que:

$$\widehat{P}[1, b^a) = \sum_{k=0}^{n-1} \widehat{P}\left[b^{k/n}, b^{(k+a)/n}\right)$$

Al igual que en la demostración del teorema 2.5.5 utilicemos la escala logarítmica en lugar de la lineal, es decir.

$$\begin{aligned} \overline{P}[0, a) &= \sum_{k=0}^{n-1} \overline{P}\left[\log_b(b^{k/n}), \log_b(b^{(k+a)/n})\right) \\ &= \sum_{k=0}^{n-1} \overline{P}\left[\frac{k}{n}, \frac{k+a}{n}\right) \\ &= \sum_{k=0}^{n-1} \overline{P}(T_n^{-1}[k, k+a)) \end{aligned}$$

Se observa que la base invariante corresponde efectivamente a la invarianza de \overline{P} con respecto a la multiplicación (nx) módulo 1, para ser más específicos se obtiene la sumatoria de medidas de probabilidad invariantes por T_n , de esto nace la siguiente definición.

Definición 2.7.2. Si una medida de probabilidad de Borel \overline{P} en $[0,1)$ es invariante bajo la transformación nx para todo $n \in \mathbb{N}$ entonces

$$\overline{P}[0, a) = \sum_{k=0}^{n-1} \overline{P}\left(\left[\frac{k}{n}, \frac{k+a}{n}\right)\right); \text{ para } a \in (0,1)$$

Antes de pasar a la prueba del teorema principal, presentamos la siguiente proposición la cual será la clave para su demostración.

Proposición 2.7.3 Una medida de probabilidad de Borel \bar{P} en $[0,1)$ es invariante bajo la transformación nx para todo $n \in \mathbb{N}$ si y sólo si

$$\bar{P} = q\delta_0 + (1-q)\lambda \text{ para } q \in [0,1] \quad \mathbf{2.7.3.1}$$

Donde: λ es la medida de Lebesgue en $[0,1)$; $\lambda[a,b) = b-a$; $a < b \leq 1$

$$\delta_0 \text{ es la medida Dirac (Borel) en } 0; \delta_0[a,b) = \begin{cases} 1; & 0 \in [a,b) \\ 0; & 0 \notin [a,b) \end{cases}$$

Prueba

i) Si $\bar{P} = q\delta_0 + (1-q)\lambda$ entonces \bar{P} es una transformación invariante ante nx

Es fácil observar la invarianza de \bar{P} , simplemente apliquemos la definición 2.7.2, así;

$$\begin{aligned} \sum_{k=0}^{n-1} \bar{P}\left(\left[\frac{k}{n}, \frac{k+a}{n}\right)\right) &= q \sum_{k=0}^{n-1} \delta_0\left(\left[\frac{k}{n}, \frac{k+a}{n}\right)\right) + (1-q) \sum_{k=0}^{n-1} \lambda\left(\left[\frac{k}{n}, \frac{k+a}{n}\right)\right) \\ &= q \left\{ \delta_0\left(\left[0, \frac{a}{n}\right)\right) + \delta_0\left(\left[\frac{1}{n}, \frac{1+a}{n}\right)\right) + \cdots + \delta_0\left(\left[\frac{n-1}{n}, \frac{n-1+a}{n}\right)\right) \right\} \\ &\quad + (1-q) \sum_{k=0}^{n-1} \lambda\left(\left[\frac{k}{n}, \frac{k+a}{n}\right)\right) \\ &= q \{1+0+\cdots+0\} + (1-q) \sum_{k=0}^{n-1} \left(\frac{k+a}{n} - \frac{k}{n}\right) \\ &= q + (1-q) \sum_{k=0}^{n-1} \frac{a}{n} \\ &= q + (1-q)a \\ &= \bar{P}([0,a)) \end{aligned}$$

Que es justo lo que queríamos demostrar.

ii) Si \bar{P} es una transformación invariante ante nx entonces $\bar{P} = q\delta_0 + (1-q)\lambda$

Sabemos que la medida de probabilidad para una medida de probabilidad de Borel en $B[0,1)$ está determinada únicamente por sus coeficientes de Fourier, es decir;

$$\phi_n = \int_0^1 e^{2\pi i n x} d\bar{P}(x); \quad d\bar{P}(x) = \bar{P}(x) dx$$

Por otra parte notemos que $|e^{2\pi i n x}| \leq 1$, con lo que los coeficientes de Fourier se reducen a:

$$\begin{aligned} \phi_n &= \int_0^1 e^{2\pi i n x} d\bar{P}(x) \\ &\leq \int_0^1 d\bar{P}(x) = 1 \end{aligned}$$

Por ser \bar{P} , la función de probabilidad definida en $[0,1)$, y en la expresión anterior no interviene n , podemos decir que cada ϕ_n está acotada por un $q \in [0,1]$ y con lo que además podemos inferir que aproximadamente todos los ϕ_n son aproximadamente iguales a q , es decir, $\phi_n \approx q$.

Por otra parte, obsérvese que

$$\begin{aligned} \bar{P}(\{0\}) &= \int_0^1 e^{2\pi i 0 x} d\bar{P}(x) \\ &= \phi_0 \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \phi_n = q \end{aligned}$$

Ahora utilicemos el teorema de descomposición de Jordan, el cual nos dice que existirá la medida de probabilidad atómica \bar{P}_A y la medida de probabilidad continua \bar{P}_C en $[0,1)$ tales que $\bar{P} = q\bar{P}_A + (1-q)\bar{P}_C$; para $q \in [0,1]$.

Sin embargo, por ser $\bar{P} = q\bar{P}_A + (1-q)\bar{P}_C$ invariante ante la transformación nx , debe cumplirse lo siguiente

$$\begin{aligned}\bar{P}([0, a)) &= q\bar{P}_A([0, a)) + (1-q)\bar{P}_C([0, a)) \\ &= q\sum_{k=0}^{n-1}\bar{P}_A\left(\left[\frac{k}{n}, \frac{k+a}{n}\right)\right) + (1-q)\sum_{k=0}^{n-1}\bar{P}_C\left(\left[\frac{k}{n}, \frac{k+a}{n}\right)\right) \\ &= \sum_{k=0}^{n-1}\bar{P}\left(\left[\frac{k}{n}, \frac{k+a}{n}\right)\right)\end{aligned}$$

Lo cual implicará obviamente

$$\begin{aligned}\bar{P}_A([0, a)) &= \sum_{k=0}^{n-1}\bar{P}_A\left(\left[\frac{k}{n}, \frac{k+a}{n}\right)\right) \\ \bar{P}_C([0, a)) &= \sum_{k=0}^{n-1}\bar{P}_C\left(\left[\frac{k}{n}, \frac{k+a}{n}\right)\right)\end{aligned}$$

Puesto que \bar{P}_A es atómica y las expresiones anteriores se cumplen para todo $n \in \mathbb{N}$ y toda $a \in (0,1)$, no queda más alternativa que \bar{P}_A sea atómica en 0 , es decir,

$$\begin{aligned}\bar{P}_A(E) &= \begin{cases} 0; & 0 \notin E \\ 1; & 0 \in E \end{cases} \\ &= \delta_0\end{aligned}$$

La explicación de la afirmación anterior es muy simple, supongamos que \bar{P}_A es atómica para algún $\xi \in (0, a)$, entonces nótese que difícilmente $\xi \in \left[\frac{k}{n}, \frac{k+a}{n}\right)$, para todo $n \in \mathbb{N}$ y algún $k \in \mathbb{N}$, en consecuencia se cumple la ecuación **2.7.3.1**

Por el otro lado

$\bar{P}_C([0, a]) = \sum_{k=0}^{n-1} \bar{P}_C\left(\left[\frac{k}{n}, \frac{k+a}{n}\right]\right)$, esto implica lógicamente que \bar{P}_C debe ser una medida de lineal (las funciones lineales son las únicas que se pueden descomponer como sumas), supongamos que $\bar{P}_C([\xi_1, \xi_2]) = x(\xi_2 - \xi_1) + y$ para $x, y \in \mathbb{R}$ (la razón por la cual tomamos la diferencia de los valores extremos del intervalo es sencillamente porque ésta nos estaría proporcionando como una especie de "medida" o "magnitud" del intervalo), pero como

$$\begin{aligned} \bar{P}_C([0, a]) &= \sum_{k=0}^{n-1} \bar{P}_C\left(\left[\frac{k}{n}, \frac{k+a}{n}\right]\right) \\ &= \sum_{k=0}^{n-1} \left(x \left(\frac{k+a}{n} - \frac{k}{n} \right) + y \right) \\ &= \sum_{k=0}^{n-1} \left(x \left(\frac{a}{n} \right) + y \right) \\ &= xa + ny \\ &= \bar{P}_C([0, a]) \\ &= x(a-0) + y \end{aligned}$$

Esto se cumplirá si y sólo si

$$ny = y \Leftrightarrow y = 0$$

$$xa = x(a-0)$$

Como la segunda de las expresiones anteriores es cierta para cualquier valor de x , podemos tomar la medida lineal de probabilidad cuya pendiente es igual a 1, es

$$\begin{aligned} \bar{P}_C([0, a]) &= a - 0 \\ \text{decir,} \quad &= \lambda([0, a]) \end{aligned}$$

Con lo cual se completa la prueba.

También pudimos haber hecho lo siguiente;

$$\begin{aligned}\bar{P}(\{0\}) &= q*1 + (1-q)*0 \\ &= q*\bar{P}_A(\{0\}) + (1-q)*\bar{P}_C(\{0\}) \\ \Leftrightarrow \begin{cases} q = q*\bar{P}_A(\{0\}) \Rightarrow \bar{P}_A(\{0\}) = \delta_0 = 1 \\ (1-q)*0 = (1-q)*\bar{P}_C(\{0\}) \Rightarrow \bar{P}_C(\{0\}) = \lambda(\{0\}) = 0 \end{cases}\end{aligned}$$

Que es justo lo que queríamos demostrar.

Teorema 2.7.4 La medida de Probabilidad P en $(\mathbb{R}^+, \mathfrak{M}_b)$ es base invariante si y sólo si para $q \in [0,1]$

$$P = qP_* + (1-q)P_b$$

Donde P_* y P_b corresponde a los ejemplos 2.5.2 y 2.5.3

Prueba:

i) si $P = qP_* + (1-q)P_b$, entonces P es base invariante.

Por la solución de los ejemplos 2.6.2 y 2.6.3, sabemos que ambas medidas de probabilidad son de base invariante, ahora por ser P una combinación lineal convexa de medidas de base invariante resultará ser también una medida de probabilidad de base invariante.

Luego P es base invariante

ii) si P es base invariante, entonces $P = qP_* + (1-q)P_b$.

Supongamos que P es base invariante y sea \bar{P} la medida de probabilidad correspondiente en la escala logarítmica, la cual como observamos anteriormente es:

$$\bar{P}[0, a) = \sum_{k=0}^{n-1} \bar{P}(T_n^{-1}[k, k+a))$$

Por ser P base invariante resulta que \bar{P} es invariante ante la transformación nx módulo 1, utilizando la proposición 2.6.3, se tiene que

$$\begin{aligned} \bar{P}([a, b')) &= q\delta_0([a, b')) + (1-q)\lambda([a, b')) \\ &= q\delta_0([a, b')) + (1-q)(b'-a) \end{aligned}$$

Sin embargo, notemos por un lado que al utilizar la relación 2.5.5.1,

$$\begin{aligned} \lambda([a, b')) &= b' - a \\ &= \log_b(b^{b'}) - \log_b(b^a) \\ &= \hat{P}([b^a, b^{b'})) \\ &= \hat{P}(\langle [b^a, b^{b'}] \rangle_b) \end{aligned}$$

Mientras que por el otro lado

$$\begin{aligned} \delta_0([a, b')) &= \begin{cases} 1; & 0 \in [a, b') \\ 0; & 0 \notin [a, b') \end{cases} \\ &= \begin{cases} 1; & 1 \in [\log_b(b^a), \log_b(b^{b'})) \\ 0; & 1 \in [\log_b(b^a), \log_b(b^{b'})) \end{cases} \\ &= \hat{P}_*([\log_b(b^a), \log_b(b^{b'}))) \\ &= P_*(\langle [\log_b(b^a), \log_b(b^{b'})) \rangle_b) \end{aligned}$$

Es decir, δ_0 y λ , son respectivamente las medidas de probabilidad de 2.6.2 y 2.6.3 en la escala logarítmica, de lo cual se deduce inmediatamente que $P = qP_* + (1 - q)P_b$, que es justo lo que queríamos demostrar. Con lo cual se completa la prueba.

Teorema 2.7.5 (Escala invariante implica base invariante) Si la medida de Probabilidad P en $(\mathbb{R}^+, \mathfrak{M}_b)$ es escala invariante entonces P es base invariante.

Demostración:

La prueba es inmediata de los teoremas 2.5.5 y 2.7.4

Recuérdese que la inversa no es cierta, pues en los ejemplo 2.5.3 y 2.6.3 observamos claramente que P_* es base invariante pero no escala invariante.

CAPÍTULO III
APLICACIÓN DE LA
LEY DE BENFORD

1.1 Introducción.

En la actualidad, la Ley de Benford se utiliza con frecuencia en la detección de datos “fraudulentos” o “inventados”, muchos investigadores la han utilizado con éxito en detectar algunas irregularidad o anomalías en los datos, algunos ejemplos de tales aplicaciones son los siguientes: identificación de fraude en declaración de renta, irregularidades en datos financieros, fraudes fiscales, procesos de auditorias, y últimamente en la detección de fraude electoral, por mencionar algunos.

