DIPLOMA EN MATEMÁTICA
INSTITUTO DE PERFECCIONAMIENTO Y ESTUDIOS SUPERIORES CFE/ANEP-UDELAR

Ley de Benford

Trabajo final

Prof. María Caputi Zunini

Orientador: Dr. Marcelo Lanzilotta

2016

Agradecimientos

A Marcelo Lanzilotta, por permitirme trabajar con libertad, por ser un orientador paciente y estar siempre dispuesto al diálogo y a evacuar mis dudas.

A Nicolás, por su amor cotidiano y apoyo permanente.

A Mateo, por su alegría y cariño.

A mis padres y hermana, por apoyarme siempre en todos mis proyectos personales y académicos.

A mis compañeros y docentes de los cursos del Diploma, por las muchas tertulias los viernes al medio día. En especial a Alicia por compartir tantas horas de estudio, viaje y amistad.

A mis amigas de siempre, por estar, escuchar y dar ánimos. En especial a So por ayudarme con la presentación.

A Damián y Patricia, por sus comentarios y por abrirme las puertas de su hogar en Salto.

A la familia: mis padres, suegros y a todas las tías, tíos y primos por su invalorable ayuda con Mateo. En especial, las tías Gimena, Vito y Sole por sus amorosos cuidados.

Índice

Resumen

La Ley de Benford establece, contrariamente a la intuición, que, en algunos conjuntos de datos numéricos, la frecuencia de aparición del primer dígito significativo no es uniforme. La frecuencia con la que aparece cada dígito sigue una proporción particular que se explicita en la denominada por Benford [1] en 1938 como "Ley de los números anómalos". El 1 aparece como primer dígito significativo un 30,1% de las veces, el 2 un 17,6%, el 3 un 12,5%, el 4 un 9,7%, el 5 un 7,9%, el 6 un 6,7%, el 7 un 5,8%, el 8 un 5,1% y el 9 un 4,6%, aproximadamente. Esta ley permaneció como una mera "curiosidad estadística" por varias décadas. En 1992 fue catapultada a la luz e interés público por el contador norteamericano Mark Nigrini, quien en su tesis de doctorado la utilizó para detectar fraudes en las declaraciones fiscales.

En este trabajo se estudia la Ley de Benford como una distribución de probabilidad, primero en base 10, y luego en cualquier base natural mayor a 2 con una aproximación más formal. Así, siguiendo a Hill [3] se reconstruye la σ-álgebra de sucesos que será dominio para la función de probabilidad. Se demuestra la condición necesaria y suficiente para que una variable aleatoria siga la Ley de Benford y que la Ley de Benford es la única ley de probabilidad sobre la mantisa invariante frente a cambios de escala.

En el texto se exponen métodos estadísticos para analizar si un conjunto de datos sigue o no la Ley de Benford. Se ejemplifican dichos métodos con conjuntos de datos clásicos que siguen la ley y se analiza cumplimiento de la Ley de Benford en los resultados del censo realizado en Uruguay en 2011.

Palabras clave: Ley de Benford; Primer dígito significativo.

Introducción

En este trabajo se busca estudiar la Ley de Benford de modo de ir avanzando en abstracción y formalidad al pasar los capítulos. En una primera instancia se exponen los antecedentes históricos que refieren a la construcción de la Ley de Benford. Así, se revisan sintéticamente sus primeras referencias en las publicaciones de Simon Newcome y más tarde de Frank Benford. Citando luego los aportes más contemporáneos de Hill y Nigrini que serán retomados en los siguientes capítulos.

Se comienza el trabajo con la Ley de Benford para el primer dígito en el sistema decimal. Allí, se modela la ley primero como una distribución de probabilidad discreta y luego como una distribución de probabilidad absolutamente continua. Bajo esta segunda aproximación, se prueban algunos resultados trascendentes como la condición necesaria y suficiente para que una variable aleatoria siga la Ley de Benford y que la Ley de Benford es la única ley de probabilidad sobre la mantisa invariante frente a cambios de escala.

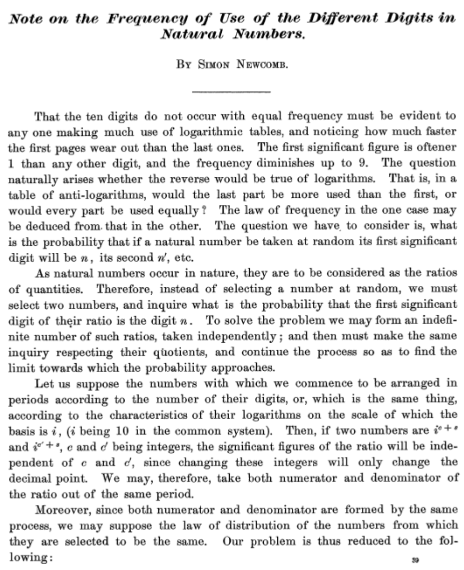
En otro escalón de la abstracción, se generaliza la Ley a los primeros k dígitos significativos y se estudian algunas particularidades de esta. Luego, se pone el foco en los primeros dos dígitos significativos por tener aplicación en la detección de fraude.

Aplicando lo expuesto hasta el momento, se analizan algunos métodos que pueden ser utilizados para estudiar si un conjunto de datos sigue, o no, la Ley de Benford. Allí, se exponen la aplicación del test de bondad de ajuste de Chi cuadrado y el Mean Absolute Deviation Test. A continuación, se realizan aplicaciones de estos test a algunos ejemplos clásicos de conjuntos de datos que siguen la Ley de Benford. A saber: la sucesión de Fibonacci, la sucesión de números de Lucas y la sucesión de potencias naturales de 2. También se analiza el cumplimiento de la Ley de Benford en el censo de la población uruguaya de 2011.

En el noveno capítulo, se busca enunciar la Ley de Benford para los primeros k dígitos significativos generalizándola para cualquier base natural mayor a 1. Para esto se construye de forma rigurosa una σ-álgebra de sucesos que será dominio para la función de probabilidad. Se prueban algunas propiedades de la σ-álgebra y se reescribe y prueba en este contexto que la Ley de Benford es la única ley de probabilidad invariante frente a cambios de escala.

Capítulo 1. Antecedentes históricos

La primera publicación sobre la que posteriormente sería llamada "Ley de Benford" fue autoría del astrónomo norteamericano Simon Newcomb y data de 1881. En un artículo de dos páginas titulado "Note on the Frequency of Use of the Different Digits in Natural Numbers" publicado en American Journal of Mathematics el autor describe sus observaciones en el uso de las tablas de logaritmos. Newcomb detecta que las hojas más desgastadas eran las primeras. Así, enuncia de forma coloquial una ley de distribución para la frecuencia de aparición de los números naturales. "La ley de probabilidad de la ocurrencia de los números es tal que las mantisas de sus logaritmos son equiprobables."[1] El artículo de Newcomb pasa desapercibido por la comunidad matemática de la época, quedando como una mera curiosidad.



Note on the Frequency of Use of the Different Digits in Natural Numbers.

BY SIMON NEWCOMB.

That the ten digits do not occur with equal frequency must be evident to any one making much use of logarithmic tables, and noticing how much faster the first pages wear out than the last ones. The first significant figure is oftener 1 than any other digit, and the frequency diminishes up to 9. The question naturally arises whether the reverse would be true of logarithms. That is, in a table of anti-logarithms, would the last part be more used than the first, or would every part be used equally? The law of frequency in the one case may be deduced from that in the other. The question we have to consider is, what is the probability that if a natural number be taken at random its first significant digit will be n, its second n', etc.

As natural numbers occur in nature, they are to be considered as the ratios of quantities. Therefore, instead of selecting a number at random, we must select two numbers, and inquire what is the probability that the first significant digit of their ratio is the digit n. To solve the problem we may form an indefinite number of such ratios, taken independently; and then must make the same inquiry respecting their quotients, and continue the process so as to find the limit towards which the probability approaches.

Let us suppose the numbers with which we commence to be arranged in periods according to the number of their digits, or, which is the same thing, according to the characteristics of their logarithms on the scale of which the basis is i, (i being 10 in the common system). Then, if two numbers are i^{c+s} and $i^{c'+s}$, c and c' being integers, the significant figures of the ratio will be independent of c and c', since changing these integers will only change the decimal point. We may, therefore, take both numerator and denominator of the ratio out of the same period.

Moreover, since both numerator and denominator are formed by the same process, we may suppose the law of distribution of the numbers from which they are selected to be the same. Our problem is thus reduced to the following:

39

Frank Benford en 1938, aparentemente sin conocer el artículo de Newcomb, realiza la misma observación en las tablas de logaritmos. Éstas eran muy utilizadas para realizar rápidamente multiplicaciones y divisiones antes de la era de las calculadoras. Benford tenía contacto frecuente con dichas tablas, ya que fue un físico que trabajó en el centro de investigaciones de General Electric en Nueva York. Para comprobar empíricamente la llamada por Benford como "Ley de los números anómalos" recolectó 20.229 números extraídos de las más diversas fuentes (entre otros: periódicos, revistas, estadísticas de béisbol, longitudes de ríos, tablas de constantes matemáticas y físicas) y sin ayuda de calculadoras o procesadores de datos realizó los cálculos de las proporciones en las que aparecían los primeros 9 números enteros positivos. Esos datos los recogió en una tabla que fue retomada en Nigrini ([9], pág. 4) quien afirma que la misma contiene algunos errores, entre otros de redondeo.

La publicación de Benford en Proceedings of the American Philosophical Society tuvo una repercusión mucho mayor que la de Newcomb, según Hill ([5], pág.

[1] Newcomb, S. (1881). *Note on the Frequency of Use of the Different Digits in Natural Numbers*. American Journal of Mathematics , 4 (1/4), Pág. 40.

359), debido a estar previo al artículo de H. A. Bethe, M. E. Rose y L. P. Smith, titulado "The Multiple Scattering of Electrons", que luego sería muy renombrado. Habiendo pasado desapercibida la publicación de Newcomb, la bautizada por Benford como "Ley de los números anómalos" sería conocida como "Ley de Benford".

Mark Nigrini, un contador estadounidense, se encuentra por primera vez con la Ley de Benford en un curso de doctorado en 1989 en la Universidad de Cincinnati (EEUU). Según Nigrini [9] interesado por el tópico y tras leer muchos de los artículos que había publicados en esa materia decide contactarse con Ralph A Raimi, quien era el autor más citado en la temática del momento. Raimi consideraba que la ley no tenía fines prácticos. En ese momento, Nigrini [9] si bien vislumbraba que podía usarse la ley para la detección de anomalías en las declaraciones de impuestos, descarta la aplicación práctica por no tener los medios informáticos necesarios para poder procesar "gran cantidad" de datos. Sin embargo, a principios de 1991, Nigrini con ayuda de John Byant decide buscar un modelo matemático de cómo los contribuyentes engañan al fisco y de cómo la Ley de Benford puede ayudar a detectarlos. Esta idea sería la semilla de la tesis de doctorado de Nigrini publicada en 1992 y de muchas otras publicaciones del autor sobre el tópico. La tesis de Nigrini es la que lanza a la fama a la Ley de Benford. A partir de allí según Joseph T. Wells (autor del prólogo de Nigrini [9]) se elaboraron otras aplicaciones de la Ley de Benford.

Desde la publicación de Benford fueron numerosos los intentos por dar una prueba matemática rigurosa de la Ley. Pero, ésta recién fue conseguida por Hill en 1995. Según Hill ([5], pág.360-361) hubo que sortear fundamentalmente dos escollos: el primero, algunos conjuntos de datos satisfacen la Ley de Benford mientras que otros no, y no se logró encontrar un método por el cual se pudiese construir conjuntos que la satisfacen y conjuntos que no; el segundo, la prueba debe satisfacer la axiomática moderna que rige la teoría de probabilidades.

Capítulo 2. Preliminares

Definición axiomática de Probabilidad[2]

Sea Ω un conjunto no vacío, llamado espacio de sucesos elementales. Cada elemento de Ω es llamado suceso elemental.

σ-álgebra

Sea $\mathscr{A}$ un subconjunto no vacío del conjunto de partes de Ω que verifica si un conjunto pertenece a $\mathscr{A}$, su complemento también pertenece a $\mathscr{A}$; y que la unión de cualquier conjunto numerable de elementos de $\mathscr{A}$ es un elemento de $\mathscr{A}$.

Axiomas de la teoría de probabilidades

Un espacio de probabilidad es una terna (Ω, $\mathscr{A}$, *P*), donde Ω es un conjunto no vacío, $\mathscr{A}$ es una la σ-álgebra de Ω y *P* es una función definida para cada elemento de la σ-álgebra $\mathscr{A}$ que verifica el siguiente sistema de axiomas.

<u>Axioma 1</u>: A cada suceso A de $\mathscr{A}$ le corresponde un número no negativo *P*(A), llamado probabilidad del suceso A.

<u>Axioma 2</u>: $P(\Omega)=1$.

<u>Axioma 3</u>: Si $A_1, A_2, \ldots$ es un conjunto finito o numerable de sucesos disjuntos dos a dos, entonces $P\left(\bigcup_i A_i\right)=\sum_i P(A_i)$.

σ-álgebra generada

Si *A* es una colección de subconjuntos de Ω, la intersección de todas las σ-álgebras que contienen a *A* es también una σ-álgebra, denotada por $\langle A\rangle$ y denominada σ-álgebra generada por *A*. Esta es, por construcción, la menor σ-álgebra que contiene a la colección *A*.

[2] En base a Petrov, V., & Mordecki, E. (2003). *Teoría de Probabilidades.* Moscú: Editorial URSS. Pág. 11-13.

σ-álgebra de Borel

La σ-álgebra generada por la colección de todos los intervalos abiertos finitos se denomina álgebra de Borel (sobre $\mathbb{R}$) que notaremos $\mathcal{B}$ (o $\mathcal{B}(\mathbb{R})$ si hubiera lugar a confusión). El álgebra de Borel es la mínima σ-álgebra sobre $\mathbb{R}$ que contiene a los subconjuntos cerrados, a los intervalos abiertos o cerrados, a los intervalos semiabiertos de la forma (a, b], o a los intervalos de la forma (−∞, b]. Un elemento de esta σ-álgebra es llamado conjunto boreliano.

Si $E \subset \mathbb{R}$, se define $\mathcal{B}(E)$ como la σ-álgebra generada por los conjuntos que son abiertos relativos en *E*. Se puede verificar que $\mathcal{B}(E) = \{B \cap E : B \in \mathcal{B}(\mathbb{R})\}$. Llamaremos borelianos de E a los elementos de esta σ-álgebra.

Además, se verifica que la preimagen de un boreliano de $\mathcal{B}(\mathbb{R})$ por una función real continua es un boreliano $\mathcal{B}(\mathbb{R})$. En particular, si la función es $f : \mathbb{R} \to \mathbb{R}, f(x) = \frac{1}{a}x$, con a real positivo, se verifica que si $A \in \mathcal{B}(\mathbb{R}^+)$, entonces $aA \in \mathcal{B}(\mathbb{R}^+)$.

Variable aleatoria

Sea (Ω, $\mathcal{A}$, *P*) un espacio de probabilidades. Se denomina variable aleatoria a una función real $X : \Omega \to \mathbb{R}$ que verifique que el suceso $\{\omega \in \Omega : X(\omega) \leq x\}$ es un elemento de la σ-álgebra $\mathcal{A}$, para todo *x* real.

Función de distribución de una variable aleatoria

Se denomina función de distribución de la variable aleatoria *X* a la función $F : \mathbb{R} \to [0,1],\ \mathrm{F}(x) = P(X \leq x)$.

Variables aleatorias con distribución discreta

Se dice que una variable aleatoria *X* tiene distribución discreta si existe un conjunto $E \subset \mathbb{R}$ numerable, tal que se verifica que $P(X \in E) = 1$.

Variables aleatorias con distribución absolutamente continua

Se dice que una variable aleatoria *X* tiene distribución absolutamente continua si su función de distribución *F* puede representarse en la forma: $F(x)=\int_{-\infty}^{x} p(u)du$ para todo *x* real, donde *p* es una función no negativa e integrable. La función *p* se denomina función densidad de la distribución de la variable aleatoria *X*.

Distribución uniforme

Se dice que una variable aleatoria tiene distribución uniforme en el intervalo *(a,b)*, donde $a<b$, si su densidad es: $p:\mathbb{R}\to\mathbb{R},\ p(u)=\begin{cases}\dfrac{1}{b-a} & \text{si u} \in (a,b)\\ 0 \text{ en otro caso}\end{cases}$.

Por lo tanto, su función de distribución será:

$$F:\mathbb{R}\to\mathbb{R}, F(u)=\begin{cases}0,\ \text{si } u\le a\\ \dfrac{u-a}{b-a},\ \text{si u}\in(a,b)\\ 1,\ \text{si } u\ge b\end{cases}.$$

Esta función es continua para todo real, y derivable para todo real distinto de *a* y *b*.

Chi cuadrado[3]

Sea *X* una variable aleatoria con distribución discreta que toma valores $x_1,\ldots, x_k$ con probabilidades $p_1,\ldots,p_k$.

Una medida de la discrepancia entre las frecuencias observadas en un cierto fenómeno y esperadas para la distribución de probabilidad de *X* viene proporcionada por el estadístico χ^2, dado por $\chi^2=\sum_{j=1}^{k}\frac{\left(f_j-e_j\right)^2}{e_j}$, siendo f_j la frecuencia observada, *N* la cantidad de datos y $e_j=N.p_j$ la frecuencia esperada para la distribución de probabilidad de *X*.

[3] En base a Spiegel, M. *Estadística*. Madrid: McGraw Hill. Pág.204-205.

Si $\chi^2 = 0$, las frecuencias observadas y teóricas coinciden exactamente. Si $\chi^2 > 0$ las frecuencias observadas y teóricas no coinciden exactamente. A valores más grandes de χ^2, mayor discrepancia entre las frecuencias observadas y esperadas.