En este capítulo nos enfocamos principalmente en los fraudes electorales como campo de aplicación, ya que como bien sabemos las elecciones “fraudulentas” y disputas sobre los resultados de una elección no son nada nuevo, generando en muchos casos un gran descontento y/o violencia en aquellos sectores en los cuales el candidato “perdedor” presenta una apelación de los resultados, donde no se da la credibilidad a la veracidad de los datos.

En esta investigación aplicamos la Ley de Benford como un método estadístico basado en la frecuencia de aparición de los dígitos del recuento de votos que puede ser útil para detectar fraude u otras anomalías, muchos autores han puesto a prueba la Ley en el recuento de votos, obteniendo muy buenos resultados, incluso el Dr. Nigrini, indica que este fenómeno se ajusta claramente a la Ley de Benford.

En la primera sección se presenta brevemente dos pruebas planteadas por el Dr. Nigrini que servirán para determinar que también se ajustan el conjunto de datos en estudio a la Ley de Benford, que son conocidas más bien como pruebas de bondad de ajuste; en la segunda sección se presentan algunos comentarios concerniente a conteo de votos y la Ley de Benford; en la tercera sección se muestra una aplicación práctica de la Ley de Benford en las Elecciones Presidenciales de El Salvador en el año 2004, se describe primeramente la base de datos empleada y

algunas limitantes sobre la misma, luego se proceden a analizar el comportamiento de los dos primeros dígitos significativos; y por último se presentan las conclusiones de la aplicación.

1.2 Prueba de bondad de ajuste de una distribución observada a la Ley de Benford.

La Ley de Benford como ya mencionamos en los capítulos anteriores especifica que en una colección de números los dígitos no deben ocurrir con igual frecuencia, sino que el 1 aparece como dígito inicial más frecuente, seguido del 2, etc. hasta el 9 que es el menos probable, un comportamiento muy similar se presenta para el segundo y demás dígitos sólo que en esta vez la diferencia entre las probabilidades de dichos dígitos son menos apreciables.

Sin embargo, podríamos plantearnos la siguiente interrogante ¿Cómo podríamos detectar o saber si una colección de números se ajusta o no a la Ley de Benford?, en vista que la Ley de Benford no es más que una distribución de probabilidad la respuesta que se esperaría en muchas persona es simplemente graficar la frecuencia de aparición de los dígitos y compararlas con la Ley de Benford y luego observar que tan alejados o cercanos estén ambas distribuciones, sin embargo, realizar tal actividad podría no ser lo conveniente debido a que personas con diferentes puntos de vista podrían llegar a diferentes conclusiones; por otra parte existen estadísticos que nos brindan una serie de test o pruebas de bondad de ajuste capaces de indicarnos si una colección de números se ajusta o no a la Ley de Benford; de hecho aunque no se especifique formalmente que la detección de datos “inventados” es una prueba de hipótesis (en la cual la hipótesis nula es que no existe manipulación), en base a ello se decide si a existido o no “manipulación”.

Estos test simplemente se encargan de examinar que también se ajusta (alejan) los datos observados a la distribución teórica de Benford, el Dr. Mark Nigrini, el cual es uno de los expertos y pioneros en el tema propone las siguientes pruebas: Z, χ^2 (Chi-Cuadrado).

En el caso del estadístico Z , la prueba de bondad de ajuste recoge las desviaciones de los datos observados vs. La Ley de Benford para cada uno de los dígitos. De este modo, permite comprobar si la frecuencia de aparición de cualquier dígito (0,1,2,...9) en una posición dada (primero, segundo, etc.), es mayor de la esperada. El estadístico Z proporcionado por Nigrini está definido por:

$$Z_i = \frac{\left| p_o^i - p_e^i \right| - \frac{1}{2n}}{\sqrt{p_e^i \left(\frac{1 - p_e^i}{n} \right)}} \quad (3.2.1)$$

Donde: p_e^i denota la proporción esperada para el i -ésimo valor del dígito en estudio.

p_o^i denota la proporción observada para el i -ésimo valor del dígito en estudio.

n el número de observaciones.

El término $\frac{1}{2n}$ es una corrección de continuidad que únicamente se utiliza cuando su valor es menor que el primer término del numerador. Las proporciones esperadas y observadas deben obtenerse a partir de la correspondiente función de distribución obedeciendo al tipo de prueba que se realiza (primer dígito, segundo, etc.) presentadas anteriormente en el capítulo primero.

En la actualidad, se considera que la corrección de continuidad es excesivamente conservadora y la mayoría de trabajos desaconsejan que se aplique para muestras grandes $[np_e \text{ y } n(1-p_e) \geq 5]$ (Haviland, 1900). Es por ello que es preferible no realizar esta corrección de continuidad y aplicar la siguiente ecuación.

$$Z_i = \frac{|p_0^i - p_e^i|}{\sqrt{p_e^i \left(\frac{1-p_e^i}{n} \right)}} \quad (3.2.2)$$

Además el Dr. Nigrini propone el estadístico χ^2 (Chi-Cuadrado), para calcular la bondad de ajuste de todos los dígitos respecto a lo esperado por la Ley de Benford. Así, las desviaciones de todos los dígitos se sumarían y el valor de este sumatorio es comparado con el valor crítico $\chi_{(\alpha, k-1 \text{ gl})}^2$. La fórmula para el cálculo del estadístico es la habitual:

$$\chi^2 = \sum_{i=0}^9 \frac{(O_i - E_i)^2}{E_i} \quad (3.2.3)$$

Donde: O_i es la frecuencia observada que el dígito tome el valor i -ésimo.

E_i es la frecuencia esperada que el dígito tome el valor i -ésimo.

Para el caso de la prueba del primer dígito los valores O_0 y E_0 serán iguales a cero.

El Dr. Nigrini propone utilizar una u otra prueba de conformidad en función de si se examina la conformidad respecto a la distribución teórica de un dígito (prueba Z) o del conjunto de ellos (prueba χ^2), el estadístico Z , es un caso particular del

χ^2 para 1 grado de libertad. La primera permitiría detectar sobre apariciones de un dígito concreto mientras que la segunda localizará patrones desajustados. Es de observar que las pruebas no son excluyentes sino complementarias; hallar desajustes en los datos respecto a los valores esperados nos haría sospechar del conjunto del registro, se haya utilizado un test u otro.

3 La Ley de Benford en el recuento de votos y algunas limitantes.

El método estadístico que se describe en esta sección es utilizado para ayudar a descubrir “fraude” en una elección electoral. Este método se distingue de otros ya que no requiere que tengamos covariantes para las cuales podemos asumir razonablemente que los votos estén relacionados por jurisdicciones políticas. Este método está basado en las pruebas de distribución de los dígitos en el recuento de votos reportados, de esta manera todo lo que se necesita es el recuento mismo de los votos. Siendo que el método está basado en poca información, no puede en sí diagnosticar que la presencia de una anomalía ha sido el resultado o consecuencia de “fraude” o “irregularidad”. Así que el método se entiende más bien como un indicador para llevar una investigación más exhaustiva en los lugares donde se ha encontrado la anomalía; por ejemplo, el conteo manual de boletas.

Una pregunta fundamental es ¿por qué debemos esperar que la Ley de Benford sea aplicable para los datos del recuento de votos?, un resultado fundamental es que en general la Ley de Benford no se lleva a cabo para datos que son absolutamente aleatorios ni muy condicionados (Raimi 1976; Hill 1995). Esta propiedad es una base para su uso propuesto en el descubrimiento de fraude financiero. Si alguien usa números tomados directamente de una tabla de números aleatorios para rellenar registros financieros falsificados, entonces los dígitos ocurrirán con igual frecuencia violando el principio de la Ley de Benford. El caso

positivo para usar la Ley de Benford con datos financieros confía en los orígenes supuestamente complicados de datos financieros.

Conjuntos de datos que siguen la Ley de Benford cuando los elementos son el resultado de las variables aleatorias tomadas de diversas fuentes que se han multiplicado, dividido, o elevado a una potencia. Esto ayuda a explicar porque ciertos conjuntos de números en contabilidad parecen a menudo seguir estrechamente una distribución de Benford. Los recuentos numéricos de conteo son a menudo el resultado de un proceso matemático. Un ejemplo simple podría ser una cuenta por cobrar que es el número de artículos vendidos (que provienen de una distribución) multiplicado por el precio del artículo (que viene de otra distribución). (Durtschi et al. 2004, 20-21).

Volviendo al caso del número de votos obtenidos por un candidato se observa que efectivamente se encuentra esa escala intermedia, de hecho para que un candidato se declare como ganador no puede obtener aleatoriamente votos en cada municipio, es decir, no puede ganar una elección teniendo la misma probabilidad de obtener pocos o mucho votos; por otra parte el total de votos obtenidos se encuentra condicionado a tomar el valor mínimo de cero y un máximo al total de la población empadronada (suponiendo que todos voten).

Además como ya es conocido la Ley es aplicable a Censos Poblacionales; de este modo el número de votos obtenidos por un candidato en una elección en realidad representa un "Censo Poblacional" donde simplemente se registra su población (votos obtenidos) en el territorio. Por lo que podemos asegurar que el conteo de votos obtenidos por un cierto candidato en una elección deberá seguir cercanamente el comportamiento de la Ley de Benford ya sea en el primer o segundo dígito.

La clase de complejidad que puede producir recuentos con dígitos que siguen la Ley de Benford se refiere a procesos que son mezclas estadísticas (por ejemplo, Janvresse y de la Rue (2004)), las que significan que porciones aleatorias de los datos vienen de diferentes distribuciones estadísticas. Sin embargo, hay algunas restricciones que se aplican a la dimensión de las mezclas. De esta manera, si deseamos creer que en general debe esperarse que la Ley de Benford describa los dígitos del recuento de votos, necesitamos tener un proceso realista de comportamiento que involucre mezclas entre un número pequeño de distribuciones.

El método de **Prueba del dígito** está basado en la perspectiva de que los segundos dígitos para el recuento de votos deben satisfacer la Ley de Benford (Hill 1995). Aunque algunos han propuesto usar la Ley de Benford de la distribución del segundo dígito para probar votos fraudulentos (el ej., Pericchi y Torres 2004), prominentes monitoreos de elección han disputado fuertemente tales propuestas (Carter Center 2005). Se sugiere que un enfoque cercano sobre el acto de votar sugiere un modelo estadístico que muy a menudo produce recuentos con segundos dígitos que tiene la distribución especificada por la Ley de Benford. Lo que tenemos a menudo en los recuentos de datos de una votación no es precisamente la Ley de Benford sino un proceso que se parece fuertemente a la distribución de la Ley del Benford producido en los segundos dígitos.

Otro problema importante de interés es que “si debe esperarse que la Ley de Benford sea aplicable a todos los dígitos en el recuento de votos informados”. En particular, para los datos a nivel de municipio hay buenas razones para dudar que los primeros dígitos del recuento de votos vayan a satisfacer la Ley de Benford. Brady (2005) desarrolla una versión de este argumento. El punto básico es que a menudo los municipios se diseñan para incluir aproximadamente el mismo número de votantes. Si un candidato tiene aproximadamente el mismo nivel de

apoyo en todos los municipios, lo cual significa que parte de los votos de los candidatos es aproximadamente el mismo en todos los municipios, entonces el recuento de los votos tendrá como primer dígito el mismo en todos los municipios. Asimismo debido a que el número de votos obtenido por un candidato en una Urna o Centro de Votación nunca puede ser mayor al total de la población empadronada limitando de esta manera el valor del primer dígito, por lo que en estos casos se recomienda utilizar el análisis del segundo dígito para determinar la presencia de supuestas anomalías.

Generalmente se acostumbra realizar únicamente las pruebas del primer y/o segundo dígito para la detección de anomalías, debido a que en naciones pequeñas o grandes se contará con muy poca información, es decir, su división política es de muy pocos municipios por lo que el análisis no serán del todo significativo independientemente que dígito se utilice, recordemos que en el caso de los dos primeros dígitos se observan 90 casos posibles por lo que es necesario que se dispongan una amplia cantidad de información, la cual es únicamente disponible para votos a nivel de urna .

Para que la Ley de Benford sea satisfecha por los dígitos del número de votos, claramente depende de la ocurrencia de una distribución fortuita de tamaños de los municipios. De esta manera, resulta más conveniente utilizar los segundos dígitos significativos del recuento de votos, para que por lo menos no haya ningún golpe bajo similar que contradiga los argumentos de la Ley.

3.4.4 Aplicación de la Ley de Benford en las Elecciones Presidenciales de El Salvador en el 2004.

Para la realización de esta aplicación hemos tomando como guía de trabajo, análisis de los resultados electorales a partir de la Ley de Benford desarrollado por Emilio Calderón de La Facultad de Ciencias de la UNAM (entre otros), para realizar un breve análisis de la Ley de Benford en el conteo de votos en las Elecciones Presidenciales de El Salvador 2004, dicha información se encuentra disponible en la Junta de Vigilancia Electoral en el siguiente sitio www.jve.gob.sv, en tal sitio se puede encontrar información referente a elecciones presidenciales de años anteriores incluso referente a las elecciones de diputados y alcaldes del País.

Nuestro análisis se centra en el estudio del patrón de los dos primeros dígitos significativos del resultado del conteo de votos para las elecciones presidenciales tanto a nivel nacional como a nivel de departamento; los candidatos que participaron para las elecciones presidenciales en el 2004 fueron: Héctor Silva, Shafick Handal, Rafael Machuca y Antonio Elías Saca, los cuales pertenecen respectivamente a los siguientes partidos políticos: La coalición CDU-PDC, FMLN, PCN y ARENA. Los resultados aquí mostrados son por partido político y no como candidato, pues en nuestro país se manejan así las elecciones.

3.4.1 Descripción de los Datos.

Decidimos trabajar con la base de datos correspondiente a las Elecciones Presidenciales de 2004, debido a la importancia de la misma y en víspera a que estamos a las puertas de nuevos comicios, además que como todos recordarán el enorme descontento que se generó cuando el Tribunal Supremo Electoral (TSE) dio a conocer los resultados.

Cabe mencionar que tenemos una limitante con respecto al tamaño y calidad de la información disponible, lamentablemente sólo disponemos con información a nivel de municipio (total de votos) nos resultó imposible contar con información a nivel de centro de votación o mejor aún a nivel de urna por lo que los resultados y análisis aquí mostrados son a nivel de departamento y no a nivel de municipio como debería ser, es por eso que los resultados que a continuación se muestran podría parecer un “Insulto” para algunas personas, por lo que más bien nuestro trabajo es un ejemplo ilustrativo de aplicación de la Ley de Benford y no un experimento científico.

Vale la pena mencionar que trabajamos con esos datos suponiendo la veracidad de los mismos, es decir, no podemos verificar de ninguna manera si esos datos son reales. Los resultados que se muestran son votos válidos y no sabemos si la cantidad de votos anulados fue significativamente, ni tampoco en que condiciones o criterios fueron anulados, en resumen no hay forma de garantizar la transparencia de la información, al menos con la información con la que contamos, en la tabla 3.1 se muestra un resumen de la información:

Como puede apreciarse contamos con el consolidado del escrutinio de votos válidos de cada partido político que participó en los comicios electorales de 2004, por departamento. Nosotros estamos interesados en estudiar únicamente el consolidado del escrutinio de votos válidos por cada partido político de los 262 municipios del país.

Tabla 3.1. Resumen del Escrutinio Final de las Elecciones Presidenciales del 2004.

DEPARTAMENTO	NÚMERO DE MUNICIPIOS POR DEPARTAMENTO	VOTOS VÁLIDOS				
		ARENA	FMLN	PCN	CDU-PDC	TOTAL
AHUACHAPÁN	12	66,942	33,142	5,772	4,186	110,042
SANTA ANA	13	123,521	66,912	4,154	8,443	203,030
SONSONATE	16	91,103	54,039	8,143	8,847	162,132
CHALATENANGO	33	44,800	26,359	1,866	1,835	74,860
LA LIBERTAD	22	162,648	86,087	4,780	10,615	264,130
SAN SALVADOR	19	388,684	301,328	7,942	33,599	731,553
CUSCATLÁN	16	56,797	28,769	2,952	2,408	90,926
SAN VICENTE	13	33,083	23,288	2,374	1,409	60,154
CABAÑAS	9	35,140	14,166	2,057	1,256	52,619
LA PAZ	22	66,322	38,046	4,627	3,941	112,936
USulután	23	71,314	47,390	2,597	2,170	123,471
SAN MIGUEL	20	87,262	56,136	4,031	5,320	152,749
MORAZÁN	26	37,060	18,896	2,667	2,682	61,305
LA UNIÓN	18	49,760	17,961	7,819	2,026	77,566
TOTAL	262	1314,436	812,519	61,781	88,737	2277,473
PORCENTAJE DE VOTOS OBTENIDOS POR PARTIDO		58%	36%	3%	4%	100%

Como podemos observar en la última fila de la tabla 1, el partido ARENA ganó las elecciones Presidenciales del 2004, con el 58% del total de votos válidos, su principal rival el FMLN obtuvo un 36% del total de votos válidos, la diferencia entre ambos partidos fue de 22% de los votos a favor de ARENA, lo cual lo adjudica como claro vencedor en las elecciones presidenciales del 2004, de hecho un dato curioso de la elección es que el partido ARENA ganó en 242 municipios, es decir, en el 92.3% del territorio nacional obtuvo la mayor cantidad de votos.

Al observar con un poco más de detalle cada dato presentado en la tabla 1, nos damos cuenta que el partido ARENA dominó la elección en todos los departamentos al obtener por lo menos el 56.1 % de los votos de cada departamento, solamente en los departamentos de San Salvador y San Vicente obtuvo respectivamente el 53% (388,684) y 54.9% (33,083), el departamento de Cabañas fue donde obtuvo el mayor apoyo pues recibió el 66.7% (35,140) del total de votos por departamento; por su parte el partido FMLN obtuvo menos del 37 por ciento del total de votos de cada departamento con excepción de los departamentos de San Salvador y San Vicente con el 41.1% (301,328) y 38.7% (23,288) del total de votos, en el departamento de La Unión fue en el que obtuvo el peor apoyo pues obtuvo a penas el 23.1% (17,961) de los votos; los restantes partidos políticos no obtuvieron más del 5.5% del total de votos en cada uno de los departamentos, a excepción del PCN que obtuvo el 10.08% (7,819) del total de votos registrados en el departamento de La Unión.

3.4.2 Análisis del primer dígito significativo del recuento de votos para las Elecciones Presidenciales del 2004.

Al obtener la frecuencia de aparición del primer dígito significativo del recuento de votos en las Elecciones Presidenciales del 2004, obtenemos los resultados presentados en la tabla 3.2, en ella se muestran la frecuencias observadas y esperadas del primer dígito significativo por cada partido político.

Tabla 3.2. Frecuencias Esperadas y Observadas del Primer Dígito Significativo del recuento de votos por Partido Político y del Total de Votos Válidos

DÍGITO	TEÓRICA	ARENA	FMLN	PCN	CDU-PDC	TOTAL
1	79	91	71	82	70	70
2	46	52	49	45	41	61
3	33	32	33	29	31	38
4	25	19	21	24	30	24
5	21	14	24	19	15	17
6	18	15	17	18	21	12
7	15	21	21	14	20	18
8	13	8	13	15	20	10
9	12	10	13	15	14	12

Nótese que la frecuencia observada por parte de ARENA que mejor se ajustan a las frecuencias esperadas, corresponde a los dígitos $d = 3, 6$ y 9 con un diferencia de 1 y 2 municipios respectivamente, por otra parte los dígitos que peor se ajustan corresponden a $d = 1$ y 5 con una diferencia de 12 y 7 municipios, las diferencias para los restantes dígitos son de no más de 6 municipios; mientras las frecuencias observadas que mejor se ajustan con las esperadas para el FMLN, son la de los dígitos $d = 3, 6, 8$ y 9 con una diferencia de un municipio para los dígitos 6 y 9, para los dígitos 3 y 8 es perfecta, mientras que los dígitos que peor se ajustan son $d = 1, 4$ y 7 con una diferencia de 8, 4 y 6 respectivamente, en los restantes dígitos el ajuste no es del todo mal pues la diferencia es de exactamente 3 municipios (uno más que los que considerábamos un buen ajuste para ARENA); por parte del PCN obtenemos que las frecuencias observadas que mejor se ajustan a las esperadas son las que corresponden a los dígitos $d = 2, 4, 6,$ y 7 con una diferencia de 1 municipio para todos los dígitos a excepción del 6 el cual es perfecta, mientras que para los restantes dígitos el ajuste no es tan malo pues presenta diferencia máxima de 4 municipios para el 3, 3 para el 1 y 2 en el resto (5 y 8), podríamos decir que son los datos que mejor se ajustan a la Ley Benford; mientras que para el CDU-PDC los dígitos que mejor se ajustan son los $d = 3$ y 9 con una diferencia de 2

municipios, el resto tiene diferencias mayores o iguales a 3 pero menores a 10; finalmente el total de votos es el peor ajuste, la diferencia mínima corresponde a 9 (ajuste perfecto) y la máxima para el 2 (15 municipios).

De la tabla 3.2 también podemos observar que el dígito uno es el que peor ajuste tiene para todos los partidos políticos, mientras que el tres es uno de los que ha obtenido menores errores o que cuyas frecuencias observadas se encuentran lo suficientemente cercanas a las esperadas.

En la tabla 3.3 se muestra el estadístico Z (obtenido a partir de la ecuación 3.2.2) para cada dígito, si recordamos este estadístico recoge las desviaciones de los datos observados vs. la Ley de Benford para cada uno de los dígitos.

Tabla 3.3. Estadístico Z Para el Primer Dígito Significativo.

DÍGITO	ARENA	FMLN	PCN	CDU-PDC	TOTAL
1	1.634	1.060	0.422	1.195	1.195
2	0.951	0.465	0.184	0.833	2.411
3	0.137	0.050	0.698	0.324	0.984
4	1.335	0.917	0.290	0.963	0.290
5	1.543	0.745	0.399	1.315	0.857
6	0.628	0.133	0.114	0.855	1.369
7	1.535	1.535	0.316	1.270	0.742
8	1.515	0.113	0.448	1.850	0.954
9	0.588	0.299	0.890	0.595	0.003

Hemos señalado de color rojo aquellos valores mayores a la unidad, nos indican que los dígitos correspondientes por cada partido político, no se ajustan muy bien debido a una mayor desviación entre los datos observados y los esperados por la Ley de Benford. Notemos que el partido ARENA tiene cinco dígitos con valores Z mayores a la unidad y corresponden a los valores 1, 4, 5, 7 y 8, también el valor de 2 es bastante próximo a la unidad; para el FMLN se tienen únicamente dos valores

mayores a la unidad los cuales corresponde a los dígitos 1 y 7; el PCN no presenta ningún valor grande, esto confirma el buen ajuste que se mencionaba y mostraba en la tabla 2; el CDU-PDC presenta cuatro valores mayores a la unidad que corresponden a los dígitos 1, 5, 7 y 8; en cambio en el total de votos válidos se presenta tres valores mayores a la unidad que corresponden a los dígitos 1, 2 y 6, el segundo con un valor mayor a 2.

Por otra parte notemos que para el caso del primer dígito presenta valores mayores a la unidad para todos los partidos políticos a excepción del PCN; el dígito 7 también presenta valores mayores a la unidad para los tres partidos políticos; el dígito 2 no presenta valores mayores a la unidad para ningún partido a excepción del total de votos válidos donde obtiene un valor de 2.41, un comportamiento bastante similar lo presenta el dígito 6; los dígitos 3 y 9 son los que presentan un mejor ajuste. En la tabla 3.4 se muestran las distribuciones teóricas para el primer dígito significativo del escrutinio de votos sólo para recordar.

Tabla 3.4. Distribución del Primer Dígito Significativo (Esperadas)

d	1	2	3	4	5	6	7	8	9
P(d₁)	0.3010	0.1761	0.1249	0.0969	0.0792	0.0669	0.0580	0.0512	0.0458

Con el estadístico Z, podemos ver la conformidad respecto a la distribución teórica de un dígito en particular, hemos visto que el conjunto de dígitos por parte del PCN se ajustan bastante bien a los esperados, para confirmar esto utilizamos el estadístico χ^2 (Chi-Cuadrado) el cual nos permite ver la conformidad de la bondad de ajuste de todos los dígitos respecto a lo esperado por la Ley de Benford.

En la tabla 3.5 se muestran los estadísticos Chi-cuadrado para la distribución del primer dígito en el recuento de votos a nivel de departamento y también a nivel nacional, estas estadísticas pueden compararse con el valor crítico $\chi^2_{(\alpha, k-1 gl)}$.

Entonces se podrá concluir a un nivel de significancia del 5% y (9-1) grados de libertad que los valores observados serán confiables si χ^2 (Chi-Cuadrado) es menor al valor crítico de 15.5.

Tabla 3.5. Estadístico Chi-Cuadrado Para el Primer Dígito.

DEPARTAMENTO	VOTOS ARENA	VOTOS FMLN	VOTOS PCN	VOTOS CDU-PDC	CONSOLIDADO DE VOTOS
Ahuachapán	10.21	6.57	13.58	10.17	13.73
San Salvador	1.33	2.94	6.21	5.32	6.87
Cabañas	15.25	7.67	5.54	34.38	6.97
Chalatenango	4.25	10.27	5.58	14.03	2.72
Cuscatlán	6.96	7.32	7.73	15.66	5.46
La Libertad	12.48	9.76	5.67	8.03	7.06
La Paz	5.97	2.39	8.83	9.54	4.22
La Unión	7.21	3.21	12.59	3.35	3.55
Morazán	4.38	10.70	11.30	6.74	6.68
San Miguel	13.78	3.48	17.84	4.10	12.69
San Vicente	7.82	16.94	4.24	10.82	19.43
Santa Ana	6.11	4.12	6.94	8.92	5.50
Sonsonate	4.18	13.37	4.14	1.89	10.02
Usulután	4.90	10.23	5.44	1.27	5.79
Nivel Nacional	11.52	4.57	1.85	9.88	10.52

Si se realizará un análisis a nivel de país, podríamos concluir que aparentemente no se presenta ninguna anomalía o alteración en el conteo de votos, pues ninguno de los valores Chi-Cuadrado resulta ser mayor al valor crítico de 15.5 para todos los partidos políticos como para el total de votos válidos, de hecho se encuentran en un margen bastante aceptable.