El valor *k* será denominado como la cantidad de grados de libertad.

Test Chi Cuadrado para la Bondad de ajuste

El test de Chi Cuadrado puede utilizarse para determinar la calidad del ajuste mediante distribuciones teóricas de distribuciones empíricas. En la práctica, las frecuencias esperadas se calculan sobre la base de una hipótesis H_0. Al contrastar una cierta hipótesis, la máxima probabilidad que estamos dispuestos a correr el riesgo de rechazarla cuando debiera ser aceptada se llama nivel de significación y es denotado por α. Es frecuente un nivel de significación de 0,05 o 0,01, es decir del 5% o del 1%. Si bajo tal hipótesis el valor calculado para χ^2 es mayor que algún valor crítico (tal como $\chi^2_{.95}$ o $\chi^2_{.99}$, que son los valores críticos de los niveles de significación (α) 0,05 y 0,01 respectivamente), debemos concluir que las frecuencias observadas difieren significativamente de las frecuencias esperadas y rechazaremos H_0 al correspondiente nivel de significación α; en caso contrario la aceptaremos. Este procedimiento se llama test de Chi Cuadrado de hipótesis.

Lema de Zorn

Si < es un orden parcial del conjunto A no vacío, en el que toda cadena tiene una cota superior, entonces A tiene un elemento maximal en <.

Capítulo 3. Ley de Benford o ley del primer dígito significativo

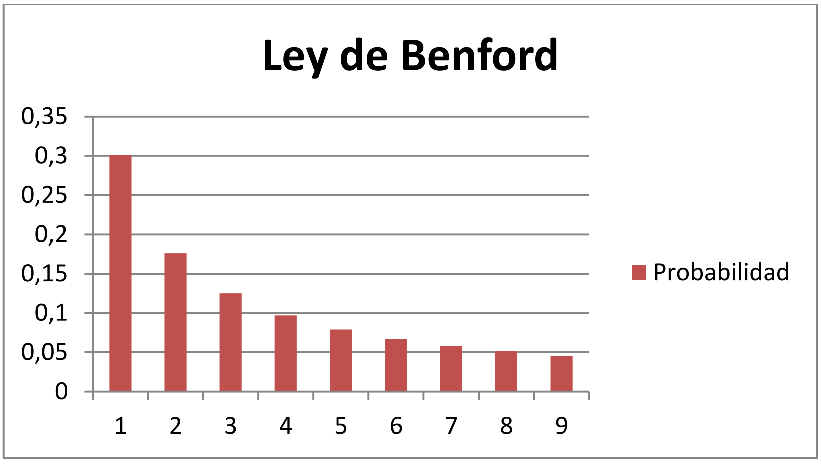
3.1 La Ley de Benford como distribución de probabilidad discreta. Una primera aproximación

La Ley de Benford establece una distribución para los primeros dígitos significativos de un número real positivo. Siendo el primer dígito significativo de un real positivo el dígito distinto de cero que aparece más a la izquierda en su expresión decimal. Por ejemplo, el primer dígito significativo de 73,15 es 7, el de 0,045 es 4 y el de e es 2.

Una primera descripción de la Ley de Benford como una ley de probabilidad podría tabularse como sigue.

Dígito d	1	2	3	4	5	6	7	8	9
Probabilidad de que el primer dígito significativo sea d	0,301	0,176	0,125	0,097	0,079	0,067	0,058	0,051	0,046

Se observa que 0 no se considera como primer dígito significativo.



Con más formalidad podremos enunciar la Ley:

$$Probabilidad\left(Primer\ dígito\ significativo = d\right) = \log_{10}\left(1+\frac{1}{d}\right) \quad d = 1\ldots 9\,.$$

Así definida la Ley es una distribución de probabilidad discreta, ya que,

$\log_{10}\left(1+\frac{1}{d}\right)\geq 0$ y $\sum_{d=1}^{d=9}\log_{10}\left(1+\frac{1}{d}\right)=1$, pues, $\sum_{d=1}^{d=9}\log_{10}\left(1+\frac{1}{d}\right)=\sum_{d=1}^{d=9}\log_{10}\left(\frac{d+1}{d}\right)=$

$=\log_{10}\prod_{d=1}^{9}\frac{d+1}{d}=\log_{10}\frac{10}{1}=1$.

Proposición 3.1.1

$$P\left(Primer\ dígito\ significativo=1\right)=\sum_{i=5}^{9}P\left(Primer\ dígito\ significativo=i\right).$$

Demostración.

$$\sum_{i=5}^{9}P\left(Primer\ dígito\ significativo=i\right)=\sum_{i=5}^{9}\log_{10}\left(1+\frac{1}{i}\right)=\log_{10}\left(\prod_{i=5}^{9}\left(\frac{i+1}{i}\right)\right)=$$

$$=\log_{10}\left(\frac{6}{5}.\frac{7}{6}.\frac{8}{7}.\frac{9}{8}.\frac{10}{9}\right)=\log_{10}\left(\frac{10}{5}\right)=\log_{10}2=P\left(Primer\ dígito\ significativo=1\right).\square.$$

Proposición 3.1.2

$$P\left(Primer\ dígito\ significativo=1\right)=\sum_{i=2}^{3}P\left(Primer\ dígito\ significativo=i\right).$$

Demostración.

$$P\left(Primer\ dígito\ significativo=2\right)+P\left(Primer\ dígito\ significativo=3\right)=$$

$$=\log_{10}\left(1+\frac{1}{2}\right)+\log_{10}\left(1+\frac{1}{3}\right)=\log_{10}\left(\frac{3}{2}.\frac{4}{3}\right)=\log_{10}\left(\frac{4}{2}\right)=$$

$$=\log_{10}2=P\left(Primer\ dígito\ significativo=1\right).\square.$$

Proposición 3.1.3

$$P\left(Primer\ dígito\ significativo=4\right)=1-3\log_{10}2\ .$$

Demostración.

$$P\left(Primer\ dígito\ significativo=4\right)=1-P\left(Primer\ dígito\ significativo\neq 4\right)=$$

$$=1-\sum_{i=1}^{3}P\left(Primer\ dígito\ significativo=i\right)-\sum_{i=5}^{9}P\left(Primer\ dígito\ significativo=i\right)=$$

$$=1-2\log_{10}2-\log_{10}2=1-3\log_{10}2.\square$$

Si bien los anteriores enunciados de la Ley de Benford, que responden a la caracterización clásica, funcionan de manera intuitiva, es necesario para avanzar "matemáticamente" definir con precisión el concepto de primer dígito significativo y ahondar en la σ-álgebra de sucesos que tomaremos como dominio para la función de probabilidad. Lo que se realizará en el Capítulo 9.

3.2 La Ley de Benford (generalizada) con distribución de probabilidad absolutamente continua. Una segunda aproximación

Definición 3.2.1. Función mantisa. Definimos la función Mantisa, $M:\mathbb{R}^{+}\rightarrow[1{,}10)$ tal que $M(x)=m$, siendo *m* el único número perteneciente a [1,10) que verifica $x=m.10^{n}$ para algún $n\in\mathbb{Z}$. Además a *m* lo llamaremos mantisa de x.

Por ejemplo, la mantisa de 0,05 es 5, la de 612 es 6,12 y la de $\frac{1}{3}$ es $3{,}\widehat{3}$.

La probabilidad del primer dígito significativo *d* de un entero positivo *x*, según la Ley de Benford, sería $\log_{10}\left(1+\frac{1}{d}\right)$ con $d=1\ldots9$.

¿Qué significa que el primer dígito significativo del real *x* sea *d* en términos de su mantisa? Quiere decir que $M(x)\in[d,d+1)$.

Entonces, $P\left(M(x)\in[d,d+1)\right)=\log_{10}\left(1+\frac{1}{d}\right)$ con $d=1\ldots9$.

La generalización propuesta por Antoine Nectourx [7] a la que refiere como Ley de Benford (generalizada) afirma que una variable aleatoria *X* la cual consiste en elegir al azar un número positivo y calcular su mantisa (*X* toma valores en [1,10)), sigue la Ley de Benford si su función de densidad es $f:\mathbb{R}\rightarrow\mathbb{R},\ f(u)=\begin{cases}\frac{1}{u.\log 10} & \text{si } u\in[1{,}10)\\ 0 \text{ en otro caso}\end{cases}$.

Queremos estudiar la distribución del primer dígito significativo, ¿Qué significa que el primer dígito significativo del real *x* sea *d* en términos de nuestra variable aleatoria? Quiere decir que $X(x)\in[d,d+1)$.

La probabilidad de que $X\in[d,d+1)$ es $P(d\leq X<d+1)=\int_{d}^{d+1}f(x)dx$.

Entonces, admitiendo que X tiene como densidad la función f definida anteriormente,

$$P(d \leq X < d+1) = \int_{d}^{d+1} \frac{1}{x.\log 10}\,dx = \frac{1}{\log 10}\left(\log(d+1) - \log(d)\right) = \frac{\log\left(\frac{d+1}{d}\right)}{\log 10} = \log_{10}\left(1+\frac{1}{d}\right).$$

Del resultado anterior se evidencia la generalización.

La función de distribución para una variable aleatoria que sigue la Ley de Benford sería entonces: $F:\mathbb{R}\to\mathbb{R}, F(u) = \begin{cases} 0, \text{si } u<1 \\ \log_{10}\text{u, si u} \in [1,10) \\ 1, \text{ si } u \geq 10 \end{cases}$.

3.2.1 Condición necesaria y suficiente para que una variable aleatoria siga la Ley de Benford

<u>Proposición 3.2.1.1</u> Sea X una variable aleatoria absolutamente continua sobre [1,10). La variable aleatoria X sigue la Ley de Benford si, y solo si, la variable aleatoria $Z = \log_{10} X$ tiene una distribución uniforme en [0,1).

Demostración.

Sean G y F las funciones de distribución de las variables aleatorias Z y X, respectivamente. Entonces, $G(z) = P(0 \leq Z < z)$ y $F(x) = P(1 \leq X < x)$.

Entonces, $G(z) = P(0 \leq Z < z) = P(0 \leq \log_{10} X < z) = P(1 \leq X < 10^z) = F(10^z)$.

Si X sigue la Ley de Benford, entonces,

$$G(z) = F(10^z) = \left\{\begin{cases} 0, \text{ si } z<0 \\ \log_{10}(10^z), \text{ si } z \in [0,1) \\ 1, \text{ si } z \geq 1 \end{cases}\right\} = \begin{cases} 0, \text{ si } z<0 \\ z, \text{ si } z \in [0,1) \\ 1, \text{ si } z \geq 1 \end{cases}.$$

Por lo tanto, Z tiene distribución uniforme en [0,1).

Si Z tiene distribución uniforme en [0,1), entonces, $G(z) = \begin{cases} 0, \text{ si } z<0 \\ z, \text{ si } z \in [0,1) \\ 1, \text{ si } z \geq 1 \end{cases}$.

Como $G(z)=F\left(10^{z}\right)$. Se deduce que $F(x)=\begin{cases}0, \text{ si } x<1\\ \log_{10} x, \text{ si } x\in[1,10)\\ 1, \text{ si } x\geq 10\end{cases}$. Por lo cual, X sigue la Ley de Benford. □.

3.2.2 La Ley de Benford es invariante frente a cambios de escala

Proposición 3.2.2.1 Sea X una variable aleatoria en [1,10) que sigue la Ley de Benford. Sea c un número real positivo y sea $Y=cX$. Entonces, la mantisa M de la variable aleatoria Y tiene la misma distribución que la de X.

Demostración.

Sea c un real positivo, se puede escribir $c=m.10^{r}$, donde $m\in[1,10)$ es la mantisa de c.

Las mantisas de cX y mX coinciden. Por tanto basta analizar el caso $c\in[1,10)$.

Como X es una variable aleatoria que sigue la Ley de Benford su función de distribución es $F:\mathbb{R}\to\mathbb{R}, F(u)=\begin{cases}0, \text{si } u<1\\ \log_{10}\text{u, si u}\in[1,10)\\ 1, \text{ si } u\geq 10\end{cases}$.

Sea G la función de distribución de la variable aleatoria M. Se probará que G es tal que $G(x)=\begin{cases}0, \text{ si } x<1\\ \log_{10} x, \text{ si } x\in[1,10)\\ 1, \text{ si } x\geq 10\end{cases}$.

Si $x<1$, $G(x)=P(M\leq x)=0$, puesto que la mantisa de un número sea menor que 1 es el suceso imposible.

Si $x\geq 10$, $G(x)=P(M\leq x)=1$, ya que $\{M\leq x\}$ es el suceso seguro, puesto que la mantisa de un número positivo es un real perteneciente a [1,10).

Si $x\in[1,10)$, se calculará $G(x)=P(M\leq x)$.

Como M es la mantisa de $Y=cX$, y cX toma valores en [c,10c), se tiene que $M = cX$ si $cX < 10$ y $M = \frac{cX}{10}$ si $cX \geq 10$.

Si $x \in [1, c)$, la única posibilidad de que la mantisa de cX pertenezca a [1,c) es que $cX \in [10, 10c)$. Entonces, la mantisa de cX, es igual a $\frac{cX}{10}$.

Por lo tanto, $P(M \leq x) = P\left(1 \leq \frac{cX}{10} \leq x\right) = P\left(\frac{10}{c} \leq X \leq \frac{10x}{c}\right) = F\left(\frac{10x}{c}\right) - F\left(\frac{10}{c}\right) =$
$= \log_{10}\left(\frac{10x}{c}\right) - \log_{10}\left(\frac{10}{c}\right) = \log_{10} x + \log_{10}\left(\frac{10}{c}\right) - \log_{10}\left(\frac{10}{c}\right) = \log_{10} x.$

Si $x \in [c, 10)$, $P(M \leq x) = P(M \leq c) + P(c < M \leq x)$.

Calculemos $P(M \leq c)$:

Como $P(M \leq c) = P(1 \leq M \leq c)$, análogamente a lo realizado para el caso anterior la única posibilidad de que $M \in [1, c)$ es que $cX \in [10, 10c)$. Entonces, $M = \frac{cX}{10}$.

$P(M \leq c) = P(1 \leq M \leq c) = P\left(1 \leq \frac{cX}{10} \leq c\right) = P\left(\frac{10}{c} \leq X \leq 10\right) = \log_{10} 10 - \log_{10}\left(\frac{10}{c}\right) =$
$= \log_{10} c.$

Calculemos $P(c < M \leq x)$:

Para que la mantisa de cX pertenezca a [c,10), como cX toma valores en [c,10c) y $c \geq 1$, entonces, $cX < 10$ y $M = cX$.

$P(c < M \leq x) = P(c < cX \leq x) = P\left(1 < X \leq \frac{x}{c}\right) = \log_{10}\left(\frac{x}{c}\right) - \log_{10} 1 = \log_{10} x - \log_{10} c.$

Por lo tanto, $G(x) = P(M \leq x) = P(M \leq c) + P(c < M \leq x) = \log_{10} x. \square$.

3.2.3 La Ley de Benford es la única ley de probabilidad invariante frente a cambios de escala

<u>Proposición 3.2.3.1</u> Sea X una variable aleatoria absolutamente continua que toma valores en [1,10). Si X es invariante frente a cambios de escala, entonces X sigue la Ley de Benford.

Demostración.

Sea F la función de distribución de X. Entonces, $F(x) = P(X \le x) = P(1 \le X \le x)$.

Si $x<1$, $F(x) = P(1 \le X \le x) = 0$, puesto que $\{1 \le X \le x\}$ es el suceso imposible.

Si $x \ge 10$, $F(x) = P(1 \le X \le x) = 1$, ya que $\{1 \le X \le x\}$ es el suceso seguro.

Si $x \in [1,10)$, $\{1 \le X \le x\}$ y $\{c \le cX \le cx\}$ son el mismo suceso para $c \in \mathbb{R}^+$. Por lo tanto, $F(x) = P(1 \le X \le x) = P(c \le cX \le cx)$.

Podemos elegir $c \in [1,10)$ de modo que $cx<10$ (c depende de x). Así, para $c \le cX \le cx$, cX es igual a su mantisa. Como X es invariante frente a cambios de escala, la mantisa de cX tiene la misma distribución que X. Entonces, $P(X \le x) = P(c \le cX \le cx) = F(cx) - F(c)$. Por tanto, $F(x) = F(cx) - F(c)$.

Si $c = 1+\varepsilon$, $F(x) = F(x(1+\varepsilon)) - F(1+\varepsilon)$ y como $F(1)=0$.

$\dfrac{F(x(1+\varepsilon)) - F(x)}{x\varepsilon} = \dfrac{F(1+\varepsilon) - F(1)}{x\varepsilon}$. Si $\varepsilon \to 0$, se deduce que $F'(x) = \dfrac{F'(1)}{x}$.

Como $F(1)=0$, $F(x) = F'(1).\ln x$.

Como $F(10)=1$, $F(10) = F'(1).\ln 10 \quad \Rightarrow \quad F'(1) = \dfrac{1}{\ln 10}$.

Por lo tanto, $F(x) = \dfrac{1}{\ln 10} \ln x = \log_{10} x$.

Concluimos así que X sigue la Ley de Benford. □.

Capítulo 4. Ley general de probabilidad para los primeros k dígitos significativos de un número en base 10

Hill [2] propone una Ley general de probabilidad para los primeros k dígitos significativos de un número. Análogamente a como se definió la mantisa de un real positivo pueden definirse las funciones $D_1 : \mathbb{R}^+ \to \{1,2,3,4,5,6,7,8,9\}$ y $D_i : \mathbb{R}^+ \to \{0,1,2,3,4,5,6,7,8,9\}$ para $i \in \mathbb{N},\ i \geq 2$ tal que D_i hace corresponder a un real positivo su i-ésimo dígito significativo en base 10. Se observa que D_1 coincide con la parte entera de la función mantisa.

<u>Definición 4.1 Ley general para los primeros k dígitos significativos de un número real positivo en base 10</u>

Dado $k \in \mathbb{Z}^+$, para todo d_1 con $d_1 \in \{1,2,3,4,5,6,7,8,9\}$,
para todos d_j con $d_j \in \{0,1,2,3,4,5,6,7,8,9\}$, $j = 2...k$

$$P\left(\bigcap_{i=1}^{k}\{D_i = d_i\}\right) = \log_{10}\left(1 + \left(\sum_{i=1}^{k} d_i.10^{k-i}\right)^{-1}\right).$$

Esta Ley incluye como caso particular a la Ley de Benford para el primer dígito significativo.

<u>Proposición 4.1</u> Sea $n \in \mathbb{N}, n \geq 2$ y siendo D_n el enésimo dígito significativo de un real positivo. Entonces, $P(\{D_n = d\}) = \sum_{j=10^{n-2}}^{10^{n-1}-1} \log_{10}\left(1 + \frac{1}{10j + d}\right)$ con $d = 0...9$.

Demostración.[4]

Dados $d \in \{0,1,2,3,4,5,6,7,8,9\}$ y $n \in \mathbb{N},\ n \geq 2$

$$P(\{D_n = d\}) = P\left(\bigcup_{i_1=1}^{9}\{D_1 = i_1\} \cap \bigcup_{i_2=0}^{9}\{D_2 = i_2\} \cap ... \cap \bigcup_{i_{n-1}=0}^{9}\{D_{n-1} = i_{n-1}\} \cap \{D_n = d\}\right) =$$

[4] La notación $i_1...i_{n-1}d$ simboliza al real $i_1.10^n + i_2.10^{n-1} + i_3.10^{n-2} + ... + i_{n-1}.10^1 + d$. Es decir, $i_1,...,i_{n-1},d$ son los dígitos de ese número real. Esta notación se utilizará en otras oportunidades a lo largo de este trabajo.

$$P\left(\bigcup_{i_1=1,i_2=0,\ldots,i_{n-1}=0}^{9}\left\{D_1=i_1,D_2=i_2,\ldots,D_{n-1}=i_{n-1},D_n=d\right\}\right)=$$

$$\sum_{i_1=1,i_2=0,\ldots,i_{n-1}=0}^{9}P\left(\left\{D_1=i_1,D_2=i_2,\ldots,D_{n-1}=i_{n-1},D_n=d\right\}\right)$$

$$\sum_{i_1=1,i_2=0,\ldots,i_{n-1}=0}^{9}\log_{10}\left(1+\frac{1}{i_1\ldots i_{n-1}d}\right)=\sum_{j=10^{n-2}}^{10^{n-1}-1}\log_{10}\left(1+\frac{1}{10j+d}\right).\ \square.$$

Un corolario que se puede deducir de la Ley general de probabilidad para los primeros k dígitos significativos es que existe una dependencia entre los dígitos significativos de un real positivo. Por ejemplo, la probabilidad de que el segundo dígito significativo sea 2, es diferente a la probabilidad de que el segundo dígito significativo sea 2 sabiendo que el primer dígito significativo es 1, y es a su vez diferente a la probabilidad de que el segundo dígito significativo sea 2 sabiendo que el primer dígito significativo es 2.

$$P(D_2=2)=\sum_{j=1}^{9}\log_{10}\left(1+\left(10j+2\right)^{-1}\right)\cong 0{,}109\,.$$

$$P\left(D_1=1,D_2=2\,/\,D_1=1\right)=\frac{\log_{10}\left(1+\left(10+2\right)^{-1}\right)}{\log_{10}\left(1+\frac{1}{1}\right)}\cong 0{,}115\,.$$

$$P\left(D_1=2,D_2=2\,/\,D_1=2\right)=\frac{\log_{10}\left(1+\left(10.2+2\right)^{-1}\right)}{\log_{10}\left(1+\frac{1}{2}\right)}\cong 0{,}110\,.$$

Esta dependencia se hace más débil a medida que crece la distancia entre los dígitos. Calculemos la probabilidad de que un real positivo tenga primer dígito 1 y enésimo dígito 1, y veamos que ésta decrece.

Proposición 4.2 La sucesión $\left(a_n\right)$: $a_n=P\left(\left\{D_1=1,D_n=1\right\}\right)$ es decreciente.

Demostración.

$$P\left(\left\{D_1=1,D_n=1\right\}\right)=P\left(D_1=1\cap\bigcup_{i_2=0}^{9}\left\{D_2=i_2\right\}\cap\ldots\cap\bigcup_{i_{n-1}=0}^{9}\left\{D_{n-1}=i_{n-1}\right\}\cap\left\{D_n=1\right\}\right)=$$

$$P\left(\bigcup_{i_2=0,\ldots i_{n-1}=0}^{9}\left\{D_1=1,D_2=i_2,\ldots,D_{n-1}=i_{n-1},D_n=1\right\}\right)=$$

$$=\sum_{i_2=0,\ldots i_{n-1}=0}^{9}P\left(\left\{D_1=1,D_2=i_2,\ldots,D_{n-1}=i_{n-1},D_n=1\right\}\right)=$$

$$=\sum_{i_2=0,\ldots i_{n-1}=0}^{9}\log_{10}\left(1+\frac{1}{1i_2\ldots i_{n-1}1}\right)=\log_{10}\left(\prod_{i_2=0,\ldots i_{n-1}=0}^{9}\left(1+\frac{1}{1i_2\ldots i_{n-1}1}\right)\right).$$

Veamos que (a_n) es decreciente.

$$a_{n+1}-a_n=P\left(\left\{D_1=1,D_{n+1}=1\right\}\right)-P\left(\left\{D_1=1,D_n=1\right\}\right)=$$

$$\log_{10}\left(\prod_{i_2=0,\ldots i_n=0}^{9}\left(1+\frac{1}{1i_2\ldots i_n1}\right)\right)-\log_{10}\left(\prod_{i_2=0,\ldots i_{n-1}=0}^{9}\left(1+\frac{1}{1i_2\ldots i_{n-1}1}\right)\right).$$

Entonces,

$$\forall n,\ n\in\mathbb{N},\ a_{n+1}-a_n\le 0\ \Leftrightarrow\ \prod_{i_2=0,\ldots i_n=0}^{9}\left(1+\frac{1}{1i_2\ldots i_n1}\right)\le\prod_{i_2=0,\ldots i_{n-1}=0}^{9}\left(1+\frac{1}{1i_2\ldots i_{n-1}1}\right)$$

Si $n=2$ se debe probar que $\left(1+\frac{1}{101}\right)\left(1+\frac{1}{111}\right)\left(1+\frac{1}{121}\right)\ldots\left(1+\frac{1}{191}\right)\le\left(1+\frac{1}{11}\right)$

$$\frac{102}{101}\cdot\frac{112}{111}\cdot\frac{122}{123}\ldots\frac{192}{191}\le\frac{12}{11}=\frac{120}{110}=\frac{111}{110}\cdot\frac{112}{111}\cdot\frac{113}{112}\ldots\frac{120}{119}$$

Se observa que el segundo factor en ambos productos coincide y que se verifica que: $\frac{192}{191}<\frac{182}{181}<\ldots<\frac{132}{131}<\frac{122}{121}<\frac{120}{119}<\frac{119}{118}<\ldots<\frac{113}{112}$ .Por lo tanto, para probar la desigualdad alcanza con comprobar que $\frac{102}{101}\cdot\frac{192}{191}\le\frac{111}{110}\cdot\frac{113}{112}$ lo que es cierto.

Si $n=3$ se debe probar que $\prod_{i_2=0,i_3=0}^{9}\left(1+\frac{1}{1i_2i_31}\right)\le\prod_{i_2=0}^{9}\left(1+\frac{1}{1i_21}\right)$. Para probarlo es suficiente probar las siguientes desigualdades:

$$\left(1+\frac{1}{1001}\right)\left(1+\frac{1}{1011}\right)\left(1+\frac{1}{1021}\right)...\left(1+\frac{1}{1091}\right)\leq\left(1+\frac{1}{101}\right)$$
$$\left(1+\frac{1}{1101}\right)\left(1+\frac{1}{1111}\right)\left(1+\frac{1}{1121}\right)...\left(1+\frac{1}{1191}\right)\leq\left(1+\frac{1}{111}\right)$$
....
$$\left(1+\frac{1}{1901}\right)\left(1+\frac{1}{1911}\right)\left(1+\frac{1}{1921}\right)...\left(1+\frac{1}{1991}\right)\leq\left(1+\frac{1}{191}\right)$$

Analizando la primera de ellas y trabajando análogamente al caso anterior: $\frac{1002}{1001}.\frac{1012}{1011}.\frac{1022}{1023}...\frac{1092}{1091}\leq\frac{102}{101}=\frac{1020}{1010}=\frac{1011}{1010}.\frac{1012}{1011}.\frac{1013}{1012}...\frac{1020}{1019}$

Se observa que el segundo factor en ambos productos coincide y que se verifica que: $\frac{1092}{1091}<\frac{1082}{1081}<...<\frac{1032}{1031}<\frac{1022}{1021}<\frac{1020}{1019}<\frac{1019}{1018}<...<\frac{1013}{1012}$ .Por lo tanto, para probar la desigualdad alcanza con comprobar que $\frac{1002}{1001}.\frac{1092}{1091}\leq\frac{1011}{1010}.\frac{1013}{1012}$.

Este razonamiento puede realizarse con las demás desigualdades y se obtendrá que es suficiente comprobar una desigualdad de la forma: $\frac{m-9}{m-10}.\frac{m+81}{m+80}\leq\frac{m}{m-1}.\frac{m+2}{m+1}$, siendo m el numerador del primer factor de la derecha de la desigualdad.

La que es equivalente a $\frac{-70m^2+1528m+729}{(m-10)(m+80)(m-1)(m+1)}\leq 0$ que tiene solución $(-\infty,-80]\cup(-1,\alpha]\cup(1,10)\cup[\beta,+\infty)$ siendo α y β las raíces de $-70m^2+1528m+729=0$. Como la parte entera de β es 22 y el menor m posible es 1011, se concluye lo que se quería probar.

En general se trabajará análogamente. Para demostrar que $\prod_{i_2=0,...i_n=0}^{9}\left(1+\frac{1}{1i_2...i_n1}\right)\leq\prod_{i_2=0,...i_{n-1}=0}^{9}\left(1+\frac{1}{1i_2...i_{n-1}1}\right)$ es suficiente probar las siguientes desigualdades:

$$\left(1+\frac{1}{10...001}\right)\left(1+\frac{1}{10...011}\right)\left(1+\frac{1}{10...021}\right)...\left(1+\frac{1}{10...091}\right)\leq\left(1+\frac{1}{10...01}\right)$$

$$\left(1+\frac{1}{10...0101}\right)\left(1+\frac{1}{10...0111}\right)\left(1+\frac{1}{10...0121}\right)...\left(1+\frac{1}{10...0191}\right)\leq\left(1+\frac{1}{10...11}\right)$$

....

$$\left(1+\frac{1}{19...9901}\right)\left(1+\frac{1}{19...9911}\right)\left(1+\frac{1}{19...9921}\right)...\left(1+\frac{1}{19...9991}\right)\leq\left(1+\frac{1}{19...91}\right)$$

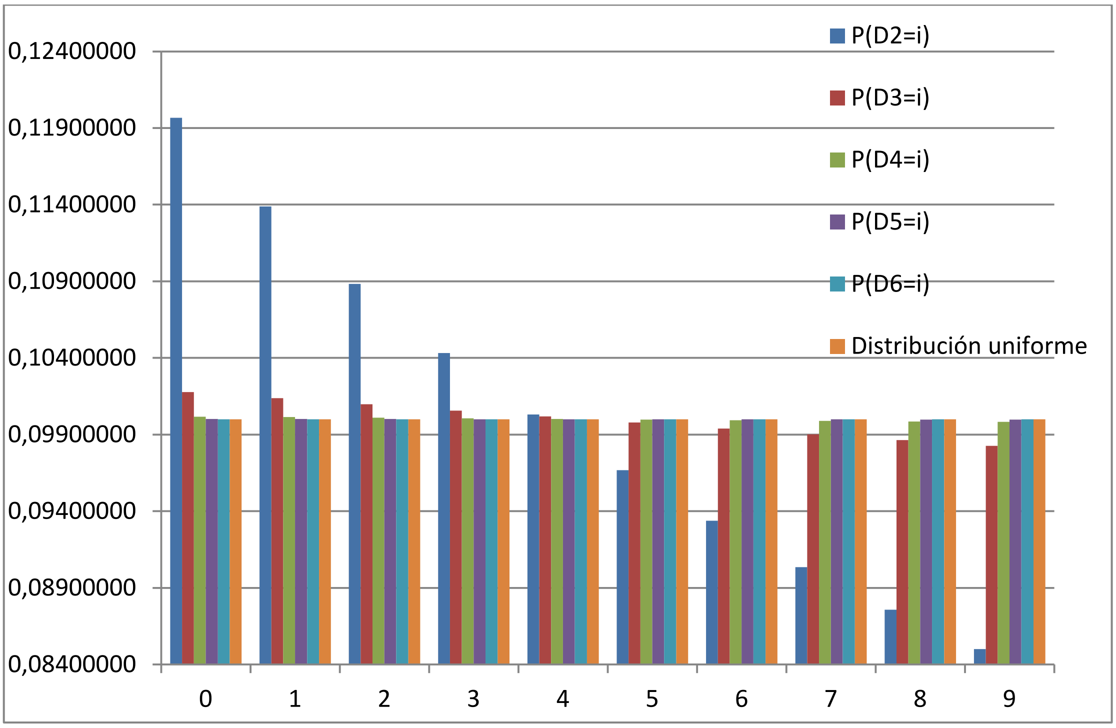
Trabajando del mismo modo que en los casos anteriores se ve que es suficiente probar que se verifica una desigualdad de la forma: $\frac{m-9}{m-10}.\frac{m+81}{m+80}\leq\frac{m}{m-1}.\frac{m+2}{m+1}$, como en todas las desigualdades el menor m posible es mayor a 22, se concluye lo que se quería demostrar. □

Otras implicancias son que $P\left(\{D_1=i,D_n=j\}\right)$ tiende a $P\left(\{D_n=j\}\right)$, cuando $n\to+\infty$, para $i\in\{1,2,3,4,5,6,7,8,9\}$, $j\in\{0,1,2,3,4,5,6,7,8,9\}$ y que la distribución del enésimo dígito tiende a la distribución uniforme a medida que n crece. Se ofrecerá al lector una prueba de la última afirmación (para el caso *d*=0) que fuera enunciada por Hill en varias oportunidades pero cuya demostración no pudo ser encontrada en la bibliografía. Antes de eso se visualizará la situación para *n* entre 2 y 6.

En la siguiente tabla se muestran las probabilidades $P\left(D_n=i\right)$ con $n=2...6$ e $i=0...9$ de acuerdo a la Ley general de probabilidad para los primeros k dígitos significativos de un número. En la misma podemos visualizar que si bien para los primeros tres dígitos la diferencia con la distribución uniforme es significativa, deja de serlo cuando *n* es mayor o igual a 4. Nigrini ([9], pág.6) plantea que en conjuntos de datos con tres o más cifras significativas, a partir del cuarto dígito significativo el error con la distribución uniforme es despreciable a fines prácticos.

	Posición del dígito significativo					
Dígito	Primero	Segundo	Tercero	Cuarto	Quinto	Sexto
0		0,11967927	0,10178436	0,10017615	0,10001759	0,10000176
1	0,30103000	0,11389010	0,10137598	0,10013689	0,10001368	0,10000137
2	0,17609126	0,10882150	0,10097220	0,10009767	0,10000977	0,10000098
3	0,12493874	0,10432956	0,10057293	0,10005850	0,10000586	0,10000059
4	0,09691001	0,10030820	0,10017809	0,10001937	0,10000195	0,10000020
5	0,07918125	0,09667724	0,09978758	0,09998029	0,09999804	0,09999981
6	0,06694679	0,09337474	0,09940131	0,09994124	0,09999414	0,09999941
7	0,05799195	0,09035199	0,09901921	0,09990224	0,09999023	0,09999902
8	0,05115252	0,08757005	0,09864118	0,09986328	0,09998632	0,09999863
9	0,04575749	0,08499735	0,09826716	0,09982437	0,09998241	0,09999824

A continuación, aparece un gráfico en el que se muestran las probabilidades $P\left(D_n = i\right)$ *con* $n = 2...6$ e $i = 0...9$ que fueron antes tabuladas.



<u>Proposición 4.3</u> Dado $n \in \mathbb{N},\ n \geq 2$. Entonces, $P\left(\left\{D_n = 0\right\}\right) \to \frac{1}{10},\ n \to +\infty$.

Demostración.

$$P\left(D_n = 0\right) = \sum_{j=10^{n-2}}^{10^{n-1}-1} \log_{10}\left(1 + \frac{1}{10j}\right) = \frac{1}{\log_e 10} \cdot \sum_{j=10^{n-2}}^{10^{n-1}-1} \log_e\left(1 + \frac{1}{10j}\right).$$

Sea $\mu = \dfrac{1}{\log_e 10}$

Veamos que $\displaystyle\sum_{j=10^{n-2}}^{10^{n-1}-1} \log_e\left(1 + \frac{1}{10j}\right) \sim \sum_{j=10^{n-2}}^{10^{n-1}-1} \frac{1}{10j}$, para $n \to +\infty$.

Como $\quad \log_e\left(1 + \dfrac{1}{10j}\right) = \dfrac{1}{10j} + \dfrac{1}{2}.\left(\dfrac{1}{10j}\right)^2 + R(j)$, con $R(j) \to 0$, si $j \to +\infty$,

entonces, $\displaystyle\left[\sum_{j=10^{n-2}}^{10^{n-1}-1} \log_e\left(1 + \frac{1}{10j}\right) - \sum_{j=10^{n-2}}^{10^{n-1}-1} \frac{1}{10j}\right] \sim \sum_{j=10^{n-2}}^{10^{n-1}-1} \frac{1}{2}\left(\frac{1}{10j}\right)^2 = \frac{1}{200}.\sum_{j=10^{n-2}}^{10^{n-1}-1} \frac{1}{j^2}$.