En cambio si nos concentramos a nivel de departamento nos damos cuenta que: el departamento de Cabañas presenta el valor de Chi-Cuadrado igual a 34.38 (más que el doble del valor crítico) para el CDU-PDC y un valor bastante próximo al valor crítico para ARENA; en los departamentos de Cuscatlán, San Miguel y San Vicente presentan un valor Chi-Cuadrado mayor al valor crítico correspondiente a los partidos CDU-PDC, PCN y FMLN con valores 15.66, 17.84 y 16.94 respectivamente, en este último departamento también el total de votos presenta un valor Chi-Cuadrado de 19.43.

Tanto el PCN como el FMLN presentan en un departamento el valor χ^2 calculado mayor al valor crítico, en los restantes departamentos ambos partidos obtienen valores bastante bajos con lo que podríamos pensar que el ajuste en ellos es muy bueno, sin embargo, para el PCN, 10 de sus 14 valores son menores a 10, mientras que para el FMLN son 5 y uno bastante próximo; ARENA no presenta ninguno más, sin embargo, existe un valor bastante próximo al valor crítico; el CDU-PDC presenta dos valores mayores y uno de los cuales es más que el doble del valor crítico; en el total de votos solamente hay un valor Chi-Cuadrado mayor al valor crítico.

En los siguientes gráficos se muestran las curvas de las distribuciones observadas vrs la teórica del primer dígito significativo para cada partido político y del total de votos válidos, esto con el fin de visualizar gráficamente los resultados hasta ahora mostrados.

Gráficos: Distribuciones observadas y teóricas de la Ley de Benford, para el primer dígito significativo, del recuento de votos por Partidos Político.

Figura 3.1.

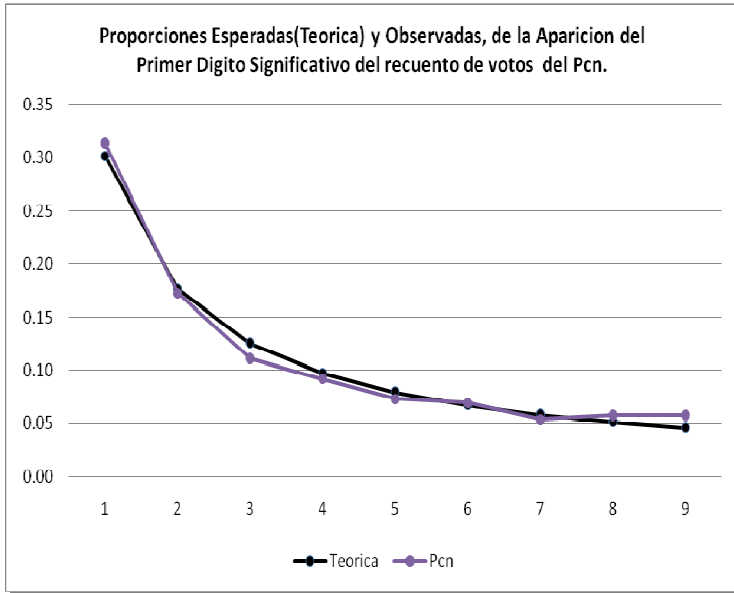


Figura 3.2.

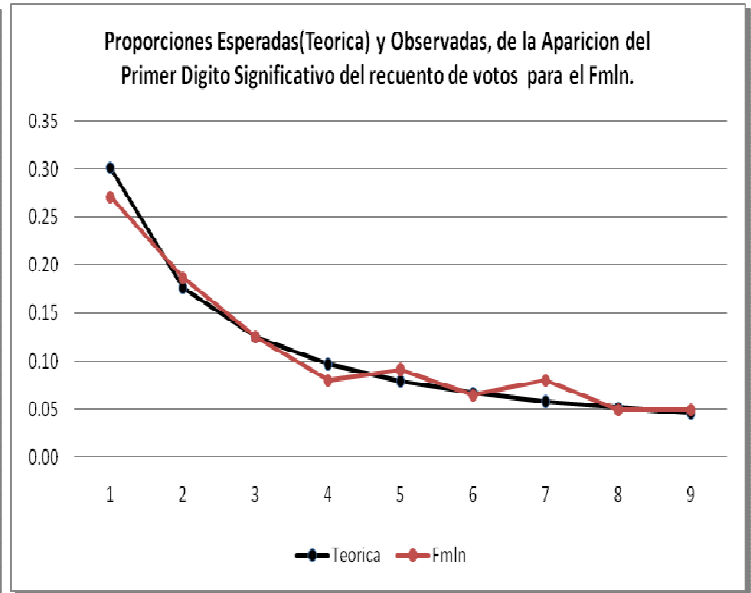


Figura 3.3.

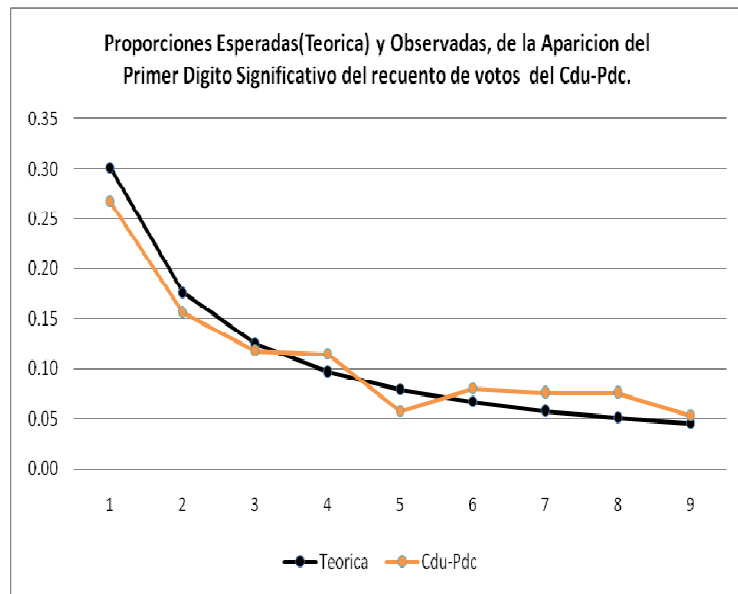


Figura 3.4.

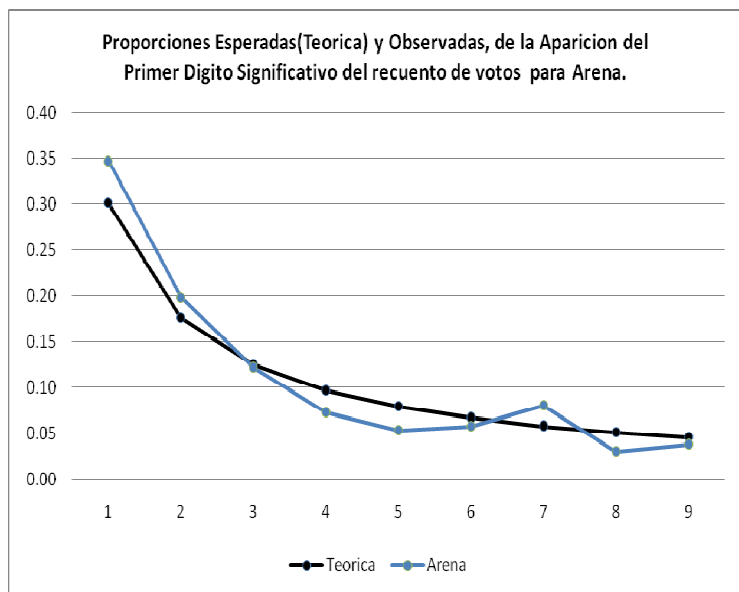
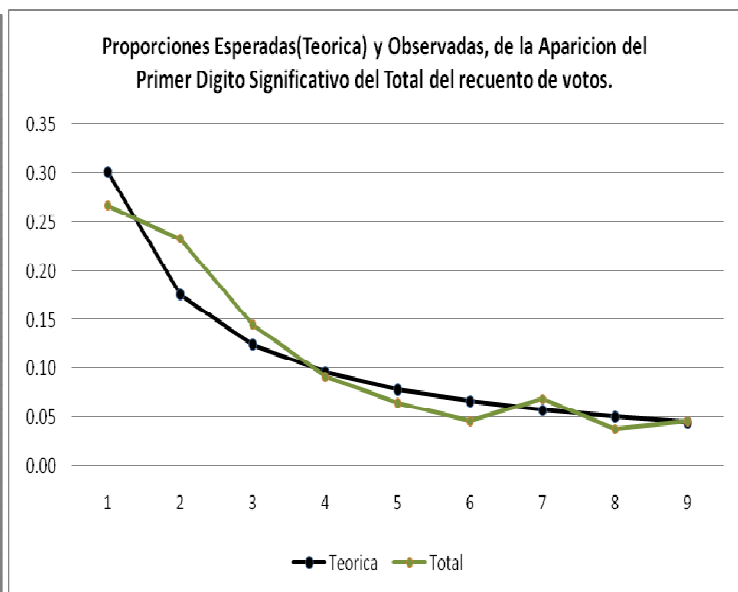


Figura 3.5.



En todas las curvas se observa claramente que la distribución del primer dígito correspondiente para el partido PCN, es la que mejor se ajusta a la Ley de Benford, el ajuste para el FMLN a simple vista no podría considerarse tan malo, mientras que las curvas correspondiente a ARENA, CDU-PDC y total de votos presentan mayor diferencia a la presentada por Benford, estos resultados están bastante próximos con los expuestos anteriormente.

El análisis que hemos realizado hasta el momento es el resultado de estudiar el escrutinio de votos obtenidos por cada partido político, del total de los 262 municipios que hay en el país. Pero como lo mencionamos anteriormente la prueba del primer dígito significativo, no nos indica en sí realmente se ha presentado un fraude o alguna irregularidad en las elecciones para esos departamentos, pero como también ya lo mencionamos nos da un indicio, para llevar un estudio más exhaustivo en esos departamentos, quien sabe tal vez sea necesario comparar el porcentaje de votos nulos con los válidos y los obtenidos en los departamento por cada partido político y así poder observar con más detalle que es lo que realmente está ocurriendo, quien sabe tal vez podría ser

indispensable realizar un nuevo conteo de los votos para estos departamentos, y así confirmar o rechazar la posibilidad de fraude.

3.4.3 Análisis del segundo dígito significativo del recuento de votos para las Elecciones Presidenciales del 2004.

Para confirmar o fortalecer los resultados obtenidos con la prueba del primer dígito significativo realizamos un análisis bastante similar sólo que en esta ocasión para el segundo dígito significativo en el recuento de votos de las elecciones presidenciales del año 2004.

Las proporciones observadas del segundo dígito significativo del recuento de votos de las Elecciones Presidenciales de 2004 de nuestro país, se muestran en la tabla 3.6 para cada partido político, además se muestra la proporción esperadas del segundo dígito significativo, obtenidas a través de la siguiente ecuación.

$$P(d_2) = \sum_{d_1=1}^9 \log_{10} \left(1 + (10d_1 + d_2)^{-1} \right); \quad d = 0,1,2,\dots,9 \quad (3.4.3.1)$$

Tabla 3.6. Proporciones Esperadas y Observadas del Segundo Dígito Significativo del recuento de votos por Partido Político y del Total de Votos Válidos

DÍGITOS	TEÓRICA	ARENA	FMLN	PCN	CDU-PDC	TOTAL
0	0.120	0.080	0.160	0.141	0.103	0.141
1	0.114	0.099	0.088	0.107	0.095	0.095
2	0.109	0.107	0.130	0.084	0.122	0.076
3	0.104	0.103	0.092	0.122	0.107	0.134
4	0.100	0.103	0.092	0.115	0.111	0.107
5	0.097	0.122	0.122	0.103	0.084	0.115
6	0.093	0.115	0.073	0.076	0.073	0.088
7	0.090	0.092	0.080	0.057	0.103	0.092
8	0.088	0.103	0.092	0.069	0.050	0.092
9	0.085	0.076	0.073	0.073	0.080	0.061

Nótese que las proporciones observadas por parte de ARENA que mejor se ajustan a las proporciones esperadas, corresponde a los dígitos $d = 2, 3$ y 7 con una diferencia de $0.002, 0.001$ y 0.002 , las cuales equivalen a $1, 0$ y 0 municipios respectivamente, por otra parte los dígitos que peor se ajustan corresponden a $d = 0$ y 5 con una diferencia de 0.04 y 0.025 , los cuales también equivalen respectivamente a 10 y 7 municipios, las diferencias para los restantes dígitos no resulta ser mayor a 0.022 la cual también equivale a 6 municipios; mientras las proporciones observadas que mejor se ajustan con las esperadas para el FMLN, son la correspondientes a los dígitos $d = 4$ y 8 con diferencias de 0.008 y 0.004 , las cuales equivalen a 2 y 1 municipio, mientras que los dígitos que peor se ajustan son $d = 0, 1$ y 5 con diferencias de $0.04, 0.026$ y 0.025 que equivalen a $11, 7$ y 7 respectivamente, en los restantes dígitos el ajuste no es del todo mal pues la diferencia máxima es de 0.021 y la mínima de 0.01 que equivale a 5 y 3 municipios respectivamente; por parte del PCN obtenemos que las proporciones observadas que mejor se ajustan a las esperadas son las que corresponden a los dígitos $d = 1$ y 5 con diferencias de 0.007 y 0.006 respectivamente que equivale a 2 municipios, para los restantes dígitos observemos que ahora el ajuste no es tan bueno como el que se presentaba para el primer dígito, pues la diferencia máxima la cual corresponde al valor de 7 es de 0.033 que corresponde a 9 municipios, mientras que la mínima es el dígito 9 con un valor de 0.012 (equivale a 3 municipios); mientras que para la coalición CDU-PDC los dígitos que mejor se ajustan son los $d = 3$ y 9 (que también fueron en el primer dígito) con una diferencia de 0.003 y 0.005 que equivalen a 1 municipio, en el resto de dígitos se tienen diferencias mayores o iguales a 0.013 (dígitos 5 y 7) pero menores a 0.038 (dígito 8), que equivalen respectivamente a 2 y 10 municipios; finalmente en el total de votos válidos se observa al igual que en el primer dígito que el ajuste a Benford no es tan bueno, la diferencia mínima corresponde a 7 (ajuste perfecto) y la máxima para el 2 (9 municipios).