$\frac{1}{200}.\sum_{j=10^{n-2}}^{10^{n-1}-1}\frac{1}{j^2}\leq\frac{1}{200}.\sum_{j=10^{n-2}}^{+\infty}\frac{1}{j^2}$ y $\frac{1}{200}.\sum_{j=10^{n-2}}^{+\infty}\frac{1}{j^2}\to 0$ para $n\to+\infty$ por ser la cola de una serie convergente. Por lo tanto, $\left[\sum_{j=10^{n-2}}^{10^{n-1}-1}\log_e\left(1+\frac{1}{10j}\right)-\sum_{j=10^{n-2}}^{10^{n-1}-1}\frac{1}{10j}\right]\to 0$ para $n\to+\infty$.

Entonces, $P(D_n=0)\sim\mu.\sum_{j=10^{n-2}}^{10^{n-1}-1}\frac{1}{10j}=\frac{\mu}{10}.\sum_{j=10^{n-2}}^{10^{n-1}-1}\frac{1}{j}$, para $n\to+\infty$.

Como $\log_e m\ \sim\ \sum_{j=1}^{m}\frac{1}{j}$, para $m\to+\infty$.

$$P(D_n=0)\sim\frac{\mu}{10}.\left[\sum_{j=1}^{10^{n-1}-1}\frac{1}{j}-\sum_{j=1}^{10^{n-2}-1}\frac{1}{j}\right]\sim\frac{\mu}{10}.\left[\log_e\left(10^{n-1}-1\right)-\log_e\left(10^{n-2}-1\right)\right]\sim$$

$$\sim\frac{\mu}{10}.\left[\log_e\frac{10^{n-1}-1}{10^{n-2}-1}\right]\sim\frac{\mu}{10}.\left[\log_e\frac{10^{n-1}}{10^{n-2}}\right]=\frac{\mu}{10}.\left[\log_e 10\right]=\frac{1}{10}.\square.$$

Capítulo 5. Primeros dos dígitos

5.1 Ley de Benford para los primeros dos dígitos

Representando el primer dígito significativo con D_1, el segundo dígito con D_2 y con D_1D_2 los primeros dos de un entero positivo, la Ley de Benford para los primeros dos dígitos podría enunciarse como $P\left(D_1D_2 = d_1d_2\right) = \log_{10}\left(1+\frac{1}{d_1d_2}\right)$ con $d_1 = 1\ldots9$ y $d_2 = 0\ldots9$.

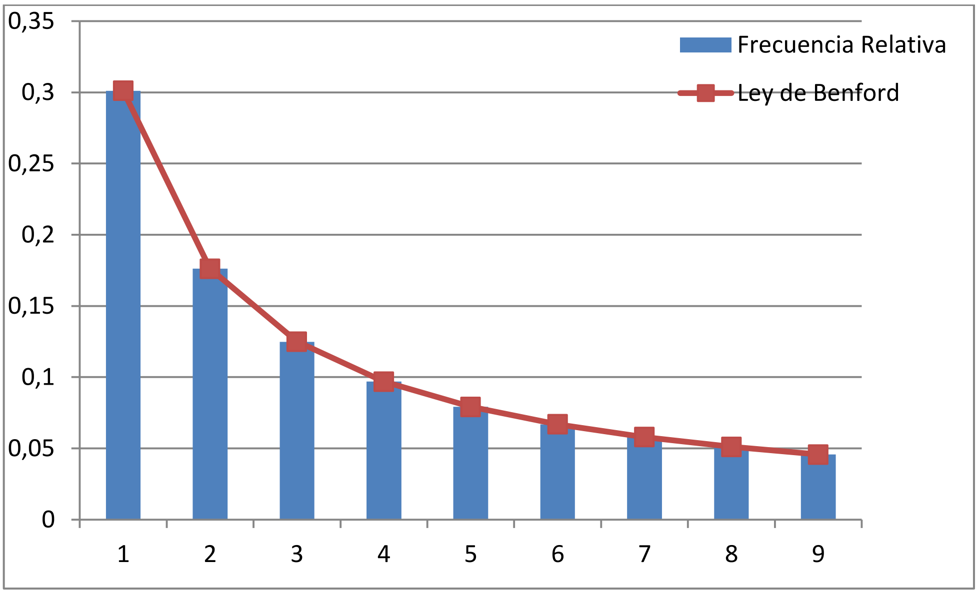
Se observa que el segundo dígito significativo a diferencia del primero puede tomar el valor 0.

Nigrini ([9], pág.15-19) bajo el título "Love at first sight" explica porqué usar solo el primer dígito a veces no es suficiente al analizar un conjunto de datos de dos o más dígitos. Allí, construye un contraejemplo que se sintetizará a continuación.

Número	Primer dígito	Frecuencia	Frecuencia Relativa	Probabilidad según Benford para el primer dígito
1,0	1	3011	0,3011	0,301029996
2,0	2	1761	0,1761	0,176091259
3,0	3	1249	0,1249	0,124938737
4,0	4	969	0,0969	0,096910013
5,0	5	792	0,0792	0,079181246
6,0	6	669	0,0669	0,06694679
7,0	7	580	0,058	0,057991947
8,0	8	512	0,0512	0,051152522
9,0	9	457	0,0457	0,045757491

En la tabla anterior se muestra un conjunto de 10.000 datos, los mismos toman exactamente nueve valores diferentes. Se muestran, también, la frecuencia con la que aparecen, la frecuencia relativa y la probabilidad según la Ley de Benford para el primer dígito.

En el siguiente gráfico se representan las frecuencias relativas de aparición de los datos y las probabilidades dadas por la Ley de Benford. En una primera instancia, podría concluirse que los datos están en tal proporción de modo que conforman la Ley de Benford. Esta última afirmación es no es cierta como veremos a continuación.

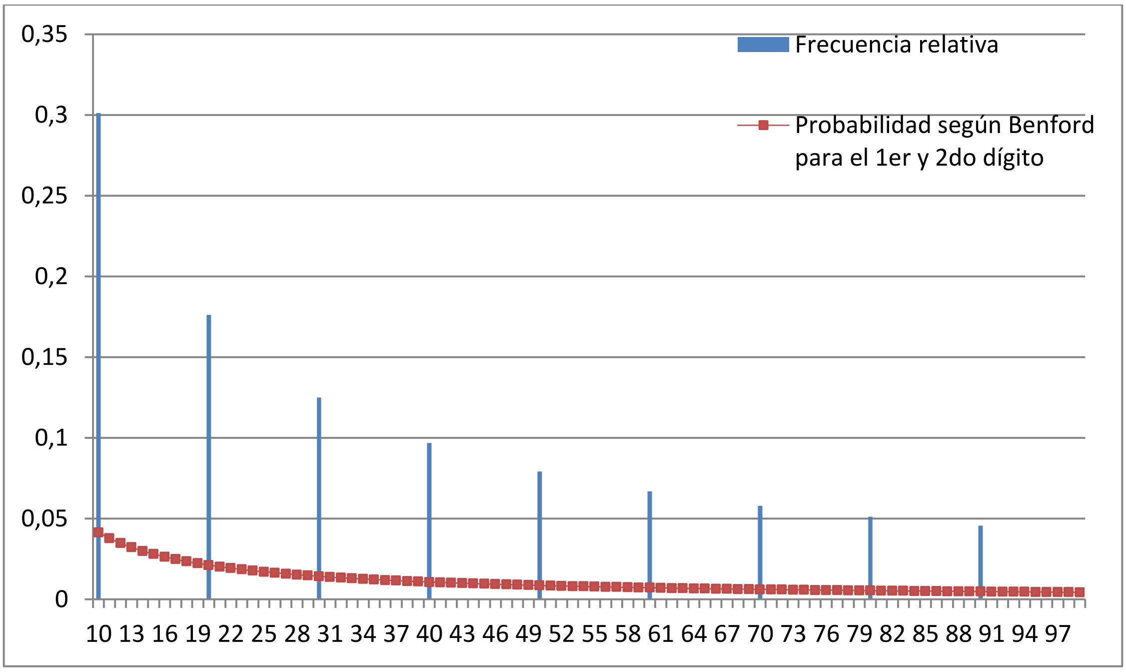


Otra forma de analizar el conjunto de datos es analizar la distribución de las mantisas de los logaritmos de los datos. De seguir la Ley de Benford, las mismas debieran estar uniformemente distribuidas, lo cual no sucede, como se visualiza en la siguiente tabla. En conclusión, el análisis con el primer dígito no es suficiente.

Número	Primer dígito	Frecuencia	Frecuencia relativa	Mantisa del logaritmo del dato
1,0	1	3011	0,3011	0,0000000
2,0	2	1761	0,1761	0,3010300
3,0	3	1249	0,1249	0,4771213
4,0	4	969	0,0969	0,6020600
5,0	5	792	0,0792	0,6989700
6,0	6	669	0,0669	0,7781513
7,0	7	580	0,058	0,8450980
8,0	8	512	0,0512	0,9030900
9,0	9	457	0,0457	0,9542425

Otra manera de demostrar que el análisis con el primer dígito no es suficiente es comparar la distribución con la Ley de Benford para los primeros dos dígitos. En la siguiente tabla se muestra en la quinta columna la probabilidad que según la Ley de Benford (para los primeros dos dígitos) tendría cada uno de los datos. Se puede observar que difiere considerablemente de la frecuencia relativa de los mismos, y esto se puede visualizar en el gráfico que aparece a continuación.

Número	Primer dígito	Frecuencia	Frecuencia relativa	Probabilidad según Benford para el 1er y 2do dígito
1,0	1	3011	0,3011	0,041392685
2,0	2	1761	0,1761	0,021189299
3,0	3	1249	0,1249	0,010995384
4,0	4	969	0,0969	0,010723865
5,0	5	792	0,0792	0,008600172
6,0	6	669	0,0669	0,007178585
7,0	7	580	0,058	0,006160309
8,0	8	512	0,0512	0,005395032
9,0	9	457	0,0457	0,004798883



En suma, si bien los datos considerados en primer análisis mostrarían gráficamente una amplia conformidad respecto a la Ley de Benford para el primer dígito, luego de un análisis más profundo se puede probar que no lo hacen. Más aún, al considerar los primeros dos dígitos se evidencia la no conformidad con la Ley.

Capítulo 6. ¿Cómo comprobar si un conjunto de datos sigue la Ley de Benford?

Nigrini ([9], pág. 20) afirma que a través de simulaciones ha comprobado que la conformidad con la ley de Benford requiere de grandes conjuntos de datos (de por lo menos 1000) que tengan por lo menos cuatro dígitos significativos. Sin embargo, si en el conjunto de datos aparecen algunos, pero no demasiados números, con dos o tres dígitos, la distribución, como por ejemplo en los datos de un censo aún se ajusta a la Ley de Benford.

Nigrini ([9], pág. 20) sugiere no comparar una distribución con la Ley de Benford para los primeros dos dígitos en conjuntos de menos de trescientos datos. Para esos casos sugiere utilizar la Ley de Benford para el primer dígito.

6.1 Chi Cuadrado

El test de bondad de ajuste de Chi Cuadrado χ^2 permite comprobar si ciertos datos siguen una cierta distribución de probabilidad con un cierto error α. Para el caso de la Ley de Benford para el primer dígito, tendremos 8 grados de libertad y si buscamos hacer el test con un error del 5%, entonces se aceptará que los datos siguen la Ley de Benford si $\chi^2 < 15,51$ y se rechazará en otro caso.

Para el caso de la Ley de Benford para los primeros dos dígitos, tendremos 89 grados de libertad y si buscamos hacer el test con un error del 5%, entonces se aceptará que los datos siguen la Ley de Benford si $\chi^2 < 112,02$ y se rechazará en otro caso.

El estadístico Chi Cuadrado depende de la cantidad de datos. Nigrini ([9], pág 154) afirma que el test de Chi Cuadrado sufre de un “exceso de poder” cuando la cantidad de datos es muy grande. Es decir, que casi siempre χ^2 será mayor que su valor crítico, lo que implicaría que los datos no conforman la Ley de Benford aunque las diferencias sean mínimas. Afirma que puede empezar a notarse este problema cuando los conjuntos tienen más de 5000 datos y que con 25000 o más datos se requeriría casi la perfección para que pasaran el test de bondad de ajuste de Chi Cuadrado.

6.2 Mean Absolute Deviation Test (MAD Test)

Nigrini ([9], pág. 158) afirma que para desentenderse del problema de la cantidad de datos (que en las auditorías de fraude fiscal suelen ser muchos) se debe utilizar un estadístico que no dependa de ese parámetro.

Propone utilizar el Mean Absolute Deviation Test (MAD Test), definiendo el estadístico $MAD = \frac{\sum_{j=1}^{K} |o_j - p_j|}{K}$, siendo K la cantidad de categorías que toma la variable, o_j la frecuencia relativa observada y p_j la probabilidad esperada para la distribución de probabilidad. Cuanto mayor es MAD, mayor es la diferencia promedio entre la frecuencia relativa observada y la probabilidad. En el caso de comparar una distribución de datos con la Ley de Benford para el primer dígito K tomará el valor 9, con la Ley de Benford para el primer y segundo dígito K será 90.

A diferencia del Test de Chi Cuadrado no hay una grilla con valores críticos objetivos según un cierto error u otros parámetros. En Nigrini ([9], pág. 160) aparece una tabla con ciertos valores críticos para decidir la conformidad o no conformidad de un conjunto de datos con la Ley de Benford. La misma fue desarrollada en base a las experiencias empíricas de Nigrini en el trabajo diario con este tópico en diferentes áreas.

TABLE 7.1 Critical Values and Conclusions for Various MAD Values

Digits	Range	Conclusion
First Digits	0.000 to 0.006	Close conformity
	0.006 to 0.012	Acceptable conformity
	0.012 to 0.015	Marginally acceptable conformity
	Above 0.015	Nonconformity
Second Digits	0.000 to 0.008	Close conformity
	0.008 to 0.010	Acceptable conformity
	0.010 to 0.012	Marginally acceptable conformity
	Above 0.012	Nonconformity
First-Two Digits	0.0000 to 0.0012	Close conformity
	0.0012 to 0.0018	Acceptable conformity
	0.0018 to 0.0022	Marginally acceptable conformity
	Above 0.0022	Nonconformity [5]

[5] Extraido de Nigrini, M. (2012). *Benford's Law Applications for Forensic Accounting, Auditing, and Fraud Detection* . Wiley. Pág.160.

Dicha tabla proporciona valores críticos para la comparación con la Ley de Benford para el primer dígito significativo, para el segundo dígito significativo, y para los primeros dos dígitos significativos. En algunas ocasiones las conclusiones pueden diferir según el Test MAD utilizado sea el del primer dígito, el segundo o el de los primeros dos dígitos. Allí, será necesario entonces, buscar explicaciones de acuerdo a las características particulares del conjunto de datos. Con respecto a esto, Nigrini sugiere una posición más bien pesimista en casos de resultados encontrados.

Nigrini ([9], pág.170) concluye que el test de Chi Cuadrado funcionará bien en conjuntos con una cantidad pequeña de datos. Para otros conjuntos con mayor cantidad de datos sugiere el uso del Test MAD siguiendo los parámetros dados por los valores críticos que aparecen en la tabla citada anteriormente.

6.3 Últimos dos dígitos

Como se demostró en la proposición 4.3 la frecuencia de los dígitos tiende a uniformizarse a medida que avanzamos hacia la derecha en la posición del dígito. Tomando esto como base, Nigrini ([9], pág.129) afirma para propósitos prácticos a partir del tercer dígito significativo la probabilidad de los dígitos es uniforme. Utilizando esto crea el Test de los últimos dos dígitos que servirá para detectar invención de números en determinadas situaciones.

¿Cuáles son los últimos dos dígitos de un número? En general, esta pregunta carece de sentido, pero lo que importa en el análisis forense de resultados es qué dígitos serían los apropiados para este análisis. Según Nigrini ([9], pág.129) para cuestiones que involucren dinero los centavos serían los apropiados candidatos a últimos dos dígitos, y que, para cuestiones en donde puede darse invención o fraude con números naturales las decenas y las unidades serán los apropiados.

El Test de los dos últimos dígitos es utilizado en datos donde se buscan signos de invención de datos. Por ejemplo, censos poblacionales, resultados electorales, inventarios, números en las deducciones de impuestos, pesos de pescados, lectura de temperaturas en estaciones meteorológicas, utilización de cupones de descuento.

Capítulo 7. Algunos ejemplos clásicos de conjuntos de datos que siguen la Ley de Benford

7.1 Sucesión de Fibonacci

La sucesión de Fibonacci puede definirse de forma recursiva para todo *n*, $n \in \mathbb{N}^*$, $f_1 = a,\ f_2 = b,\ f_{n+2} = f_n + f_{n+1}$, con *a* y *b* naturales cualesquiera. Siendo la sucesión de Fibonacci clásica la definida con *a*=*b*=*1*.