De la tabla 6 también podemos observar que el dígito cero es el que peor ajuste tiene en todos los partidos políticos, para los restantes valores las diferencias han tenido un valor bastante significativo en algunos partidos mientras que en otros no.

En la tabla 3.7 se muestra el estadístico Z (obtenido a partir de la ecuación 3.2.2) para cada dígito, si recordamos este estadístico recoge las desviaciones de los datos observados vrs. La Ley de Benford para cada uno de los dígitos.

Tabla 3.7. Estadístico Z Para el Segundo Dígito Significativo.

DÍGITOS	ARENA	FMLN	PCN	CDU-PCN	TOTAL
0	1.971	2.026	1.074	0.829	1.074
1	0.747	1.330	0.358	0.941	0.941
2	0.101	1.089	1.292	0.692	1.689
3	0.068	0.674	0.943	0.135	1.549
4	0.148	0.469	0.765	0.559	0.354
5	1.395	1.395	0.349	0.696	0.976
6	1.175	1.160	0.948	1.160	0.311
7	0.071	0.576	1.869	0.717	0.071
8	0.887	0.231	1.080	2.173	0.231
9	0.503	0.724	0.724	0.281	1.389

Hemos señalado nuevamente de color rojo aquellos valores mayores a la unidad, nos indican que los dígitos correspondientes para cada partido político, no se ajustan muy bien debido a una mayor desviación entre los datos observados y los esperados por la Ley de Benford. Notemos que el partido ARENA tiene tres dígitos con valores Z mayores a la unidad (2 menos que para el primer dígito) y corresponden a los valores 0, 5 y 6, y además no se encuentran valores próximos a la unidad; el FMLN ahora tiene cinco valores mayores a la unidad (3 más que en el primer dígito) los cuales corresponde a los dígitos 0, 1, 2, 5 y 6, de hecho uno de esos valores es mayor a 2 (para 0); el PCN que en el primer dígito no presentaba valores mayores a la unidad ahora tiene cuatro y corresponden a los dígitos 0, 2, 7

y 8, esto confirma que ahora el ajuste no es del todo bueno, además de tener dos valores bastantes próximos a la unidad; el CDU-PDC presenta dos valores mayores a la unidad (uno de ellos mayor a 2) que corresponden a los dígitos 6 y 8, y uno bastante próximo a la unidad; en cambio en el total de votos válidos se presenta cuatro valores mayores a la unidad que corresponden a los dígitos 0, 2, 3 y 9, además de dos valores bastante cercano a la unidad.

Por otra parte, notemos que el dígito cero presenta valores mayores a la unidad en todos los partidos políticos a excepción del CDU-PDC, de hecho para el FMLN presenta un valor mayor a 2; el dígito 6 también presenta valores mayores a la unidad para todos los partidos políticos a excepción del PCN y también para el total de votos válidos; para cada uno de los dígitos 2, 5 y 8 presentan exactamente valores mayores a la unidad para dos partidos políticos, en ellos dos corresponden al FMLN y al PCN y uno para ARENA y el CDU-PDC que tiene un valor igual a 2.17, podríamos decir que el ajuste en general para todos los dígitos no es bueno porque presenta valores relativamente grandes, en algunos casos bastante próximos a la unidad, para ser más exactos 18 de los 50 valores mostrados en la tabla son mayores a la unidad, y dos de ellos mayores a dos.

En la tabla 3.8 se muestran los estadísticos Chi-cuadrado para la distribución del segundo dígito, en el recuento de votos a nivel de departamento y también a nivel nacional, estas estadísticas pueden compararse con el valor crítico $\chi^2_{(\alpha, k-1 gl)}$. Entonces se podrá concluir a un nivel de significancia del 5% y (10-1) grados de libertad que los valores observados serán confiables si χ^2 (Chi-Cuadrado) es menor al valor crítico de 16.92.

Tabla 3.8. Estadístico Chi-Cuadrado Para el Segundo Dígito Significativo.

DEPARTAMENTO	VOTOS ARENA	VOTOS FMLN	VOTOS PCN	VOTOS CDU-PDC	CONSOLIDADO DE VOTOS
Ahuachapán	6.03	9.14	10.89	4.92	7.35
San Salvador	7.87	12.78	6.15	11.01	8.97
Cabañas	8.54	11.11	14.45	3.77	3.41
Chalatenango	10.49	6.27	12.37	12.23	17.88
Cuscatlán	8.10	4.13	18.44	14.07	7.00
La Libertad	7.69	12.31	4.43	1.70	7.43
La Paz	24.15	1.54	13.38	4.64	10.69
La Unión	9.78	10.37	6.38	2.93	2.39
Morazán	9.79	2.99	11.16	3.77	17.00
San Miguel	8.68	16.38	5.79	20.11	9.15
San Vicente	11.82	5.38	8.86	5.99	13.05
Santa Ana	6.83	15.00	4.40	10.70	10.96
Sonsonate	11.32	5.22	5.42	5.33	7.00
Usulután	23.27	4.20	7.83	8.09	5.86
Nivel Nacional	7.91	10.65	9.59	8.62	9.37

Si se realizará un análisis a nivel de país, podríamos concluir que aparentemente no se presenta ninguna anomalía o alteración en el conteo de votos, pues ninguno de los valores Chi-Cuadrado resulta ser mayor al valor crítico de 16.9 para todos los partidos políticos como para el total de votos válidos, de hecho se encuentran en un margen bastante aceptable.

En cambio si nos concentramos a nivel de departamento nos damos cuenta que: el departamento de Chalatenango presenta el valor de Chi-Cuadrado igual a 17.88 para el total de votos válidos y ningún valor lo suficientemente cercano al valor crítico; en los departamentos de Cuscatlán, La Paz, Morazán, San Miguel y Usulután presentan valores Chi-Cuadrado mayor al valor crítico que corresponden respectivamente a los partidos PCN, ARENA, del total de votos válidos, FMLN y

PDC (para el departamento de San Miguel) y ARENA con valores 18.44, 24.15, 17.0, 16.38, 20.11 y 23.27 respectivamente.

Tanto el PCN, FMLN y el CDU-PDC presentan exactamente en un departamento el valor χ^2 calculado mayor al valor crítico, en los restantes departamentos esos partidos obtienen valores bastante bajos con lo que podríamos pensar que el ajuste en ellos es muy bueno, sin embargo, para el PCN y FMLN 6 de sus 14 valores son menores a 10, mientras que para el PDC-CDU son 5; ARENA presenta dos valores χ^2 mayores al valor crítico en los departamentos de La Paz y Usulután; en el total de votos se encuentran dos valores χ^2 mayores al valor crítico (en los departamentos de Chalatenango y Morazán).

En los siguientes gráficos se muestra las curvas de las distribuciones observadas vrs. la teórica para el segundo dígito significativo para cada partido político y del total de votos válidos, esto con el fin de visualizar gráficamente los resultados hasta ahora mostrados.

Figura 3.6

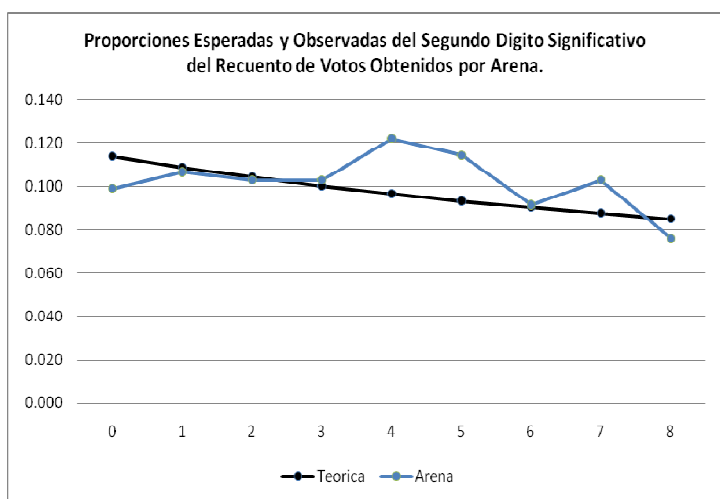


Figura 3.7

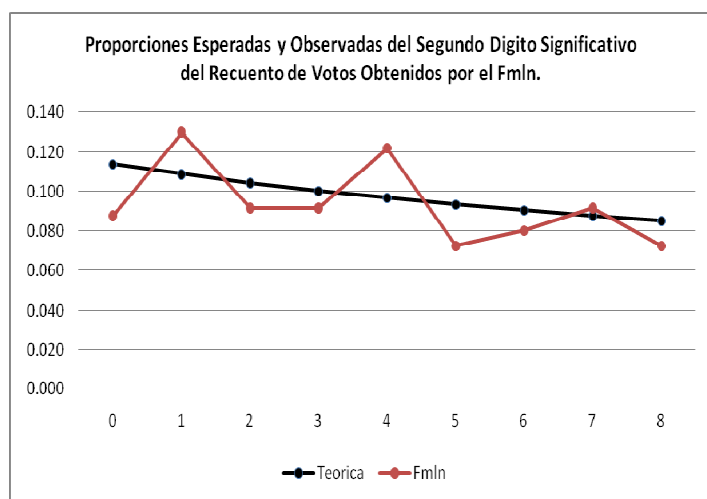


Figura 3.8

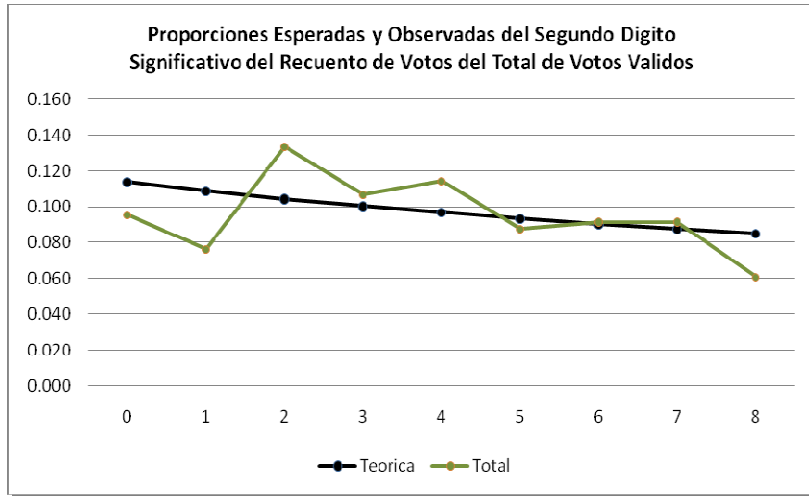


Figura 3.9

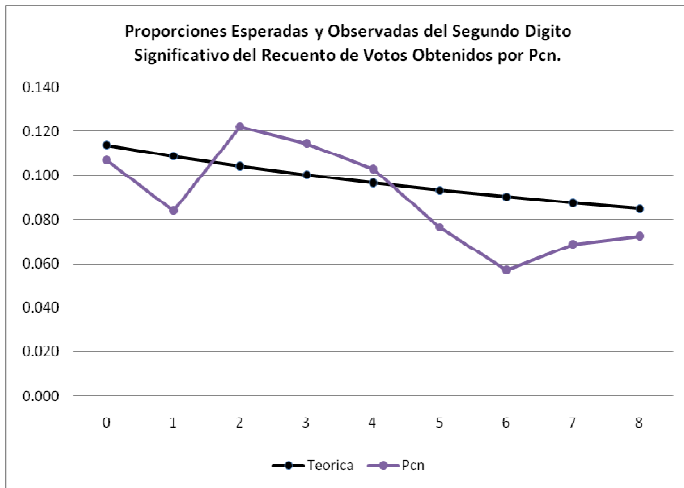
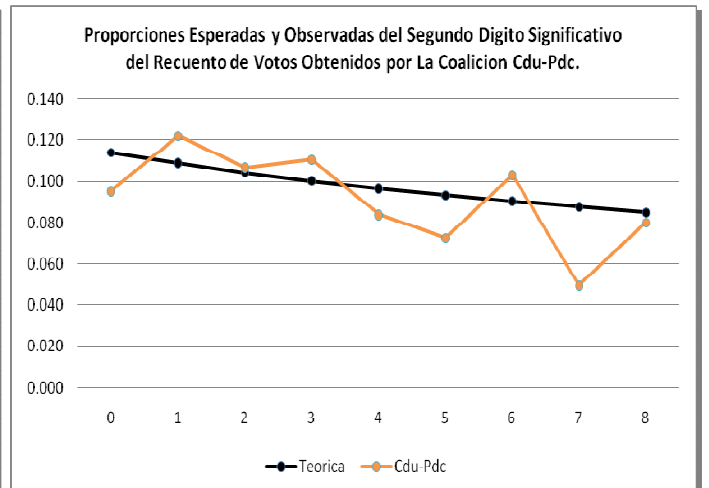


Figura 3.10



En todas las curvas se observa claramente que la distribución del segundo dígito observado para el conteo de votos difiere con las esperadas según la Ley de Benford en cada uno de los partidos políticos que participaron en los comicios así, pero mantienen la tendencia, con base a dichos gráficos señalamos que no es conveniente realizar una investigación más exhaustiva con el fin de determinar si realmente existió o no manipulación en el conteo de votos de las Elecciones Presidenciales del año 2004.