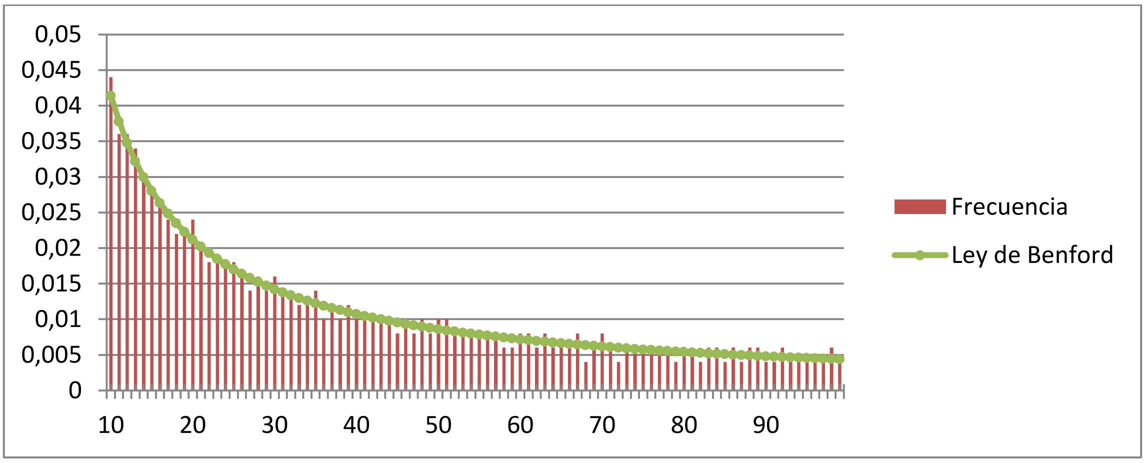
Al calcular los primeros 500 términos de la sucesión de Fibonacci clásica, y al estudiar la frecuencia del primer dígito significativo obtenemos los siguientes resultados.

Dígito	Frecuencia	Frecuencia relativa
1	151	0,302
2	88	0,176
3	63	0,126
4	47	0,094
5	40	0,080
6	33	0,066
7	29	0,058
8	27	0,054
9	22	0,044

Al aplicar el test de Bondad de ajuste de Chi-cuadrado, con 8 grados de libertad y $\alpha = 0,05$, se obtiene $\chi^2 = 0,173 < 15,51$. Por lo tanto, es aceptada la hipótesis de que este conjunto de datos, los primeros 500 términos de la sucesión de Fibonacci clásica, satisface la Ley de Benford para el primer dígito significativo.

Al aplicar el Test MAD para el primer dígito significativo obtenemos el valor 0,001 lo que describe una gran conformidad de los primeros 500 términos de la sucesión de Fibonacci clásica con la Ley de Benford para el primer dígito significativo.

El cardinal del conjunto de datos es 500 por lo que también es interesante estadísticamente estudiar la frecuencia de los primeros dos dígitos significativos. Los cuales aparecen representados en el siguiente gráfico.



Al aplicar el test de Bondad de ajuste de Chi-cuadrado, con 89 grados de libertad y $\alpha = 0,05$, se obtiene $\chi^2 = 5,69 < 112,02$. Por lo tanto, es aceptada la hipótesis de que este conjunto de datos, los primeros 500 términos de la sucesión de Fibonacci clásica, satisface la Ley de Benford para el primer y segundo dígito significativo.

Al aplicar el Test MAD para los primeros dos dígitos significativos obtenemos el valor 0,0009 lo que describe una gran conformidad de los primeros 500 términos de la sucesión de Fibonacci clásica con la Ley de Benford para el primer y segundo dígito significativo.

7.2 Números de Lucas

La sucesión de números de Lucas es un caso particular de la de Fibonacci con $a=2$ y $b=1$. La sucesión de números de Lucas aparece en el campo de la Teoría de Números.

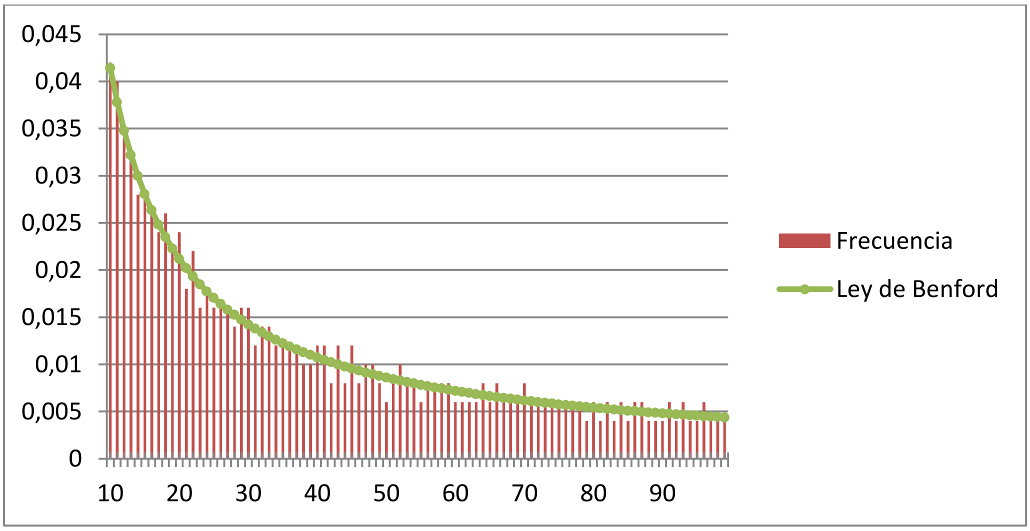
Al calcular los primeros 500 términos de la sucesión de números de Lucas, y estudiar la frecuencia del primer dígito significativo obtenemos los siguientes resultados.

Dígito	Frecuencia	Frecuencia relativa
1	151	0,302
2	88	0,176
3	62	0,124
4	50	0,100
5	39	0,078
6	32	0,064
7	30	0,060
8	25	0,050
9	23	0,046

Al aplicar el test de Bondad de ajuste de Chi-cuadrado, con 8 grados de libertad y $\alpha = 0,05$, se obtiene $\chi^2 = 0,176 < 15,51$. Por lo tanto, es aceptada la hipótesis de que este conjunto de datos, los primeros 500 términos de la sucesión de Números de Lucas, satisface la Ley de Benford para el primer dígito significativo.

Al aplicar el Test MAD para el primer dígito significativo obtenemos el valor 0,001 lo que describe una gran conformidad de los primeros 500 términos de la sucesión de Números de Lucas con la Ley de Benford para el primer dígito significativo.

El cardinal del conjunto de datos es 500 por lo que también es interesante estadísticamente estudiar la frecuencia de los primeros dos dígitos significativos. Los cuales aparecen representados en el siguiente gráfico.



Al aplicar el test de Bondad de ajuste de Chi-cuadrado, con 89 grados de libertad y $\alpha = 0,05$, se obtiene $\chi^2 = 6,69 < 112,02$. Por lo tanto, es aceptada la hipótesis de que este conjunto de datos, los primeros 500 términos de la sucesión de Números de Lucas, satisface la Ley de Benford para el primer y segundo dígito significativo.

Al aplicar el Test MAD para los primeros dos dígitos significativos obtenemos el valor 0,001 lo que describe una gran conformidad de los primeros 500 términos de la sucesión de Números de Lucas con la Ley de Benford para el primer y segundo dígito significativo.

7.3 Potencias de 2

La sucesión de las primeras 100 potencias de 2 también es un ejemplo de conjunto de datos que satisfacen la Ley de Benford

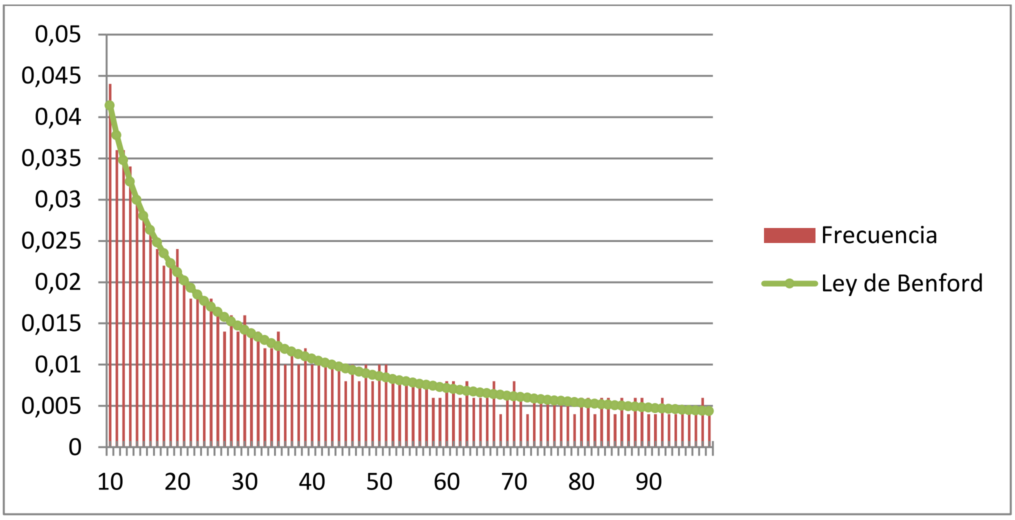
Al calcular sus primeros 100 términos, y estudiar la frecuencia del primer dígito significativo obtenemos los siguientes resultados.

Dígito	Frecuencia	Frecuencia relativa
1	30	0,300
2	17	0,170
3	13	0,130
4	10	0,100
5	7	0,070
6	7	0,070
7	6	0,060
8	5	0,050
9	5	0,050

Al aplicar el test de Bondad de ajuste de Chi-cuadrado, con 8 grados de libertad y $\alpha = 0,05$, se obtiene $\chi^2 = 0,22 < 15,51$. Por lo tanto, es aceptada la hipótesis de que este conjunto de datos, los primeros 100 términos de la sucesión de potencias de 2, satisface la Ley de Benford para el primer dígito significativo.

Al aplicar el Test MAD para el primer dígito significativo obtenemos el valor 0,004 lo que describe una gran conformidad de los primeros 100 términos de la sucesión de potencias de 2 con la Ley de Benford para el primer dígito significativo.

También podemos estudiar la frecuencia de los primeros dos dígitos significativos. Los cuales aparecen representados en el siguiente gráfico.



Al aplicar el test de Bondad de ajuste de Chi-cuadrado, con 89 grados de libertad y $\alpha = 0,05$, se obtiene $\chi^2 = 25,86 < 112,02$. Por lo tanto, es aceptada la hipótesis de que este conjunto de datos, los primeros 100 términos de la sucesión de potencias de 2, satisface la Ley de Benford para el primer y segundo dígito significativo.

Al aplicar el Test MAD para los primeros dos dígitos significativos obtenemos el valor 0,009 lo que describe una no conformidad de los primeros 100 términos de la sucesión de potencias de 2 con la Ley de Benford para el primer y segundo dígito significativo. Esto podría explicarse debido a que el conjunto de datos es relativamente pequeño. Al calcular las primeras 300 potencias de 2 naturales y aplicar el Test MAD para los primeros dos dígitos significativos obtenemos el valor 0,0044 lo que describe una no conformidad de los primeros 300 términos de la sucesión de potencias de 2 con la Ley de Benford para el primer y segundo dígito significativo. Sin embargo, Al calcular las primeras 500 potencias de 2 y aplicar el Test MAD para los primeros dos dígitos significativos obtenemos el valor 0,00088 lo que describe una gran conformidad de los primeros 500 términos de la sucesión de potencias de 2 con la Ley de Benford para el primer y segundo dígito significativo.

María Caputi Zunini
IPES - CFE/ANEP-UdelaR

Capítulo 8. Análisis del cumplimiento de la Ley de Benford: censo de la población uruguaya en 2011

En la siguiente tabla se muestra la población total de localidades uruguayas con más de 1.000 habitantes según el censo realizado en 2011[6]. Dicho censo tuvo la particularidad que, a diferencia de los anteriores, no fue realizado todo un mismo día si no que llevó varios meses en concluirse. Cabe preguntarse si los resultados del mismo serán válidos. La Ley de Benford provee una estrategia para analizar los datos recogidos.

Departamento	Localidad	Total
Artigas	**Artigas**	40.657
Canelones	**Canelones**	19.865
Cerro Largo	**Melo**	51.830
Colonia	**Colonia del Sacramento**	26.231
Durazno	**Durazno**	34.368
Montevideo	**Montevideo**	1.304.729
Paysandú	**Paysandú**	76.412
Rivera	**Vichadero**	3.698
Río Negro	**Fray Bentos**	24.406
Rocha	**Rocha**	25.422
Salto	**Salto**	104.011
Soriano	**Cardona**	4.600

A continuación aparece una tabla en donde se registran las frecuencias del primer dígito significativo del total de población de las 176 diferentes localidades uruguayas de más de 1.000 habitantes.

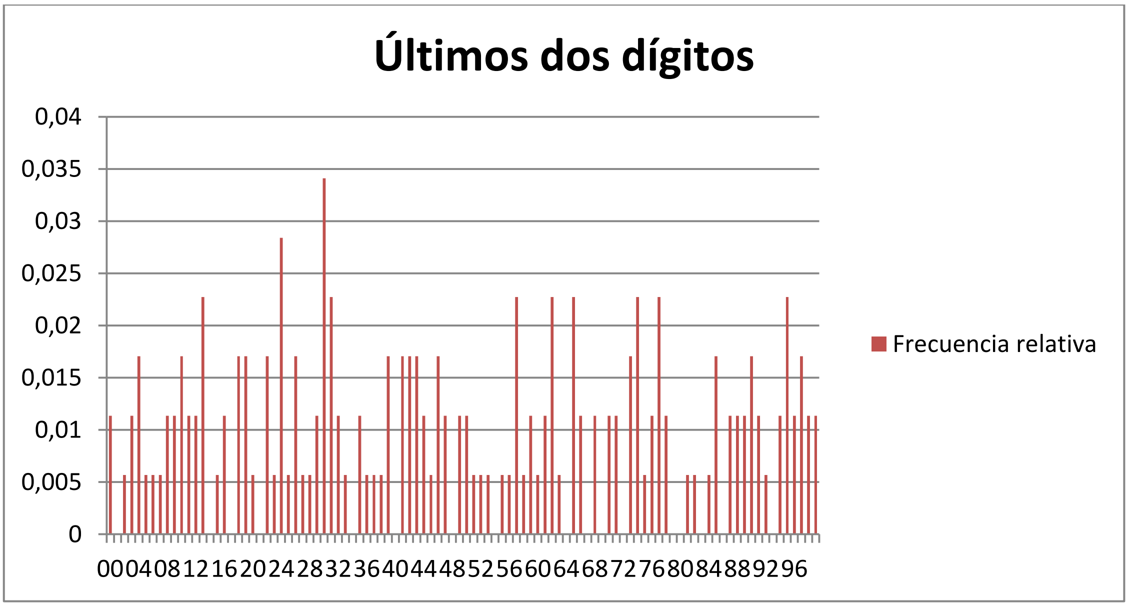
Dígito	1	2	3	4	5	6	7	8	9
Frecuencia	70	37	21	9	7	10	9	6	7

Al aplicar el test de Bondad de ajuste de Chi-cuadrado, con 8 grados de libertad y $\alpha = 0,05$, se obtiene $\chi^2 = 15,48 < 15,51$. Por lo tanto, es aceptada la hipótesis de que este conjunto de datos satisface la Ley de Benford para el primer dígito significativo.

[6] En base a la información extraída de Instituto Nacional de Estadística. (s.f.). *Instituto Nacional de Estadística Uruguay*. Recuperado el 14 de Septiembre de 2015, de http://www.ine.gub.uy/ La tabla completa se encentra en el Apéndice.

Al aplicar el Test MAD para el primer dígito significativo obtenemos el valor 0,029 lo que describe una no conformidad de la cantidad de habitantes de las localidades uruguayas con más de mil habitantes con la Ley de Benford para el primer dígito significativo. Esto podría explicarse debido a que el conjunto de datos es pequeño o tal vez se deba a otras razones como invención de datos.

Podemos utilizar el Test de los últimos dos dígitos para analizar si hay signos de invención en los datos. Consideraremos los últimos dos dígitos como las decenas y unidades de los datos. A continuación, aparece un gráfico en donde se muestra la frecuencia relativa de los últimos dos dígitos.



Se observa que hay varias diferencias entre la distribución de los últimos dos dígitos con la distribución uniforme. Pero, al aplicar el test de Bondad de ajuste de Chi-cuadrado, con 100 grados de libertad y $\alpha = 0,05$, se obtiene $\chi^2 = 92,18 < 124,34$. Por lo tanto, es aceptada la hipótesis de que la distribución de los últimos dos dígitos satisface la distribución uniforme.

Capítulo 9. La Ley de Benford para cualquier base [7]

9.1 Construcción de la σ-álgebra

Un primer paso para realizar una rigurosa construcción de la Ley de Benford es definir un correcto dominio para la función de probabilidad.

Este trabajo se restringirá a los primeros dígitos significativos de un número real positivo.

<u>Definición 9.1.1 Función mantisa</u> Dado un entero positivo *b>1*. Se define la función Mantisa, $M_b : \mathbb{R}^+ \to [1,b)$ tal que $M_b\left(x\right) = m$, siendo *m* el único número perteneciente a [1,b) que verifica $x = m.b^n$ para algún $n \in \mathbb{Z}$.

<u>Observación 9.1.1</u> M_b está bien definida ya que hay unicidad en la escritura de *x* en base *b*.

<u>Definición 9.1.2</u> Para cada $E \subset [1,b)$, sea $\left\langle E \right\rangle_b = M_b^{-1}(E)$.

<u>Proposición 9.1.1</u> Sea $E \subset [1,b)$, entonces $\left\langle E \right\rangle_b = \bigcup_{n\in\mathbb{Z}} b^n E$ siendo esta unión disjunta. [8]

Demostración.

Sea $x \in \left\langle E \right\rangle_b \Rightarrow x \in M_b^{-1}(E) \Rightarrow \exists n_0 \in \mathbb{Z},\ \exists m \in E$ tal que $x = m.b^{n_0} \Rightarrow x \in \bigcup_{n\in\mathbb{Z}} b^n E.$

Sea $x \in \bigcup_{n\in\mathbb{Z}} b^n E \Rightarrow \exists n_0 \in \mathbb{Z}$ tal que $x \in b^{n_0} E \Rightarrow \exists m \in E$ tal que $x = m.b^{n_0}$

$\Rightarrow x \in M_b^{-1}(E) \Rightarrow x \in \left\langle E \right\rangle_b$.

Por lo tanto, $\left\langle E \right\rangle_b = \bigcup_{n\in\mathbb{Z}} b^n E.$ □.

Resta ver que la unión es disjunta. Sean $m,n \in \mathbb{Z},\ m > n$

---

[7] En base a Hill, TP (1995). *Base-Invariance Implies Benford's Law*. Proceedings of the American Mathematical Society 123(3), 887-895.

[8] Sean $E \subset \mathbb{R}$ y $a \in \mathbb{R},\ \ aE := \left\{ ae \,/\, e \in E \right\}$.

$$\left.\begin{array}{l}\text{Como } E\subset[1,b),\ \text{entonces, } b^nE\subset[b^n,b^{n+1}) \text{ y } b^mE\subset[b^m,b^{m+1})\\ \text{Como } m>n,\ b^m\geq b^{n+1}.\ \text{Entonces, } [b^m,b^{m+1})\cap[b^n,b^{n+1})=\varnothing\end{array}\right\}\Rightarrow b^nE\cap b^mE=\varnothing$$

Por lo tanto, la unión es disjunta.□.

Proposición 9.1.2 Dados x_0, x_1, x_2 reales tales que $1\leq x_0<x_1<x_2<b$, se cumple que $\langle[x_0,x_2)\rangle_b=\langle[x_0,x_1)\rangle_b\uplus\langle[x_1,x_2)\rangle_b$.

Demostración.

Por proposición 9.1.1 $\langle[x_0,x_1)\rangle_b=\bigcup_{n\in\mathbb{Z}}b^n[x_0,x_1)=\bigcup_{n\in\mathbb{Z}}\left[b^nx_0,b^nx_1\right)$. Análogamente, $\langle[x_1,x_2)\rangle_b=\bigcup_{n\in\mathbb{Z}}\left[b^nx_1,b^nx_2\right)$, siendo las uniones disjuntas.

Para todo n entero $\left[b^nx_0,b^nx_1\right)\cap\left[b^nx_1,b^nx_2\right)=\varnothing$ puesto que si $x\in\left[b^nx_0,b^nx_1\right)$ entonces, $x<b^nx_1$ y por lo tanto $x\notin\left[b^nx_1,b^nx_2\right)$.

Si $n>m$, $\left[b^n,b^{n+1}\right)\cap\left[b^m,b^{m+1}\right)=\varnothing$. Como $\left[b^nx_0,b^nx_1\right)\subset\left[b^n,b^{n+1}\right)$ y $\left[b^mx_1,b^mx_2\right)\subset\left[b^m,b^{m+1}\right)$, se concluye que $\left[b^nx_0,b^nx_1\right)\cap\left[b^mx_1,b^mx_2\right)=\varnothing$. Por lo tanto, $\langle[x_0,x_1)\rangle_b\uplus\langle[x_1,x_2)\rangle_b$.

Resta ver que $\langle[x_0,x_1)\rangle_b\uplus\langle[x_1,x_2)\rangle_b=\langle[x_0,x_2)\rangle_b$.

$$\langle[x_0,x_1)\rangle_b\uplus\langle[x_1,x_2)\rangle_b=\bigcup_{n\in\mathbb{Z}}\left[b^nx_0,b^nx_1\right)\uplus\bigcup_{n\in\mathbb{Z}}\left[b^nx_1,b^nx_2\right)=\bigcup_{n\in\mathbb{Z}}\left[b^nx_0,b^nx_2\right)=\bigcup_{n\in\mathbb{Z}}b^n[x_0,x_2)$$

Entonces, $\langle[x_0,x_1)\rangle_b\uplus\langle[x_1,x_2)\rangle_b=\bigcup_{n\in\mathbb{Z}}b^n[x_0,x_2)=\langle[x_0,x_2)\rangle_b$.□.

Proposición 9.1.3 $\mathcal{M}_b=\left\{\langle E\rangle_b : E\in\mathcal{B}\left([1,b)\right)\right\}$ es una σ-álgebra en $\mathbb{R}^+$ que se denominará σ-álgebra mantisa.

Demostración.

1. Veamos que $\mathcal{M}_b\subset\mathbb{R}^+$ y que $\mathcal{M}_b$ es no vacío.

Por definición de $\langle E\rangle_b$, $\langle E\rangle_b \subset \mathbb{R}^+$, entonces, $\mathcal{M}_b \subset \mathbb{R}^+$. Por otra parte, $\mathcal{M}_b$ es no vacío puesto que $\mathrm{B}\left([1,b)\right)$ es no vacío por ser σ-álgebra.

2. Veamos que si $X \in \mathcal{M}_b$ entonces $X^c \in \mathcal{M}_b$.

Si $X \in \mathcal{M}_b$, entonces existe $E \in \mathrm{B}\left([1,b)\right)$ tal que $X = \langle E\rangle_b$. Como $E \in \mathrm{B}\left([1,b)\right)$ existe $B \in \mathrm{B}\left(\mathbb{R}\right)$ tal que $E = B \cap [1,b)$. Sea B^c el complemento de B en $\mathbb{R}$, entonces $B^c \in \mathrm{B}\left(\mathbb{R}\right)$ por ser $\mathrm{B}\left(\mathbb{R}\right)$ una σ-álgebra. Por lo tanto, $E^c = B^c \cap [1,b)$ pertenece a $\mathrm{B}\left([1,b)\right)$.

Veamos que X^c es $\bigcup_{n\in\mathbb{Z}} b^n E^c$. Para eso se probará que $X \cup \bigcup_{n\in\mathbb{Z}} b^n E^c = \mathbb{R}^+$ siendo esta unión disjunta.

Como $X = \langle E\rangle_b = \bigcup_{n\in\mathbb{Z}} b^n E$ siendo la unión disjunta, $X \cup \bigcup_{n\in\mathbb{Z}} b^n E^c = \bigcup_{n\in\mathbb{Z}} b^n E \cup \bigcup_{n\in\mathbb{Z}} b^n E^c$

Sean *m* y *n* naturales distintos, supongamos *m*>*n*, entonces $b^m \geq b^{n+1}$. Por lo tanto, $[b^m, b^{m+1}) \cap [b^n, b^{n+1}) = \varnothing$. Como $E \in \mathrm{B}\left([1,b)\right)$, entonces $b^n E \subset [b^n, b^{n+1})$ y $b^m E^c \subset [b^m, b^{m+1})$. Concluyendo así que $b^n E \cap b^m E^c = \varnothing$ para m y n distintos.

Supongamos *m*=*n*.

Sea $x \in b^n E \cap b^n E^c$. Entonces, $x = b^n.e$ con $e \in E$ y $x = b^n.e'$ con $e' \in E^c$. Por lo tanto, $e = e'$ lo que es imposible puesto que significaría que *E* y E^c tienen un elemento común. Por lo tanto, $b^n E \cap b^n E^c = \varnothing$.

Se concluye que la unión $X \cup \left(\bigcup_{n\in\mathbb{Z}} b^n E^c\right)$ es disjunta. Además,

$$X \cup \bigcup_{n\in\mathbb{Z}} b^n E^c = \bigcup_{n\in\mathbb{Z}} b^n E \cup \bigcup_{n\in\mathbb{Z}} b^n E^c = \bigcup_{n\in\mathbb{Z}} b^n.\left(E \cup E^c\right) = \bigcup_{n\in\mathbb{Z}} b^n.[1,b) = \bigcup_{n\in\mathbb{Z}} \left[b^n, b^{n+1}\right) = \mathbb{R}^+ .$$

Por lo tanto, $X^c \in \mathcal{M}_b$.

3. Veamos que si $A_m \in \mathcal{M}_b$ para todo $m \in \mathbb{N}$, entonces $\bigcup_{m\in\mathbb{N}} A_m \in \mathcal{M}_b$.

Dado m natural, $A_m \in \mathcal{M}_b$, existe $E_m \in \mathcal{B}\left([1,b)\right)$ tal que $A_m = \langle E_m \rangle_b$.

$$\bigcup_{m\in\mathbb{N}} A_m = \bigcup_{m\in\mathbb{N}} \langle E_m \rangle_b = \bigcup_{m\in\mathbb{N}}\bigcup_{n\in\mathbb{Z}} b^n.E_m = \bigcup_{n\in\mathbb{Z}}\bigcup_{m\in\mathbb{N}} b^n.E_m = \bigcup_{n\in\mathbb{Z}} b^n\left(\bigcup_{m\in\mathbb{N}} E_m\right)$$

Por ser $E_m \in \mathcal{B}\left([1,b)\right)$ para cada m natural y como $\mathcal{B}\left([1,b)\right)$ es una σ-álgebra, $\bigcup_{m\in\mathbb{N}} E_m \in \mathcal{B}\left([1,b)\right)$. Entonces, $\bigcup_{m\in\mathbb{N}} A_m = \left\langle \bigcup_{m\in\mathbb{N}} E_m \right\rangle_b \in \mathcal{M}_b$.

De lo probado en 1, 2 y 3 se deduce que $\mathcal{M}_b$ es una σ-álgebra en $\mathbb{R}^+$. □.

Proposición 9.1.4: Propiedades de la σ-álgebra mantisa

Dado *b* entero positivo mayor a 1.

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

(ii) Para todo $n \in \mathbb{N}$, $\mathcal{M}_b \subset \mathcal{M}_{b^n} \subset \mathcal{B}$.

(iii) Dados $S \in \mathcal{M}_b$ y $a > 0$, entonces, $aS \in \mathcal{M}_b$.

(iv) Dados $S \in \mathcal{M}_b$ y $n \in \mathbb{N}$, entonces, $b^n S = S$.

Demostración.

(i) Dados $n \in \mathbb{N}$ y $E \subset [1,b)$.

Se quiere probar que $\langle E \rangle_b = \bigcup_{k=0}^{n-1} \langle b^k E \rangle_{b^n}$.

Por definición $\langle b^k E \rangle_{b^n} = \bigcup_{i\in\mathbb{Z}} \left(b^n\right)^i.\left(b^k E\right) = \bigcup_{i\in\mathbb{Z}} b^{ni+k} E \Rightarrow \langle b^k E \rangle_{b^n} \subset \bigcup_{i\in\mathbb{Z}} b^i E = \langle E \rangle_b$ para todo $k = 0...n-1$. Entonces, $\langle E \rangle_b \supset \bigcup_{k=0}^{n-1} \langle b^k E \rangle_{b^n}$.

Resta ver que $\langle E \rangle_b \subseteq \bigcup_{k=0}^{n-1} \langle b^k E \rangle_{b^n}$.

Sea $b^m e \in \langle E \rangle_b$, con $m \in \mathbb{Z}$ y $e \in E$. Existen $q \in \mathbb{Z}, r \in \mathbb{N}$ tal que $m = nq + r$ con $0 \le r < n$

Así, $b^m e = b^{nq+r} e = \left(b^n\right)^q .b^r e \in \left\langle b^r E \right\rangle_{b^n} \subseteq \bigcup_{k=0}^{n-1} \left\langle b^k E \right\rangle_{b^n} \Rightarrow b^m e \in \bigcup_{k=0}^{n-1} \left\langle b^k E \right\rangle_{b^n}$.

$\therefore \bigcup_{k=0}^{n-1} \left\langle b^k E \right\rangle_{b^n} = \langle E \rangle_b$. □.

(ii) Dado $n \in \mathbb{N}$. Sea $\langle E \rangle_b \in \mathcal{M}_b$. Por (i) $\langle E \rangle_b = \bigcup_{k=0}^{n-1} \left\langle b^k E \right\rangle_{b^n}$

Veamos que $\left\langle b^k E \right\rangle_{b^n} \in \mathcal{M}_{b^n}$. Como $\langle E \rangle_b \in \mathcal{M}_b$, para todo $k = 0...n-1$, por definición de $\mathcal{M}_b$, $E \in \mathcal{B}\left(\left[1,b\right)\right)$. Como $\mathcal{B}\left(\left[1,b\right)\right) \subset \mathcal{B}\left(\mathbb{R}^+\right)$, $E \in \mathcal{B}\left(\mathbb{R}^+\right)$ y también $b^k E \in \mathcal{B}\left(\mathbb{R}^+\right)$. Como $E \in \mathcal{B}\left(\left[1,b\right)\right)$, $E \subset \left[1,b\right)$, entonces $b^k E \subset \left[b^k, b^{k+1}\right) \subset \left[1, b^n\right)$. Por lo tanto, $b^k E \in \mathcal{B}\left(\left[1,b^n\right)\right)$. De lo que se concluye que, $\left\langle b^k E \right\rangle_{b^n} \in \mathcal{M}_{b^n}$.

Como $\mathcal{M}_{b^n}$ es σ-álgebra es cerrada por uniones, entonces, $\bigcup_{k=0}^{n-1} \left\langle b^k E \right\rangle_{b^n} \in \mathcal{M}_{b^n}$. Por lo tanto, $\langle E \rangle_b \in \mathcal{M}_{b^n}$. Por lo tanto, $\mathcal{M}_b \subset \mathcal{M}_{b^n}$.□.

Veamos que $\mathcal{M}_c \subset \mathcal{B}$ para todo $c \in \mathbb{N}$, *c>2.*

Sea $\langle E \rangle_c \in \mathcal{M}_c$, entonces $\langle E \rangle_c = \bigcup_{n \in \mathbb{Z}} c^n E$ con $E \in \mathcal{B}\left(\left[1,c\right)\right)$.

Como $\mathcal{B}\left(\left[1,c\right)\right) \subset \mathcal{B}\left(\mathbb{R}^+\right)$, $E \in \mathcal{B}\left(\mathbb{R}^+\right)$. Entonces $tE \in \mathcal{B}\left(\mathbb{R}^+\right)$ para todo *t* real positivo. En particular, se cumple que $c^n E \in \mathcal{B}\left(\mathbb{R}^+\right)$para todo *n* entero. Por lo tanto, $\bigcup_{n \in \mathbb{Z}} c^n E \in \mathcal{B}\left(\mathbb{R}^+\right)$. De lo que se deduce que $\langle E \rangle_c \in \mathcal{B}\left(\mathbb{R}^+\right)$.□.

(iii) Dados $\mathcal{S} \in \mathcal{M}_b$ y $a > 0$. Por definición de $\mathcal{M}_b$ existe $E \in \mathcal{B}\left(\left[1,b\right)\right)$ tal que $\langle E \rangle_b = S$.

Como $aS = a\langle E\rangle_b = a.\bigcup_{n\in\mathbb{Z}} b^n E = \bigcup_{n\in\mathbb{Z}} b^n .aE = \bigcup_{n\in\mathbb{Z}} b^n .b^t .\alpha E = \bigcup_{r\in\mathbb{Z}} b^r .\alpha E$ siendo t y α tales que $a = b^t.\alpha$ con $\alpha \in [1,b)$ con t entero.

Como $\alpha \in [1,b)$ y $E \in \mathcal{B}([1,b))$, entonces $\alpha E \in \mathcal{B}(\mathbb{R}^+)$ y $\alpha E \subset [1,b^2)$.

Sean $E_1 = \alpha E \cap [1,b)$ y $E_2 = \alpha E \cap [b,b^2)$. Entonces, $E_1 \in \mathcal{B}([1,b))$ y $E_2 \in \mathcal{B}(\mathbb{R}^+)$.

Por lo tanto $aS = \bigcup_{r\in\mathbb{Z}} \left(b^r .E_1 \uplus b^r E_2\right) = \bigcup_{r\in\mathbb{Z}} b^r .E_1 \uplus \bigcup_{r\in\mathbb{Z}} b^r E_2 = \langle E_1\rangle_b \uplus \bigcup_{r\in\mathbb{Z}} b^r E_2$

Sea $E_3 = \frac{1}{b} E_2$. Como $E_2 \subset [b,b^2)$, entonces $E_3 \subset [1,b)$ y como $E_2 \in \mathcal{B}(\mathbb{R}^+)$ se deduce que $E_3 \in \mathcal{B}([1,b))$. Por lo tanto, $\bigcup_{r\in\mathbb{Z}} b^r E_2 = \bigcup_{r\in\mathbb{Z}} b^r E_3 = \langle E_3\rangle_b$.

Entonces, $aS = \langle E_1\rangle_b \uplus \langle E_3\rangle_b$. Como $\langle E_1\rangle_b$ y $\langle E_3\rangle_b$ son elementos de la σ-álgebra mantisa $\mathcal{M}_b$, aS también lo es.□.

(iv) Dado $S \in \mathcal{M}_b$. Por definición de $\mathcal{M}_b$ existe $E \in \mathcal{B}[1,b)$ tal que $\langle E\rangle_b = S$.

Así, $b^n S = b^n \langle E\rangle_b = b^n.\bigcup_{m\in\mathbb{Z}} b^m E = \bigcup_{m\in\mathbb{Z}} b^{n+m} E = \bigcup_{i\in\mathbb{Z}} b^i E = \langle E\rangle_b = S$. □.

Observación 9.1.2 De la Proposición 9.1.4 (iii) podemos concluir que si $\langle E\rangle_b = S$ con $E \in \mathcal{B}([1,b))$ y $a > 0$, $aS = \langle E_1\rangle_b \uplus \langle E_3\rangle_b$ siendo $E_1 = \alpha E \cap [1,b)$ y $E_3 = \frac{1}{b}\left(\alpha E \cap [b,b^2)\right)$ elementos de $\mathcal{B}([1,b))$ y α tal que $a = b^t.\alpha$ con $\alpha \in [1,b)$ con t entero.

En particular, si E es un intervalo, $E = [\delta,\gamma) \subset [1,b)$, y $a \in [1,b)$. Entonces, $1 \le a\delta < a\gamma < b$ o $1 \le a\delta < b \wedge b \le a\gamma < b^2$ o $b \le a\delta < a\gamma < b^2$.