Los resultados que se han obtenido por la prueba del primer y segundo dígito significativo sugieren unos leves problemas en el recuento de votos de las elecciones presidenciales del 2004 en algunos departamentos del país, principalmente en San Miguel, sin embargo, se necesita un análisis más fino para obtener conclusiones más claras, pero la impresión general es que una investigación más intensiva de los resultados de la elección no es necesaria.

3.4.4 Conclusiones de la aplicación.

- Con ayuda del estadístico Z hemos encontrados una serie de datos curiosos que llaman la atención, pues no presentan un ajuste adecuado a la Ley de Benford en el primer dígito significativo en el recuento de votos de las Elecciones Presidenciales de 2004, especialmente el partido ARENA es el que más desajustes presenta, pues de los 14 anomalías encontradas 5 corresponden a dicho partido; mientras que para el FMLN sólo se encontraron dos anomalías; la coalición CDU-PDC es el segundo partido con más anomalías encontradas pues se le encontraron 4; el PCN por su parte no presenta ninguna.
- Con ayuda del estadístico χ^2 hemos encontrados una serie de datos curiosos que llaman la atención, pues no presentan un ajuste adecuado en el primer dígito significativo en el recuento de votos de las Elecciones Presidenciales de 2004, especialmente el partido ARENA, FMLN y PCN presenta únicamente una anomalía, la coalición CDU-PDC mostró tener dos anomalías en el patrón de los dígitos, sólo que en esta ocasión se tiene 6 anomalías en total y no 14 como era el caso de la prueba Z .
- Con ayuda del estadístico Z hemos encontrados una serie de datos que presentan desajustes a la Ley de Benford, pues no presentan un ajuste

adecuado en el segundo dígito significativo en el recuento de votos de las Elecciones Presidenciales de 2004, especialmente el partido FMLN es el que más presenta, pues de las 18 anomalías encontradas 5 corresponden a dicho partido; mientras que para el PCN se encontraron 4 desajustes; ARENA es el tercer partido con más anomalías encontradas pues se le encontraron 3.

- A través del estadístico χ^2 hemos encontrado muy pocos datos curiosos que llaman la atención, es decir, que no presentan un ajuste adecuado en el segundo dígito significativo en el recuento de votos de las Elecciones Presidenciales de 2004, especialmente ARENA es el partido con más anomalías encontradas pues tiene dos; el resto de partidos tiene solamente una anomalía.
- En general no existe suficientes indicios para dudar de la veracidad de los resultados de las Elecciones Presidenciales de 2004, esto debido a que como ya mencionamos anteriormente el análisis del patrón de dígitos solamente es un indicador, el cual permite llevar un estudio más exhaustivo para verificar la presencia o ausencia de fraude; además de existir un buen margen de ganancia por parte del partido ganador.

APÉNDICE

Teoría de la Medida y Teoría Ergódica.

A.1. Colecciones de subconjuntos.

Definición A.1.1: Dado el espacio total Ω una clase $\mathfrak{S} \subset \wp(\Omega)$ tiene estructura de σ -Álgebra si y sólo si.

$$a) \quad \Omega \in \mathfrak{S}$$

$$b) \quad \forall A \in \mathfrak{S} \text{ es } A^c \in \mathfrak{S}$$

$$c) \quad \forall \text{ sucesión } \{A_n\}_{n \in \mathbb{N}} \subset \mathfrak{S} \text{ es } \bigcup_{n=1}^{\infty} A_n \in \mathfrak{S}$$

Cuando $\Omega = \mathbb{R}$ y $\mathfrak{S} = (a, b]$ con a y $b \in \mathbb{R}$ se denomina **σ -álgebra de Borel** y se representa por $B(\mathbb{R})$, ésta σ -álgebra no sólo está engendrada por \mathfrak{S} sino que está engendrada por cualquier de las siguientes clases de intervalos

$$\mathfrak{S}_1 = (a, b), \quad \mathfrak{S}_2 = (-\infty, b], \quad \mathfrak{S}_3 = [a, \infty), \quad \mathfrak{S}_4 = [a, b], \quad \mathfrak{S}_5 = (a, \infty), \quad \mathfrak{S}_6 = [a, b), \quad \mathfrak{S}_7 = (-\infty, b)$$

La **σ -álgebra de Borel** contiene a todos los **intervalos abiertos, semiabiertos o cerrados**, y es la **menor σ -Álgebra** que los contiene.

A.2 Medida de Lebesgue en \mathbb{R}

Sea \mathcal{V} el álgebra de todas las uniones finitas de intervalos semiabiertos de \mathbb{R} , es decir, todos los conjuntos de la forma

$$A = \bigcup_{j=1}^k [a_j, b_j)$$

Definimos una función $\mu : \mathcal{V} \rightarrow \mathbb{R}$ mediante

$$A = \sum_{j=1}^k (b_j - a_j)$$

Definición A.2.1: Dado un conjunto $A \subset \mathbb{R}$ y $x \in \mathbb{R}$, definimos el conjunto trasladado $A(x) = \{x + y : y \in A\}$

Lema A.2.2: *Invariancia bajo traslación.* $\mu(A) = \mu(A(x))$

B.1 Dinámica que preserva medida.

Definición B.1.1: sea $f : I \rightarrow J$, la función $f(x)$ es un homomorfismo si $f(x)$ es uno a uno, sobreyectiva y continua, y $f^{-1}(x)$ es también continua.

Sea (X, \mathfrak{F}) un espacio medible y sea $T : X \rightarrow X$ una transformación medible, es decir $T^{-1}(A) \in \mathfrak{F} \forall A \in \mathfrak{F}$

Definición B.1.2: Sea T medible en un espacio (X, \mathfrak{F}) . Una medida μ se dice que es invariante por T , o que T preserva μ , se dice también que T es un automorfismo del espacio de medida (X, \mathfrak{F}, μ) , si

$$\mu(T^{-1}(A)) = \mu(A) \quad \forall A \in \mathfrak{F}$$

Teorema B.1.3: Si X es un espacio métrico compacto, \mathfrak{S} es la sigma-álgebra de Borel y T es CONTINUA, entonces existe (frecuentemente infinitas) medida de probabilidad invariante.

La medida de Lebesgue m en \mathbb{R} induce una medida m_{\sim} en S^1 dada por $m_{\sim}(A) = m(\Pi^{-1}(A) \cap [0,1])$. Esta medida m_{\sim} se llama medida de Lebesgue en el círculo. Como m en \mathbb{R} es invariante por traslaciones, m_{\sim} es invariante por las rotaciones.

Definición B.1.4: Ergodicidad

Sea (X, \mathfrak{S}, μ) un espacio de medida de probabilidad y $T: X \rightarrow X$ medible que preserva μ . Se dice que T es ergódica respecto a la medida μ , o que μ es una medida ergódica para T , si todo conjunto medible A que sea invariante por T (es decir $T^{-1}(A) = A$) tiene o bien medida nula o bien medida 1.

Teorema B.1.5: La rotación irracional del círculo es ergódica respecto a la medida de Lebesgue.

ANEXO

Conteo de votos a nivel de municipio Elecciones Presidenciales 2004

DEPARTAMENTO	MUNICIPIO	ARENA	CDU.PDC	FMLN	PCN	TOTAL VALIDOS
AHUACHAPÁN	AHUACHAPAN	26,139	1,228	11,286	3,017	41,670
AHUACHAPÁN	ATIQUIZAYA	7,846	312	4,421	363	12,942
AHUACHAPÁN	SAN FRANCISCO MENENDEZ	7,568	666	3,239	463	11,936
AHUACHAPÁN	TACUBA	4,894	249	3,071	628	8,842
AHUACHAPÁN	CONCEPCION DE ATACO	3,494	77	1,784	73	5,428
AHUACHAPÁN	JUJUTLA	3,196	708	2,042	213	6,159
AHUACHAPÁN	GUAYMANGO	2,734	498	2,423	216	5,871
AHUACHAPÁN	APANECA	2,630	87	916	135	3,768
AHUACHAPÁN	SAN PEDRO PUXTLA	1,908	116	1,182	75	3,281
AHUACHAPÁN	SAN LORENZO	1,872	63	928	206	3,069
AHUACHAPÁN	TURIN	2,295	93	1,055	71	3,514
AHUACHAPÁN	EL REFUGIO	2,366	89	795	312	3,562
SANTA ANA	SANTA ANA	57,733	4,268	38,441	1,665	102,107
SANTA ANA	CHALCHUAPA	18,030	927	10,176	621	29,754
SANTA ANA	METAPAN	13,861	1,062	6,520	341	21,784
SANTA ANA	COATEPEQUE	7,536	376	2,712	335	10,959
SANTA ANA	EL CONGO	6,800	442	2,287	143	9,672
SANTA ANA	TEXISTEPEQUE	4,915	665	1,568	240	7,388
SANTA ANA	CAND. DE LA FRONTERA	6,292	334	1,795	408	8,829
SANTA ANA	SAN SEBASTIAN SALIT.	2,385	109	1,228	63	3,785
SANTA ANA	SANTA ROSA GUACH.	1,332	49	185	65	1,631
SANTA ANA	SANTIAGO DE LA FRONTERA	1,472	64	476	72	2,084
SANTA ANA	EL PORVENIR	1,647	84	879	58	2,668
SANTA ANA	MASAHUAT	598	34	316	114	1,062
SANTA ANA	SAN ANTONIO PAJONAL	920	29	329	29	1,307
SONSONATE	SONSONATE	17,904	1,672	12,004	852	32,432
SONSONATE	IZALCO	16,244	865	6,709	1,108	24,926
SONSONATE	ACAJUTLA	8,236	2,249	7,814	401	18,700
SONSONATE	ARMENIA	9,859	446	3,199	335	13,839
SONSONATE	NAHUIZALCO	7,405	422	7,260	2,327	17,414
SONSONATE	JUAYUA	6,510	291	3,022	643	10,466
SONSONATE	SAN JULIAN	3,712	258	1,854	546	6,370
SONSONATE	SONZACATE	5,421	589	2,586	320	8,916
SONSONATE	SAN ANTONIO DEL MONTE	4,485	655	3,326	204	8,670
SONSONATE	NAHUILINGO	2,186	272	1,153	90	3,701
SONSONATE	CUISNAHUAT	1,778	109	1,177	278	3,342

SONSONATE	SANTA CATARINA. MASAHUAT	1,519	312	871	457	3,159
SONSONATE	CALUCO	1,242	144	1,083	286	2,755
SONSONATE	SANTA ISABEL ISHUATAN	1,727	97	649	138	2,611
SONSONATE	SALCOATITAN	1,556	141	630	87	2,414
SONSONATE	SANTO DOMINGO DE GUZMAN	1,319	325	702	71	2,417
CHALATENANGO	CHALATENANGO	8,148	323	5,092	137	13,700
CHALATENANGO	NUEVA CONCEPCION	6,729	327	3,412	253	10,721
CHALATENANGO	LA PALMA	2,795	243	1,246	150	4,434
CHALATENANGO	TEJUTLA	3,112	128	1,652	192	5,084
CHALATENANGO	LA REINA	2,099	83	917	123	3,222
CHALATENANGO	ARCATAO	355	7	751	4	1,117
CHALATENANGO	SAN IGNACIO	1,768	50	1,153	100	3,071
CHALATENANGO	DULCE NOMBRE DE MARIA	1,306	47	595	74	2,022
CHALATENANGO	CITALA	1,097	24	502	191	1,814
CHALATENANGO	AGUA CALIENTE	2,086	151	623	42	2,902
CHALATENANGO	CONCEPCION QUEZALTEPEQUE	1,713	52	781	91	2,637
CHALATENANGO	NUEVA TRINIDAD	152	9	586	3	750
CHALATENANGO	LAS VUELTAS	78	7	595	2	682
CHALATENANGO	COMALAPA	958	37	636	29	1,660
CHALATENANGO	SAN RAFAEL	1,331	44	536	43	1,954
CHALATENANGO	LAS FLORES O SAN J.	19	7	673	1	700
CHALATENANGO	OJO DE AGUA	854	7	406	13	1,280
CHALATENANGO	NOMBRE DE JESUS	794	10	815	10	1,629
CHALATENANGO	POTONICO	399	4	299	6	708
CHALATENANGO	SAN FRANCISCO MORAZAN	380	22	557	22	981
CHALATENANGO	SANTA RITA	1,288	47	461	62	1,858
CHALATENANGO	LA LAGUNA	1,179	13	449	32	1,673
CHALATENANGO	SAN ISIDRO LABRADOR	28	7	177	3	215
CHALATENANGO	SAN ANTONIO DE LA CRUZ	403	26	204	11	644
CHALATENANGO	EL PARAISO	2,592	83	990	213	3,878
CHALATENANGO	SAN MIGUEL DE MERCEDES	595	41	358	11	1,005
CHALATENANGO	SAN LUIS DEL CARMEN	371	6	133	4	514
CHALATENANGO	CANCASQUE	346	6	296	19	667
CHALATENANGO	SAN ANTONIO LOS RANCHOS	26	3	460	1	490
CHALATENANGO	EL CARRIZAL	645	5	366	10	1,026
CHALATENANGO	SAN FERNANDO	553	7	296	9	865
CHALATENANGO	AZACUALPA	357	8	138	5	508
CHALATENANGO	SAN FRANCISCO LEMPA	244	1	204	0	449

LA LIBERTAD	SANTA TECLA	36,017	3,724	24,189	511	64,441
LA LIBERTAD	QUEZALTEPEQUE	12,003	588	8,498	418	21,507
LA LIBERTAD	CIUDAD ARCE	13,775	616	5,688	335	20,414
LA LIBERTAD	SAN JUAN OPICO	15,589	791	6,758	468	23,606
LA LIBERTAD	COLON	23,777	1,307	10,137	550	35,771
LA LIBERTAD	LA LIBERTAD	8,935	557	5,450	326	15,268
LA LIBERTAD	ANTIGUO CUSCATLAN	11,260	1,419	5,537	99	18,315
LA LIBERTAD	COMASAGUA	2,403	90	1,553	73	4,119
LA LIBERTAD	SAN PABLO TACACHICO	3,965	120	2,786	256	7,127
LA LIBERTAD	JAYAQUE	3,401	97	1,745	107	5,350
LA LIBERTAD	HUIZUCAR	2,567	69	1,045	280	3,961
LA LIBERTAD	TEPECOYO	3,141	115	1,014	149	4,419
LA LIBERTAD	TEOTEPEQUE	2,752	81	779	47	3,659
LA LIBERTAD	CHILTUIPAN	2,240	47	1,220	64	3,571
LA LIBERTAD	NUEVO CUSCATLAN	2,356	149	1,005	86	3,596
LA LIBERTAD	TAMANIQUE	2,698	76	1,341	158	4,273
LA LIBERTAD	SACACOYO	3,771	157	1,404	88	5,420
LA LIBERTAD	SAN JORGE VILLANUEVA	2,570	108	1,540	74	4,292
LA LIBERTAD	ZARAGOZA	5,140	259	2,637	349	8,385
LA LIBERTAD	TALNIQUE	1,765	56	531	40	2,392
LA LIBERTAD	SAN MATIAS	1,173	63	943	269	2,448
LA LIBERTAD	JICALAPA	1,350	126	287	33	1,796
SAN SALVADOR	SAN SALVADOR	98,586	11,027	69,542	1,822	180,977
SAN SALVADOR	CIUDAD DELGADO	25,314	2,123	20,161	523	48,121
SAN SALVADOR	MEJICANOS	32,548	3,666	34,431	599	71,244
SAN SALVADOR	SOYAPANGO	63,091	5,123	55,850	1,071	125,135
SAN SALVADOR	CUSCATANCINGO	15,965	1,289	13,435	268	30,957
SAN SALVADOR	SAN MARCOS	15,181	1,432	13,586	283	30,482
SAN SALVADOR	ILOPANGO	25,885	2,158	19,639	453	48,135
SAN SALVADOR	NEJAPA	5,673	295	3,552	157	9,677
SAN SALVADOR	APOPA	31,808	2,095	22,611	630	57,144
SAN SALVADOR	SAN MARTIN	18,052	1,030	10,453	456	29,991
SAN SALVADOR	PANCHIMALCO	8,714	379	5,218	202	14,513
SAN SALVADOR	AGUILARES	4,827	408	3,364	435	9,034
SAN SALVADOR	TONACATEPEQUE	14,450	968	9,083	278	24,779
SAN SALVADOR	SANTO TOMAS	6,615	293	4,054	148	11,110
SAN SALVADOR	SANTIAGO TEXACUANGOS	5,238	170	2,232	110	7,750
SAN SALVADOR	EL PAISNAL	2,151	97	2,161	70	4,479
SAN SALVADOR	GUAZAPA	4,787	298	3,535	157	8,777

SAN SALVADOR	AYUTUXTEPEQUE	7,025	675	7,073	177	14,950
SAN SALVADOR	ROSARIO DE MORA	2,774	73	1,348	103	4,298
CUSCATLÁN	COJUTEPEQUE	14,893	915	7,297	540	23,645
CUSCATLÁN	SUCHITOTO	2,957	117	5,273	46	8,393
CUSCATLÁN	SAN PEDRO PERULAPA	10,861	349	4,589	554	16,353
CUSCATLÁN	SAN JOSE GUAYABAL	2,476	66	1,038	84	3,664
CUSCATLÁN	TENANCINGO	1,254	61	1,135	66	2,516
CUSCATLÁN	SAN RAFAEL CEDROS	4,515	109	2,016	148	6,788
CUSCATLÁN	CANDELARIA	3,125	99	687	160	4,071
CUSCATLÁN	EL CARMEN	3,199	131	1,234	159	4,723
CUSCATLÁN	MONTE SAN JUAN	2,672	84	590	191	3,537
CUSCATLÁN	SAN CRISTOBAL	1,870	88	702	245	2,905
CUSCATLÁN	SANTA CRUZ MICHAPA	2,635	112	1,143	355	4,245
CUSCATLÁN	SAN BARTOLOME PERULAPIA	2,006	112	1,441	57	3,616
CUSCATLÁN	SAN RAMON	1,632	83	751	60	2,526
CUSCATLÁN	EL ROSARIO	990	51	281	159	1,481
CUSCATLÁN	ORATORIO	896	19	332	114	1,361
CUSCATLÁN	SANTA CRUZ ANALQUITO	816	12	260	14	1,102
SAN VICENTE	SAN VICENTE	11,582	705	7,288	924	20,499
SAN VICENTE	TECOLUCA	2,995	146	4,458	111	7,710
SAN VICENTE	SAN SEBASTIAN	3,658	77	2,108	48	5,891
SAN VICENTE	APASTEPEQUE	4,173	110	2,445	131	6,859
SAN VICENTE	SAN ESTEBAN CATARINA	748	28	1,277	22	2,075
SAN VICENTE	SAN IDELFONSO	1,260	32	986	532	2,810
SAN VICENTE	SANTA CLARA	755	21	696	29	1,501
SAN VICENTE	SAN LORENZO	1,411	42	753	35	2,241
SAN VICENTE	VERAPAZ	1,433	70	789	86	2,378
SAN VICENTE	GUADALUPE	1,442	39	365	253	2,099
SAN VICENTE	SANTO DOMINGO	1,788	87	784	43	2,702
SAN VICENTE	SAN CAYETANO ISTEPEQUE	1,183	34	807	24	2,048
SAN VICENTE	TEPETITAN	655	18	532	136	1,341
CABAÑAS	SENSUNTEPEQUE	10,025	476	3,522	444	14,467
CABAÑAS	ILOBASCO	15,401	457	5,323	871	22,052
CABAÑAS	VICTORIA	2,763	49	1,817	64	4,693
CABAÑAS	SAN ISIDRO	1,967	100	795	51	2,913
CABAÑAS	JUTIAPA	1,332	23	636	122	2,113
CABAÑAS	TEJUTEPEQUE	1,446	48	896	56	2,446
CABAÑAS	DOLORES	1,021	42	351	143	1,557
CABAÑAS	CINQUERA	144	14	362	14	534

CABAÑAS	GUACOTECHI	1,041	47	464	292	1,844
LA PAZ	ZACATECOLUCA	15,118	838	10,907	381	27,244
LA PAZ	SANTIAGO NONUALCO	7,475	443	4,136	1,372	13,426
LA PAZ	SAN JUAN NONUALCO	3,520	398	2,411	174	6,503
LA PAZ	SAN PEDRO MASAHUAT	4,868	251	3,193	124	8,436
LA PAZ	OLOCUILTA	5,852	223	3,167	600	9,842
LA PAZ	SAN PEDRO NONUALCO	2,316	144	967	102	3,529
LA PAZ	SAN FRANCISCO CHINAMECA	1,858	30	574	149	2,611
LA PAZ	SAN JUAN TALPA	1,852	118	884	82	2,936
LA PAZ	EL ROSARIO	3,196	177	1,655	112	5,140
LA PAZ	SAN RAFAEL OBRAJUELO	2,309	176	1,435	119	4,039
LA PAZ	SANTA MARIA OSTUMA	1,511	27	436	91	2,065
LA PAZ	SAN LUIS TALPA	4,655	254	2,377	105	7,391
LA PAZ	SAN ANTONIO MASAHUAT	757	325	485	103	1,670
LA PAZ	SAN MIGUEL TEPEZONTES	1,373	26	501	271	2,171
LA PAZ	SAN JUAN TEPEZONTES	985	36	291	27	1,339
LA PAZ	TAPALHUACA	664	40	600	92	1,396
LA PAZ	CUYULITAN	1,356	69	700	67	2,192
LA PAZ	PARAISO DE OSORIO	903	27	222	28	1,180
LA PAZ	SAN EMIGDIO	785	44	198	62	1,089
LA PAZ	JURUSALEN	584	9	275	80	948
LA PAZ	MERCEDES DE CEIBA	307	2	126	2	437
LA PAZ	SAN LUIS LA HERRADURA	4,078	284	2,506	484	7,352
USULUTÁN	USULUTAN	18,258	623	9,586	609	29,076
USULUTÁN	JIQUILISCO	7,316	182	7,811	132	15,441
USULUTÁN	BERLIN	3,580	110	2,652	470	6,812
USULUTÁN	SANTIAGO DE MARIA	3,798	134	3,270	301	7,503
USULUTÁN	JUCUAPA	4,268	140	2,422	133	6,963
USULUTÁN	SANTA ELENA	4,282	75	2,707	81	7,145
USULUTÁN	JUCUARAN	1,628	84	1,367	157	3,236
USULUTÁN	SAN AGUSTIN	921	18	821	9	1,769
USULUTÁN	OZATLAN	2,208	44	2,075	30	4,357
USULUTÁN	ESTANZUELAS	2,225	22	1,371	18	3,636
USULUTÁN	MERCEDES UMAÑA	2,927	62	1,561	39	4,589
USULUTÁN	ALEGRIA	1,887	40	1,466	53	3,446
USULUTÁN	CONCEPCION BATRES	2,534	43	1,121	81	3,779
USULUTÁN	SAN FRANCISCO JAVIER	1,275	23	752	28	2,078
USULUTÁN	PUERTO EL TRIUNFO	3,660	307	1,702	83	5,752
USULUTÁN	TECAPAN	1,258	29	1,380	111	2,778

USULUTÁN	SAN DIONICIO	1,207	21	421	36	1,685
USULUTÁN	EREGUAYQUIN	1,480	30	760	26	2,296
USULUTÁN	SANTA MARIA	2,271	75	1,162	54	3,562
USULUTÁN	NUEVA GRANADA	1,429	9	855	17	2,310
USULUTÁN	EL TRIUNFO	1,346	54	946	109	2,455
USULUTÁN	SAN BUENAVENTURA	995	32	801	13	1,841
USULUTÁN	CALIFORNIA	561	13	381	7	962
SAN MIGUEL	SAN MIGUEL	43,546	3,310	30,008	1,919	78,783
SAN MIGUEL	CHINAMECA	5,638	161	3,166	91	9,056
SAN MIGUEL	EL TRANSITO	4,621	63	2,594	128	7,406
SAN MIGUEL	CIUDAD BARRIOS	4,362	223	2,570	224	7,379
SAN MIGUEL	CHIRILAGUA	3,537	136	1,570	345	5,588
SAN MIGUEL	SESORI	1,720	85	901	37	2,743
SAN MIGUEL	SAN RAFAEL ORIENTE	2,905	76	2,073	182	5,236
SAN MIGUEL	MONCAGUA	4,172	156	2,894	328	7,550
SAN MIGUEL	LOLOTIQUE	2,491	557	1,967	95	5,110
SAN MIGUEL	SAN JORGE	2,808	25	1,067	94	3,994
SAN MIGUEL	CHAPELTIQUE	1,865	98	1,295	108	3,366
SAN MIGUEL	SAN GERARDO	1,245	16	624	13	1,898
SAN MIGUEL	CAROLINA	1,474	7	1,004	56	2,541
SAN MIGUEL	QUELEPA	1,317	89	625	48	2,079
SAN MIGUEL	SAN LUIS DE LA REINA	1,366	65	450	19	1,900
SAN MIGUEL	NUEVO EDEN DE SAN JUAN	743	12	358	35	1,148
SAN MIGUEL	NUEVA GUADALUPE	1,245	125	1,484	96	2,950
SAN MIGUEL	ULUAZAPA	750	38	472	54	1,314
SAN MIGUEL	COMACARAN	727	25	550	19	1,321
SAN MIGUEL	SAN ANTONIO DEL MOSCO	730	53	464	140	1,387
MORAZÁN	SAN FRANCISCO GOTERA	6,074	660	2,044	294	9,072
MORAZÁN	JOCORO	2,623	84	1,623	63	4,393
MORAZÁN	CORINTO	2,931	82	845	201	4,059
MORAZÁN	SOCIEDAD	2,149	83	1,099	172	3,503
MORAZÁN	CACAOPERA	1,411	63	1,407	527	3,408
MORAZÁN	GUATAGIAGUA	1,877	152	1,283	30	3,342
MORAZÁN	EL DIVISADERO	1,752	109	906	55	2,822
MORAZÁN	JOCOAYTIQUE	507	36	951	21	1,515
MORAZÁN	OSICALA	2,622	102	571	164	3,459
MORAZÁN	CHILANGA	2,058	121	850	98	3,127
MORAZÁN	MEANGUERA	624	52	1,865	26	2,567
MORAZÁN	TOROLA	467	15	461	30	973