Si $1 \le a\delta < a\gamma < b$, $E_1 = [a\delta, a\gamma)$ y $E_3 = \varnothing$. Entonces, $a\langle E\rangle_b = \langle [a\delta, a\gamma)\rangle_b$

Si $1 \le a\delta < b \wedge b \le a\gamma < b^2$, $E_1 = [a\delta, b)$ y $E_3 = \left[1, \frac{a\gamma}{b}\right)$. Entonces,

$a\langle E\rangle_b = \langle [a\delta, b)\rangle_b \uplus \left\langle \left[1, \frac{a\gamma}{b}\right)\right\rangle_b$

Si $b \le a\delta < a\gamma < b^2$, $E_1 = \varnothing$ y. $E_3 = \left[\frac{a\delta}{b}, \frac{a\gamma}{b}\right)$ Entonces, $a\langle E\rangle_b = \left\langle \left[\frac{a\delta}{b}, \frac{a\gamma}{b}\right)\right\rangle_b$.

Definición 9.1.3 Sea $E \in \mathcal{B}([1,b))$ y $a > 0$. Definimos $\langle aE\rangle_b = a\langle E\rangle_b$.

Observación 9.1.3 La definición 9.1.3 si bien puede ser considerada como un abuso de notación ya que aE puede no ser un elemento de $\mathcal{B}([1,b))$, tiene sentido debido a lo concluido en la observación 9.1.2.

9.2 Ley general de probabilidad para los primeros k dígitos significativos de un número en base *b*

Definición 9.2.1 i-ésimo dígito significativo del real positivo *x* representado en la base *b*. Sean $b \in \mathbb{N}$, $b \ge 2$, $i \in \mathbb{N}^*$. Se definen $D_i^b : \mathbb{R}^+ \to \{0,1,\dots,b-1\}$ tales que $D_i^b = \left(g_b^{-1}\right)_i \circ M_b$, siendo M_b la función mantisa antes definida, $(.)_i$ es la proyección en la coordenada i-ésima, $(a_1, a_2, \dots)_i = a_i$ y siendo g_b tal que $g_b(i_1, i_2, \dots) = \sum_{k=1}^{+\infty} \frac{i_k}{b^{k-1}}$, con $i_1 \in \{1,2,\dots,b-1\}$ y $i_k \in \{0,1,2,\dots,b-1\}$ para $k \in \mathbb{N}, k \ge 2$, y definiendo $\left(g_b^{-1}\right)$ como la "inversa" de g_b de manera que no haya ambigüedades (por ejemplo, en base 10: $\left(g^{-1}{}_{10}\right)(3) = \left(g^{-1}{}_{10}\right)(2,999....) = (3,0,0,0,....)$ y $\left(g^{-1}{}_{10}\right)(9,99...) = (1,0,0,0....)$ y en base 5: $\left(g^{-1}{}_{5}\right)(3) = \left(g^{-1}{}_{5}\right)(2,444....) = (3,0,0,0,....)$ y $\left(g^{-1}{}_{5}\right)(4,444...) = (1,0,0,0....)$).

En otras palabras, $D_i^b(x)$ es el i-ésimo dígito significativo del real positivo *x* representado en la base *b*.

Proposición 9.2.1 Para todo d_1 con $d_1 \in \{1,\dots,b-1\}$ y para todos d_i con $d_i \in \{0,1,\dots,b-1\}$, $i \in \mathbb{N}$, *i*>2, se cumple que $\left\{D_i^b = d_i\right\} \in \mathcal{M}_b$.

Demostración.

Analicemos el caso para i=1. Veamos qué elementos pertenecen al conjunto $E=\left\{D_1^b=d_1\right\}\cap[1,b)$. Si $x\in E$, entonces $x=d_1,j_2j_3j_4...$ con $j_n\in\{0,1,...,b-1\}$ pero no siendo $j_n=b-1$ para todo *n* mayor o igual a 2. Por lo que $E=[d_1,d_1+1)$ es un elemento de $\mathrm{B}([1,b))$ por ser un intervalo contenido en $[1,b)$.

Para describir el resto de los elementos de $\left\{D_1^b=d_1\right\}$ basta multiplicar los elementos de E por potencias enteras de *b*. Puesto que, por ejemplo, $x=d_1j_2,j_3j_4...$, $x=d_1j_2j_3,j_4...$y $x=0,d_1j_2j_3j_4...$ con $j_n\in\{0,1,...,b-1\}$ pero no siendo $j_n=b-1$ para todo *n* mayor o igual a 2, son elementos de $\left\{D_1^b=d_1\right\}$. Entonces, $\left\{D_1^b=d_1\right\}=\bigcup_{n\in\mathbb{Z}}b^nE$, como $E\in\mathrm{B}([1,b))$, se deduce que $\left\{D_1^b=d_1\right\}\in\mathcal{M}_b$.

Analicemos el caso para i=2. Como en el caso anterior veamos qué elementos pertenecen al conjunto $E=\left\{D_2^b=d_2\right\}\cap[1,b)$. Si $x\in E$, entonces $x=j_1,d_2j_3j_4...$ con $j_1\in\{1,2,...,b-1\}$ y $j_n\in\{0,1,...,b-1\}$ para *n>2*, pero no siendo $j_n=b-1$ para todo *n* mayor o igual a 3. Por lo que $E=\bigcup_{k=1}^{b-1}\left[k,d_2;k,(d_2+1)\right)$ si $d_2\neq b-1$ y $E=\bigcup_{k=1}^{b-1}\left[k,d_2;k+1\right)$ si $d_2=b-1$. En ambos casos E es unión finita de *b-1* intervalos contenidos en $[1,b)$, por lo tanto, E es un elemento de $\mathrm{B}([1,b))$.

Para describir el resto de los elementos de $\left\{D_2^b=d_2\right\}$ basta multiplicar los elementos de E por potencias enteras de *b*. Puesto que, por ejemplo, $x=j_1d_2,j_3j_4...$, $x=j_1d_2j_3,j_4...$y $x=0,j_1d_2j_3j_4...$ con $j_1\in\{1,2,...,b-1\}$ y $j_n\in\{0,1,...,b-1\}$ para *n>2*, pero no siendo $j_n=b-1$ para todo *n* mayor o igual a 3, son elementos de $\left\{D_2^b=d_2\right\}$. Entonces, $\left\{D_2^b=d_2\right\}=\bigcup_{n\in\mathbb{Z}}b^nE$, como $E\in\mathrm{B}([1,b))$,se deduce que $\left\{D_2^b=d_2\right\}\in\mathcal{M}_b$.

Analicemos el caso para i=3. Como en el caso anterior veamos qué elementos pertenecen al conjunto $E=\left\{D_3^b=d_3\right\}\cap[1,b)$. Si $x\in E$, entonces $x=j_1,j_2d_3j_4...$ con

$j_1 \in \{1,2,...,b-1\}$ y $j_n \in \{0,1,..., b-1\}$ para $n \neq 3$, pero no siendo $j_n = b-1$ para todo n mayor o igual a 4. Por lo que $E = \bigcup_{k=1,h=0}^{b-1} [k, hd_3; k, h(d_3+1))$ si $d_3 \neq b-1$ y $E = \left(\bigcup_{k=1,h=0}^{b-1,b-2} [k, hd_3; k, h+1) \right) \cup [k, (b-1)d_3; k+1)$ si $d_3 = b-1$. En ambos casos E es unión finita de *b.(b-1)* intervalos contenidos en $[1,b)$, por lo tanto, E es un elemento de $\mathcal{B}([1,b))$.

Para describir el resto de los elementos de $\{D_3^b = d_3\}$ basta multiplicar los elementos de E por potencias enteras de *b*. Puesto que, por ejemplo, $x = j_1 j_2, d_3 j_4 ...$, $x = j_1 j_2 d_3, j_4 ...$ y $x = j_1, j_2 d_3 j_4 ...$ con $j_1 \in \{1,2,...,b-1\}$ y $j_n \in \{0,1,..., b-1\}$ para $n \neq 3$, pero no siendo $j_n = b-1$ para todo *n* mayor o igual a 4, son elementos de $\{D_3^b = d_3\}$. Entonces, $\{D_3^b = d_3\} = \bigcup_{n \in \mathbb{Z}} b^n E$, como $E \in \mathcal{B}([1,b))$, se deduce que $\{D_3^b = d_3\} \in \mathcal{M}_b$.

Analicemos el caso para i natural mayor que 3. Veamos qué elementos pertenecen al conjunto $E = \{D_i^b = d_i\} \cap [1,b)$. Si $x \in E$, entonces $x = j_1, j_2 j_3 ... j_{i-1} d_i j_{i+1} ...$ con $j_1 \in \{1,2,...,b-1\}$ y $j_n \in \{0,1,..., b-1\}$ para *n* mayor o igual que 2 y $n \neq i$, pero no siendo $j_n = b-1$ para todo n mayor a i. Por lo tanto, $E = \bigcup_{j_1=1, j_2=0,..., j_{i-1}=0}^{b-1} [j_1, j_2 j_3 ... d_i; j_1, j_2 j_3 ... (d_i+1))$ si $d_i \neq b-1$ y $E = \left(\bigcup_{j_1=1, j_2=0,..., j_{i-1}=0}^{b-1, b-1,..., b-1, b-2} [j_1, j_2 j_3 ... d_i; j_1, j_2 j_3 ... (d_i+1)) \right) \cup [j_1, j_2 j_3 ... (b-1) d_i; j_1, j_2 j_3 ... (j_{i-2}+1))$ si $d_i = b-1$. En ambos casos E es unión finita de $b^{i-2}.(b-1)$ intervalos contenidos en $[1,b)$, por lo tanto, E es un elemento de $\mathcal{B}([1,b))$.

Para describir el resto de los elementos de $\{D_i^b = d_i\}$ basta multiplicar los elementos de E por potencias enteras de *b*. Puesto que, por ejemplo, $x = j_1 j_2, j_3 ... j_{i-1} d_i j_{i+1} ...$, $x = j_1 j_2 j_3 ... j_{i-1} d_i, j_{i+1} ...$ y $x = 0, j_1 j_2 j_3 ... j_{i-1} d_i j_{i+1} ...$ con $j_1 \in \{1,2,...,b-1\}$ y $j_n \in \{0,1,..., b-1\}$ para *n* mayor o igual que 2 y $n \neq i$, pero no siendo $j_n = b-1$ para todo n mayor a i, son elementos de $\{D_i^b = d_i\}$. Entonces, $\{D_i^b = d_i\} = \bigcup_{n\in\mathbb{Z}} b^n E$, como $E \in \mathrm{B}([1,b))$, se deduce que $\{D_i^b = d_i\} \in \mathcal{M}_b$. □.

Definición 9.2.2 Ley general de probabilidad para los primeros k dígitos significativos en base *b*

Esta Ley puede enunciarse utilizando las funciones D_i^b como sigue:

Dado $b \in \mathbb{N},\ b > 1$.
Dado $k,\ k \in \mathbb{Z}^+$, para todo d_1 con $d_1 \in \{1,...,b-1\}$ y para todos d_j con $d_j \in \{0,1,...,b-1\}$, $j = 2...k,\ P_b\left(\bigcap_{i=1}^{k}\{D_i^b = d_i\}\right) = \log_b\left(1+\left(\sum_{i=1}^{k} d_i.b^{k-i}\right)^{-1}\right)$.

O utilizando explícitamente los elementos de $\mathcal{M}_b$: $P_b\left(\langle[1,\gamma)\rangle_b\right) = \log_b \gamma$, para $\gamma \in [1,b)$.

La terna ($\mathbb{R}^+$, $\mathcal{M}_b$, P_b) es un espacio de probabilidad.

Observación 9.2.1: Dados $x_0,\ x_1,\ x_2$ reales tales que $1 \le x_0 < x_1 < x_2 < b$. De acuerdo a la proposición 9.1.2 se cumple que $\langle[x_0,x_2)\rangle_b = \langle[x_0,x_1)\rangle_b \uplus \langle[x_1,x_2)\rangle_b$, por ser P_b una probabilidad concluimos que $P_b\left(\langle[x_0,x_2)\rangle_b\right) = P_b\left(\langle[x_0,x_1)\rangle_b\right) + P_b\left(\langle[x_1,x_2)\rangle_b\right)$.

Observación 9.2.2 La definición 9.2.1 de P_{10} con $k=1$ coincide con la definición de la ley de Benford para el primer dígito.

Por ejemplo, si se busca calcular la probabilidad de que el primer dígito de un real positivo sea 4, según la Ley de Benford estudiada en el capítulo 3 es $\log_{10}\left(1+\frac{1}{4}\right)$. Según la definición de P_{10}, sería $P_{10}\left(\langle[4,5)\rangle_{10}\right)$. Aplicando la observación 9.2.1 $P_{10}\left(\langle[4,5)\rangle_{10}\right) = P_{10}\left(\langle[1,5)\rangle_{10}\right) - P_{10}\left(\langle[1,4)\rangle_{10}\right) = \log_{10} 5 - \log_{10} 4 = \log_{10}\left(\frac{5}{4}\right)$.

Observación 9.2.3 Se puede identificar la probabilidad P_b definida en $\mathcal{M}_b$ con su representación como probabilidad en $\mathrm{B}([1,b))$, Pues por la definición, la relación

$P_b\left(\langle E\rangle_b\right)=P(E)$, para todo $E\in\mathcal{B}\ ([1,b))$ define una correspondencia entre la probabilidad P_b en ($\mathbb{R}^+$, $\mathcal{M}_b$) y la probabilidad P en $\mathcal{B}\left([1,b)\right)$.

Observación 9.2.4: Dados x_0, x_1 reales tales que $1\le x_0<x_1<b$. Por la definición de P_b concluimos que $P_b\left(\langle[x_0,x_1]\rangle_b\right)=P_b\left(\langle[x_0,x_1)\rangle_b\right)+P_b\left(\langle\{x_1\}\rangle_b\right)$. Como $P_b\left(\langle\{x_1\}\rangle_b\right)=0$, se concluye que $P_b\left(\langle[x_0,x_1]\rangle_b\right)=P_b\left(\langle[x_0,x_1)\rangle_b\right)$. Análogamente, se concluye que $P_b\left(\langle(x_0,x_1)\rangle_b\right)=P_b\left(\langle(x_0,x_1]\rangle_b\right)=P_b\left(\langle[x_0,x_1)\rangle_b\right)$.

9.3 La Ley de Benford es invariante frente a cambios de escala

Definición 9.3.1: Una función de probabilidad P en ($\mathbb{R}^+$, $\mathcal{M}_b$) es invariante frente a cambios de escala, si existe $\alpha>0$, $\alpha\neq b^r$, $r\in\mathbb{Q}$ tal que para todo $S\in\mathcal{M}_b, P(S)=P(\alpha S)$.

Lema 9.3.1 Dados (Ω, $\mathcal{A}$, P) y (Ω, $\mathcal{A}$, Q) dos espacios de probabilidad. Sea A un subconjunto de $\mathcal{A}$, A es cerrado por intersecciones finitas y por complementos, $\langle \mathrm{A}\rangle=\mathcal{A}$ y $P|_A=Q|_A$. Entonces P=Q en $\mathcal{A}$.

Demostración.

Sea χ el mayor subconjunto de Ω que contiene a A, es cerrado por intersecciones finitas y por complementos, y verifica que $P|_\chi=Q|_\chi$. Dicho χ existe. Sea H el conjunto de todos los subconjuntos de Ω que verifican que: $\mathrm{A}\subset X$ para todo $X\in\mathrm{H}$, H es cerrado por intersecciones finitas y complementos, y que $P|_X=Q|_X$ para todo $X\in\mathrm{H}$. H es no vacío, ya que $A\in\mathrm{H}$. Sea una cadena $\mathcal{C}=\left(\chi_\lambda\right)_{\lambda\in\Lambda}$ de subconjuntos de H. $\mathcal{C}$ tiene cota superior $\bigcup_{\lambda\in\Lambda}\chi_\lambda$, puesto que por definición de unión $\chi_\lambda\subset\bigcup_{\lambda\in\Lambda}\chi_\lambda$ para todo $\lambda\in\Lambda$. Por el Lema de Zorn, se concluye que existe un elemento maximal χ en el conjunto H.

Afirmación: χ es una σ-álgebra de Ω.

Demostración.

1. χ es no vacío pues $\mathrm{A} \subset \chi$.

2. Veamos que si $X \in \chi$ y dado X^c en Ω, entonces $X^c \in \chi$.

Como $P|_{\chi} = Q|_{\chi}$, $P(X^c) = 1 - P(X) = 1 - Q(X) = Q(X^c)$. Entonces, $X^c \in \chi$.

3. Veamos que si $X_n \in \chi$ para todo $n \in \mathbb{N}$, entonces $\bigcup_{n=1}^{+\infty} X_n \in \chi$.

Sean $\overline{X_n} = X_n \setminus \left(\bigcup_{i=1}^{n-1} X_i \right)$. Entonces, $\overline{X_n} = X_n \cap \left(\bigcup_{i=1}^{n-1} X_i \right)^c$. Como $X_n \in \chi$ para todo $n \in \mathbb{N}$ y χ es cerrado por uniones finitas y complementos, se concluye que $\overline{X_n} \in \chi$ para todo $n \in \mathbb{N}$. Se observa que por construcción los conjuntos $\overline{X_n}$ son disjuntos dos a dos y que $\bigcup_{n=1}^{+\infty} X_n = \bigcup_{n=1}^{+\infty} \overline{X_n}$. Por lo tanto, $P\left(\bigcup_{n=1}^{+\infty} X_n \right) = P\left(\bigcup_{n=1}^{+\infty} \overline{X_n} \right) = \sum_{n=1}^{+\infty} P\left(\overline{X_n} \right) = \sum_{n=1}^{+\infty} Q\left(\overline{X_n} \right) = Q\left(\bigcup_{n=1}^{+\infty} \overline{X_n} \right) = Q\left(\bigcup_{n=1}^{+\infty} X_n \right)$. De lo que se deduce que $\bigcup_{n=1}^{+\infty} X_n \in \chi$.