MORAZÁN	SAN SIMON	1,314	197	895	304	2,710
MORAZÁN	DELICIAS DE CONCEPCION	1,686	58	272	71	2,087
MORAZÁN	JOATECA	1,122	13	212	73	1,420
MORAZÁN	ARAMBALA	691	34	525	10	1,260
MORAZÁN	LOLOTIQUILLO	779	351	364	29	1,523
MORAZÁN	YAMABAL	797	41	227	20	1,085
MORAZÁN	YOLOAYQUIN	1,191	45	238	29	1,503
MORAZÁN	SAN CARLOS	1,037	100	321	34	1,492
MORAZÁN	EL ROSARIO	363	21	247	21	652
MORAZÁN	PERQUIN	927	16	617	18	1,578
MORAZÁN	SENSEMBRA	566	77	204	23	870
MORAZÁN	GUALOCOCTI	701	60	272	209	1,242
MORAZÁN	SAN FERNANDO	352	17	280	60	709
MORAZÁN	SAN ISIDRO	439	93	317	85	934
LA UNIÓN	LA UNION	7,785	568	3,516	704	12,573
LA UNIÓN	SANTA ROSA DE LIMA	6,052	213	3,567	1,382	11,214
LA UNIÓN	PASAQUINA	3,651	80	1,853	145	5,729
LA UNIÓN	SAN ALEJO	2,353	104	1,240	1,553	5,250
LA UNIÓN	ANAMOROS	3,656	29	1,061	80	4,826
LA UNIÓN	EL CARMEN	3,075	75	725	490	4,365
LA UNIÓN	CONCHAGUA	6,296	306	1,847	1,315	9,764
LA UNIÓN	EL SAUCE	1,882	13	533	96	2,524
LA UNIÓN	LISLIQUE	2,129	17	183	751	3,080
LA UNIÓN	YUCUAYQUIN	1,297	223	493	436	2,449
LA UNIÓN	NUEVA ESPARTA	2,414	43	528	168	3,153
LA UNIÓN	POLOROS	2,073	19	346	141	2,579
LA UNIÓN	BOLIVAR	1,145	24	296	183	1,648
LA UNIÓN	CONCEPCION DE ORIENTE	1,648	12	481	157	2,298
LA UNIÓN	INTIPUCA	1,613	47	603	80	2,343
LA UNIÓN	SAN JOSE LAS FUENTES	812	18	224	11	1,065
LA UNIÓN	YAYANTIQUÉ	1,198	232	439	99	1,968
LA UNIÓN	MEANGERA DEL GOLFO	681	3	26	28	738

BIBLIOGRAFÍA

- Hill, T.P. (1988). Random-number guessing and the first digit phenomenon. *Psychological Reports* 62: 967-971.
- Hill, T.P. (1995a). Base-invariance implies Benford's law. *Proceedings of the American Mathematical Society* 123, 887-895.
- Hill, T.P. (1995b). The significant-digit phenomenon. *Amer. Math. Monthly* 102, 322-326.
- Hill, T.P. (1995c). A statistical derivation of the significant-digit law. *Statistical Science* 10, 354-363.
- Hill, T.P. (1997). Benford's law. *Encyclopedia of Mathematics Supplement*, vol. 1, 102.
- Hill, T.P. (1998). The first digit phenomenon. *The American Scientist* 10(4), 354-363.
- Hill, T.P. (1999a). Le premier chiffre significatif fait sa loi. *La Recherche Hors Série*, no. 2, 72-76.
- Hill, T.P. (1999b). La primera cifra significativa dicta su ley. *Mundo Científico* num. 199 (marzo), 56-59.
- Hill, T.P. (1999c). The difficulty of faking data. *Chance* 26, 8-13.
- Hill, T.P. (2005). Newton's Method Obeys Benford's Laws. Conference, Leiden University.
- Hill, T.P. (2002). Recent applications of Benford's law. Conference, Leiden University.
- Hill, T.P. and K. Schürger (2005). Regularity of digits and significant digits of random variables. *Journal of Stochastic Processes and their Applications* 115, 1723-1743.
- Nigrini, M. J. (1992). The Detection of Income Tax Evasion Through an Analysis of Digital Frequencies. Ph.D. thesis. Cincinnati, OH: University of Cincinnati.

- Nigrini, M. J. (1997). *Digital Analysis Tests and Statistics*. Allen, Texas: The Nigrini Institute, Inc. Mark_Nigrini@classic.msn.com
- Nigrini, M.J. (1999a). I've got your number. *Journal of Accountancy* 187, 79-83.
- Nigrini, M.J. (2000a). *Digital Analysis Using Benford's Law: Tests Statistics for Auditors*. Vancouver, Canada: Global Audit Publications.
- Nigrini, M.J. and L. Mittermaier (1997). The use of Benford's law as an aid in analytical procedures. *Auditing : A Journal of Practice and Theory* 16(2), 52-67.
- Dorogovtsev, S.N., Mendes, J.F.F. and J.G. Oliveira (2006). Frequency of occurrence of numbers in the World Wide Web. *Physica A* 360, 548. <http://arxiv.org/abs/physics/0504185>.
- Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society* 78, 551-572.
- Vicente Quezada Palomo (1988), *Lecciones de Calculo de Probabilidad*, Díaz de Santos
- *Apuntes de Teoría de la Medida* (2006), Herbert A. Medina, Editorial Universitaria UES.
- John Walthoe (2005). En Busca del Número Uno, www.cup.cam.ac.uk

-
- J. Aaronson (1997). *An Introduction to Infinite Ergodic Theory*, *Mathematical Surveys and Monographs* 50, American Mathematical Society, Providence RI, USA,.
 - J. Aaronson (1977). Rational ergodicity and a metric invariant for Markov shifts, *Israel Journal of Mat.*, 27, 93-123.
 - J. Aaronson (1978). On the pointwise ergodic theory of transformations preserving infinite measures, *Israel Journal of Math*, 32, 67-82.

- J. Aaronson (1981). The asymptotic distributional behavior of transformations preserving infinite measures, *J. D'Analyse Math*, 39, 203-234.
- J. Aaronson (1987). The intrinsic normalising constants of transformations preserving infinite measures, *J. D'Analyse Math.*, 49, 239-270.
- J. Aaronson and B. Weiss (1998). On the asymptotics of a 1-parameter family of infinite measure preserving transformations, *Bol. Soc. Brasil. Mat. (N.S.)*, 29, 181-193.
- J. Aaronson, H. Nakada, O. Sarig, R. Solomyak (2002). Invariant measures and asymptotics for some skew products (as in *Israel J. Math.* 128 93--134.)
- Jon Aaronson, Max Thaler, Roland Zweimueller (2005). Occupation times of sets of infinite measure for ergodic transformations (as in *ETDS* 25 959--976)
- Jon Aaronson, Tom Meyerovitch, Absolutely continuous, invariant measures for dissipative, ergodic transformations
- J. Aaronson (2000). An overview of infinite ergodic theory (as in *Descriptive set theory and dynamical systems*, 1--29, *LMS Lecture Notes*, 277)
- J. Aaronson, H. Nakada (2000). Multiple Recurrence of Markov Shifts and Other Infinite Measure Preserving Transformations (as in *Israel J. Math.* 117 285--310.)
- Marian Prendar, George Judge and Laura Schechter (2007). An Empirical Non-Parametric Likelihood Family of Data-Based Benford-Like Distributions. *AMS Classification: Primary 62E20.*
- Tomás HOBZA and Igor VAJDA (2001). On the newcomb-benford law in models of statistical data. *revista matemática complutense*. vol. XIV, num. 2, 407-420.
- Cindy Durtschi, Williams Hillison and Carl Pacini (2004). The effective use of Benford's Law to assist in detecting fraud in accounting data. *Journal of forensic accounting* 1524-5586 Vol. V pp. 17-34.

- Fernando Gales García (2007). Definiciones originales de la integral y medida de Lebesgue, *Misceláneo matemático* 44, pp 83-100.
- Christian Schäfer, Jörg Meter Schräpler, Klaus Robert Müller and Gert G. Wayner, Automatic identification of faked any fraudulent interview in survey by two different methods, www.amstat.org/Sections/Srms/Proceedings/y2004/Files/Jsm2004-000257.pdf.
- Arno Berger, Leonid A. Bunimovich and Ted Hill (2005). One dimensional dynamical systems and Bedford's Law, *American Mathematical Society*, 357 pp 197-219.
- Lisa Bloomer and Ted Hill. Random probability measures with given mean and variance, Doctoral dissertation, Georgia Institute of Tecnology.
- David Swason, Moon Jung Cho and John Eltinger, 2003, Detecting possibly fraudulent error prone survey data using Bedford's Law, www.amstat.org/Sections/Srms/Proceedings/y2003/Files/JSM2003-000205.pdf.
- Yongyi Wang and Steven Pedlow (2004) Detecting falsified cases in SCF 2004 using Bedford's Law, www.amstat.org/sections/SRMS/Proceedings/y2005/y2005.html
www.amstat.org/sections/SRMS/Proceedings/y2005/Files/JSM2005-000519.pdf.
- Julio Grafe Arias y Roberto Escuder Valles (1984). Tipos extremos de distribuciones de probabilidad y funciones de distribuciones asociadas, *Estadística española* Num 104, pp 35-47.
- Sukanto Bhattacharga and Kuldeep Eumar (2002) Condicional probability of actualy detecting of financial fraud a neutrosophic extensión to Bedford's Law, *Journal-ref: International Journal of Applied Mathematics*, Vol. 17, No 1, pp. 7-14, 2005.

- Li Zhipeng, Cong Lin and Wang Huajia (2004). Discussion on Bedford's Law and its application, arXiv:math/0408057.
 - Steven J. Miller and Mark Nigrini (2008). Order statistics and Bedford's Law, *International Journal of Mathematics and Mathematical Sciences* Volume 2008 (2008), Article ID 382948, 19 pages.
 - Steven J. Miller and Mark Nigrini (2007). The modulo 1 central limit theorem and Bedford's Law for products, *International of algebra* 2, N° 3, 119-130.
 - S.R.S Varadham (2000). *Probability theory*, ISBN 0-8218-2852-5.
 - Richard L. Wheden and Atoni Zygmund (1977). *Measure and integral*, Marcel Dekker Inc, ISBN 0-8247-6499-4.
 - William Feller (1989). *Introducción a la teoría de probabilidad y sus aplicaciones Vol II Segunda Edición*, Limusa.
 - Ya G. sinal (1937). *Dynamical systemms: ergodic with applications to dynamical systems*, New York: Springer-Verlag.
 - Robert L. Devaney(1987). *An introduction to chaotic dynamical systems*. Adisson-Wesley Publishing Company, Inc.
-
- Hill, T.P. and A. Berger (2007). "Newton's method obeys Benford's law", *American Mathematical Monthly* 114, 588-601
 - Hill, T.P. with Klaus Schürger (2005). "Regularity of Digits and Significant Digits of Random Variables," *Journal of Stochastic Processes and Their Applications* **115**, 1723-1743
 - Hill, T.P. (1996). "A note on distributions of true versus fabricated data," *Perceptual and Motor Skills* **83**, 776-778
 - Hill, T.P. (1993). "Partitioning inequalities in probability and statistics," *Stochastic Inequalities, IMS Lecture Notes* **22**, 116-132.
 - Hill, T.P. (1987). "Expectation inequalities associated with prophet inequalities," *Stochastic Analysis and Applications* **5**, 299-310.

- Hill, T.P. '(1987) 'A sharp partitioning-inequality for non-atomic probability measures based on the mass of the infimum of the measures," Probability Theory and Related Fields **75**, 143-147
- Hill, T.P. (1987)'Partitioning general probability measures," Annals of Probability **15**, 804-813.
- Oded Kafri, Entropy Principle in Direct Derivation of Benford's, Law arXiv:0901.3047.
- Bartolo Luque, Lucas Lacasa, The first digit frequencies of primes and Riemann zeta zeros tend to uniformity following a size-dependent generalized Benford's law, arXiv:0811.3302
- Dennis Jang, Jung Uk Kang, Alex Kruckman, Jun Kudo, Steven J. Miller, Chains of distributions, hierarchical Bayesian models and Benford's Law, arXiv:0805.4226
- Lutz Duembgen, Christoph Leuenberger, (2008). Explicit bounds for the approximation error in Benford's law, Journal-ref: Electronic Communications in Probability **13** 99-112
- Alex Ely Kossovsky, Towards A Better Understanding Of The Leading Digits Phenomena, arXiv:math/0612627
- Steven J. Miller, Mark J. Nigrini, The Modulo 1 Central Limit Theorem and Benford's Law for Products, arXiv:math/0607686
- Werner Hurlimann, Generalizing Benford's law using power laws: application to integer, arXiv:math/0607166
- Jeffrey C. Lagarias, K. Soundararajan (2006). Benford's law for the $3x+1$ function, Journal-ref: J. London Math. Soc. **74** 289--303
- Wendy K . Tam Cho , Brian J . Gaines (2007), Breaking the (Benford), Law The American Statistician Aug 2007, Vol. 61, No. 3, 218-223: 218-223.
- Ricardo J . Rodriguez (2004). First Significant Digit Patterns From Mixtures of Uniform Distributions The American Statistician Feb 2004, Vol. 58, No. 1, 64-71: 64-71.