De lo probado en 1, 2 y 3 se concluye que χ es una σ-álgebra de Ω. Además, como $\mathrm{A} \subset \chi$ y $\langle \mathrm{A} \rangle = \mathscr{A}$, se deduce que $\chi = \mathscr{A}$. Como $P|_{\chi} = Q|_{\chi}$, entonces, P=Q en $\mathscr{A}$. □.

<u>Teorema 9.3.1</u> P_b es la única probabilidad que es invariante frente a cambios de escala en ($\mathbb{R}^+$, $\mathcal{M}_b$).

Demostración.

Veamos primero que P_b es invariante frente a cambios de escala en ($\mathbb{R}^+$, $\mathcal{M}_b$).

Sea $\Im$ el conjunto de todas las posibles uniones finitas de intervalos contenidos en $[1,b)$.

Sea $\alpha \in [1,b)$ tal que $\alpha \neq b^r$, $r \in \mathbb{Q}$.

Sea $[\delta,\gamma)\subset[1,b)$. Se probará que $P_b\left(\alpha\left\langle[\delta,\gamma)\right\rangle_b\right)=P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\right)$.

Como $P_b\left(\left\langle[1,\gamma)\right\rangle_b\right)=\log_b\gamma,$ para $\gamma\in[1,b)$, por observación 9.2.1 se cumple que $P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\right)=P_b\left(\left\langle[1,\gamma)\right\rangle_b\right)-P_b\left(\left\langle[1,\delta)\right\rangle_b\right)=\log_b\gamma-\log_b\delta=\log_b\dfrac{\gamma}{\delta}$.

Si $1\le a\delta<a\gamma<b$, aplicando las observaciones 9.1.2 y 9.1.3 se concluye que $P_b\left(\alpha\left\langle[\delta,\gamma)\right\rangle_b\right)=P_b\left(\left\langle[\alpha\delta,\alpha\gamma)\right\rangle_b\right)$ y por observación 9.2.1 $P_b\left(\left\langle[\alpha\delta,\alpha\gamma)\right\rangle_b\right)=P_b\left(\left\langle[1,\alpha\gamma)\right\rangle_b\right)-P_b\left(\left\langle[1,\alpha\delta)\right\rangle_b\right)$. Como $P_b\left(\left\langle[1,\gamma)\right\rangle_b\right)=\log_b\gamma,$ para $\gamma\in[1,b)$, $P_b\left(\left\langle[\alpha\delta,\alpha\gamma)\right\rangle_b\right)=\log_b\alpha\gamma-\log_b\alpha\delta=\log_b\dfrac{\gamma}{\delta}=P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\right)$. Por lo tanto, $P_b\left(\alpha\left\langle[\delta,\gamma)\right\rangle_b\right)=P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\right)$.

Si $1\le a\delta<b\ \wedge\ b\le a\gamma<b^2$, aplicando las observaciones 9.1.2 y 9.1.3 se concluye que si, $P_b\left(\alpha\left\langle[\delta,\gamma)\right\rangle_b\right)=P_b\left(\left\langle[\alpha\delta,b)\right\rangle_b\right)+P_b\left(\left\langle\left[1,\dfrac{\alpha\gamma}{b}\right)\right\rangle_b\right)$. Como $P_b\left(\left\langle[\alpha\delta,b)\right\rangle_b\right)=P_b\left(\left\langle[1,b)\right\rangle_b\right)-P_b\left(\left\langle[1,\alpha\delta)\right\rangle_b\right)=\log_b b-\log_b\alpha\delta=1-\log_b\alpha\delta$ y $P_b\left(\left\langle\left[1,\dfrac{\alpha\gamma}{b}\right)\right\rangle_b\right)=\log_b\left(\dfrac{\alpha\gamma}{b}\right)=\log_b\alpha+\log_b\gamma-\log_b b=\log_b\alpha+\log_b\gamma-1$, entonces $P_b\left(\alpha\left\langle[\delta,\gamma)\right\rangle_b\right)=1-\log_b\alpha\delta+\log_b\alpha+\log_b\gamma-1=-\log_b\delta+\log_b\gamma=P_b\left(\left\langle[1,\gamma)\right\rangle_b\right)$. En suma, $P_b\left(\alpha\left\langle[\delta,\gamma)\right\rangle_b\right)=P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\right)$.

Si $b\le a\delta<a\gamma<b^2$, aplicando las observaciones 9.1.2 y 9.1.3 se concluye que $P_b\left(\alpha\left\langle[\delta,\gamma)\right\rangle_b\right)=P_b\left(\left\langle\left[\dfrac{a\delta}{b},\dfrac{a\gamma}{b}\right)\right\rangle_b\right)$ y por observación 9.2.1 $P_b\left(\left\langle\left[\dfrac{a\delta}{b},\dfrac{a\gamma}{b}\right)\right\rangle_b\right)=P_b\left(\left\langle\left[1,\dfrac{a\gamma}{b}\right)\right\rangle_b\right)-P_b\left(\left\langle\left[1,\dfrac{a\delta}{b}\right)\right\rangle_b\right)$. Como $P_b\left(\left\langle[1,\gamma)\right\rangle_b\right)=\log_b\gamma,$ para $\gamma\in[1,b)$, $P_b\left(\left\langle\left[\dfrac{\alpha\delta}{b},\dfrac{\alpha\gamma}{b}\right)\right\rangle_b\right)=\log_b\left(\dfrac{\alpha\gamma}{b}\right)-\log_b\left(\dfrac{\alpha\delta}{b}\right)=\log_b\dfrac{\gamma}{\delta}=P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\right)$.

Por lo tanto, $P_b\left(\alpha\left\langle[\delta,\gamma)\right\rangle_b\right)=P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\right)$.

Sean $[\delta,\gamma)\subset[1,b)$ y $[\beta,\varepsilon)\subset[1,b)$. Se probará que $P_b\left(\alpha\left\langle[\delta,\gamma)\cup[\beta,\varepsilon)\right\rangle_b\right)=P_b\left(\left\langle[\delta,\gamma)\cup[\beta,\varepsilon)\right\rangle_b\right)$.

Si $[\delta,\gamma)\cap[\beta,\varepsilon)\neq\varnothing$, entonces $[\delta,\gamma)\cup[\beta,\varepsilon)$ es un intervalo, lo que reduce la prueba al caso anterior.

Si $[\delta,\gamma)\cap[\beta,\varepsilon)=\varnothing$, entonces $[\delta,\gamma)\uplus[\beta,\varepsilon)$. Por proposición 9.1.2 $P_b\left(\left\langle[\delta,\gamma)\uplus[\beta,\varepsilon)\right\rangle_b\right)=P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\uplus\left\langle[\beta,\varepsilon)\right\rangle_b\right)=P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\right)+P_b\left(\left\langle[\beta,\varepsilon)\right\rangle_b\right)$.

Por proposiciones 9.1.1 y 9.1.2 y la definición 9.1.3 se deduce que $\alpha\left\langle[\delta,\gamma)\uplus[\beta,\varepsilon)\right\rangle_b \overset{\substack{\text{Proposición}\\ 9.1.2}}{=} \alpha\left(\left\langle[\delta,\gamma)\right\rangle_b\uplus\left\langle[\beta,\varepsilon)\right\rangle_b\right) \overset{\substack{\text{Proposición}\\ 9.1.1}}{=} \alpha\left(\bigcup_{n\in\mathbb{Z}}b^n[\delta,\gamma)\uplus\bigcup_{n\in\mathbb{Z}}b^n[\beta,\varepsilon)\right)=$ $=\bigcup_{n\in\mathbb{Z}}b^n\alpha[\delta,\gamma)\uplus\bigcup_{n\in\mathbb{Z}}b^n\alpha[\beta,\varepsilon) \overset{\substack{\text{Proposición 9.1.1}\\ \text{Definición 9.1.3}}}{=} \left\langle\alpha[\delta,\gamma)\right\rangle_b\uplus\left\langle\alpha[\beta,\varepsilon)\right\rangle_b$. Entonces aplicando lo probado antes se concluye que, $P_b\left(\alpha\left(\left\langle[\delta,\gamma)\uplus[\beta,\varepsilon)\right\rangle_b\right)\right)=$
$=P_b\left(\left\langle\alpha[\delta,\gamma)\right\rangle_b\uplus\left\langle\alpha[\beta,\varepsilon)\right\rangle_b\right)=P_b\left(\left\langle\alpha[\delta,\gamma)\right\rangle_b\right)+P_b\left(\left\langle\alpha[\beta,\varepsilon)\right\rangle_b\right)=$
$=P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\right)+P_b\left(\left\langle[\beta,\varepsilon)\right\rangle_b\right)=P_b\left(\left\langle[\delta,\gamma)\right\rangle_b\uplus\left\langle[\beta,\varepsilon)\right\rangle_b\right)=P_b\left(\left\langle[\delta,\gamma)\uplus[\beta,\varepsilon)\right\rangle_b\right)$. Por lo tanto, $P_b\left(\alpha\left\langle[\delta,\gamma)\cup[\beta,\varepsilon)\right\rangle_b\right)=P_b\left(\left\langle[\delta,\gamma)\cup[\beta,\varepsilon)\right\rangle_b\right)$ para todos los intervalos $[\delta,\gamma)\subset[1,b)$ y $[\beta,\varepsilon)\subset[1,b)$.

Análogamente podemos probar que $P_b\left(\alpha\left\langle\bigcup_{i=1}^{k}[\delta_i,\gamma_i)\right\rangle_b\right)=P_b\left(\left\langle\bigcup_{i=1}^{k}[\delta_i,\gamma_i)\right\rangle_b\right)$ para todos los intervalos $[\delta_i,\gamma_i)\subset[1,b)$ para $i=1...k$.

De acuerdo a la observación 9.2.4, trabajando análogamente podemos concluir que si $J_i\subset[1,b)$, J_i un intervalo, se verifica que $P_b\left(\alpha\left\langle\bigcup_{i=1}^{k}J_i\right\rangle_b\right)=P_b\left(\left\langle\bigcup_{i=1}^{k}J_i\right\rangle_b\right)$.

Se definen P y Q dos funciones de probabilidad en $\left([1,b),\mathcal{B}\left([1,b)\right)\right)$ tales que $P(E)=P_b\left(\alpha\langle E\rangle_b\right)$ y $Q(E)=P_b\left(\langle E\rangle_b\right)$ con $E\in\mathcal{B}\left([1,b)\right)$.

Sea $I \in \Im$, entonces por lo probado antes se cumple que $P_b\left(\alpha\langle I\rangle_b\right) = P_b\left(\langle I\rangle_b\right)$. Por lo tanto, $P(I) = Q(I)$ para todo $I \in \Im$. Además, $\Im$ es cerrado por intersecciones finitas y por complementos y $\langle \Im \rangle = \mathrm{B}\left([1,b)\right)$. Por el Lema 9.3.1 concluimos que $P=Q$ en $\mathrm{B}\left([1,b)\right)$. Por lo tanto, $P_b\left(\alpha\langle E\rangle_b\right) = P_b\left(\langle E\rangle_b\right)$ para todo $E \in \mathrm{B}\left([1,b)\right)$ por lo que se concluye que P_b es invariante frente a cambios de escala.

Resta ver que P_b es la única probabilidad invariante frente a cambios de escala en $(\mathbb{R}^+, \mathcal{M}_b)$.

Supongamos que existe R_b una función de probabilidad invariante frente a cambios de escala en $(\mathbb{R}^+, \mathcal{M}_b)$.

Sea R una función de probabilidad en $\left([1,b), \mathrm{B}\left([1,b)\right)\right)$ tal que $R(E) = R_b\left(\langle E\rangle_b\right)$ con $E \in \mathrm{B}\left([1,b)\right)$. R es invariante por escala ya que $R(\alpha E) = R_b\left(\langle \alpha E\rangle_b\right) \overset{\text{Definición 9.1.3}}{=} R_b\left(\alpha\langle E\rangle_b\right) = R_b\left(\langle E\rangle_b\right) = R(E)$.

Sea R una función de probabilidad en $\left([1,10), \mathrm{B}\left([1,10)\right)\right)$ tal que $R(E) = R\left(g(E)\right)$ con $E \in \mathrm{B}\left([1,10)\right)$, siendo $g : \mathrm{B}\left([1,10)\right) \to \mathrm{B}\left([1,b)\right)$, $g(x) = b^{\log_{10} x}$. R es una función de probabilidad puesto que g es una función continua y su inversa $g^{-1} : \mathrm{B}\left([1,b)\right) \to \mathrm{B}\left([1,10)\right)$, $g(x) = 10^{\log_b x}$ también es continua, por lo tanto, transforman borelianos en borelianos.

Afirmación: R es invariante frente a cambios de escala en $\left([1,10), \mathrm{B}\left([1,10)\right)\right)$. Es decir, $R(\alpha E) = R(E)$ para $E \in \mathrm{B}\left([1,10)\right)$.

Demostración.

Supongamos α como antes y $\alpha < 10$.

Por definición $g(\alpha E) = \left\{b^{\log_{10}\alpha x} : x \in E\right\} = \left\{b^{\log_{10}\alpha}.\, b^{\log_{10} x} : x \in E\right\} = b^{\log_{10}\alpha}.g(E)$, entonces $R(\alpha E) = R\left(g(\alpha E)\right) = R\left(b^{\log_{10}\alpha}.g(E)\right)$. Como R es invariante frente a cambios de escala y $b^{\log_{10}\alpha} < b$ puesto que $\alpha < 10$, $R\left(b^{\log_{10}\alpha}.g\left(E\right)\right) = R\left(g\left(E\right)\right)$. Por lo tanto, $R\left(\alpha E\right) = R\left(E\right)$.

Como R es invariante frente a cambios de escala en ($\left[1,10\right), \mathrm{B}\left(\left[1,10\right)\right)$) por teorema 3.2.3.1, R verifica la Ley de Benford en ($\left[1,10\right), \mathrm{B}\left(\left[1,10\right)\right)$). Entonces, $R\left(\left[1,\gamma\right)\right) = \log_{10}\gamma$ con $\gamma \in \left[1,10\right)$.

Por definición $R\left(E\right) = R\left(g\left(E\right)\right)$ con $E \in \mathrm{B}\left(\left[1,10\right)\right)$ entonces $R\left(\left[1,\gamma\right)\right) = R\left(g^{-1}\left(\left[1,\gamma\right)\right)\right) = R\left(\left[1,10^{\log_b \gamma}\right)\right) = \log_{10}\left(10^{\log_b \gamma}\right) = \log_b \gamma$ con $\gamma \in \left[1,b\right)$. Por lo tanto, R sigue la Ley de Benford y por definición de R, R_b sigue la Ley de Benford.

Entonces, $R_b = P_b$ en ($\mathbb{R}^+$, $\mathcal{M}_b$). Concluyéndose así que P_b es la única probabilidad invariante frente a cambios de escala en ($\mathbb{R}^+$, $\mathcal{M}_b$). □.

Capítulo 10. Aplicaciones

10.1 Testeo de modelos matemáticos

Se ha utilizado la Ley de Benford para chequear los pronósticos de modelos matemáticos. La idea subyacente a esta aplicación es que si los datos empíricos satisfacen la Ley de Benford, los resultados de los "buenos" modelos matemáticos también debieran hacerlo. Algunos modelos que fueron testeados utilizando la Ley de Benford refieren a producción económica, índice de acciones en la bolsa, cantidad de tierra empleada para diferentes usos y chequeo de listas de fallos para la detección de errores sistémicos.

10.2 Detección de fraude en la fabricación de datos en documentos financieros

El uso de la Ley de Benford para detectar fraude o falsificación de datos en las declaraciones de impuestos fue abordado por el contador norteamericano Mark Nigrini en sus tesis de doctorado. Nigrini ha recogido una extensa cantidad de evidencia empírica de la ocurrencia de la Ley de Benford en muchas áreas de la contabilidad y de la demografía y concluyó que en una amplia y variada cantidad de situaciones contables los datos reales siguen la Ley de Benford. Sin embargo, al falsear datos estos deben inventarse, lo que puede hacerse tomando números uniformemente generados o mediante una elección personal del falseador. Nigrini ha diseñado varios tests para medir la conformidad de un conjunto de datos con la Ley de Benford.

10.3 Juegos de azar

En Hill ([4], pág. 326) se describe como se ha utilizado la Ley de Benford para ganar dinero en el Massachusetts Numbers Game. Allí los jugadores primero apuestan a un número de cuatro cifras de su elección, luego, un número de cuatro cifras es generado al azar siguiendo una distribución uniforme, y luego todos los jugadores con ganancias reparten el pozo de forma equitativa. En esa situación es ventajoso identificar números que sean poco elegidos, ya que los todos los números tienen la misma probabilidad de ganar y los números poco elegidos generarán más ganancias a los apostadores. Si las personas se basan en sus experiencias personales para elegir el número a apostar, y si los números de su experiencia están distribuidos siguiendo la Ley de Benford, entonces se deben elegir números que comiencen por 8 o 9 para maximizar las posibles ganancias.

Apéndice

En la siguiente tabla se muestra la población total de las localidades uruguayas con más de 1.000 habitantes según el censo realizado en 2011 según el Instituto Nacional de Estadística de Uruguay.

Departamento	Localidad	Total	Departamento	Localidad	Total
Artigas	Artigas	40.657	Florida	Sarandí Grande	6.130
Artigas	Baltasar Brum	2.531	Florida	Veinticinco de Agosto	1.849
Artigas	Bella Unión	12.200	Florida	Veinticinco de Mayo	1.852
Artigas	Las Piedras	2.771	Flores	Trinidad	21.429
Artigas	Pintadito	1.642	Lavalleja	José Batlle y Ordóñez	2.203
Artigas	Sequeira	1.149	Lavalleja	José Pedro Varela	5.118
Artigas	Tomás Gomensoro	2.659	Lavalleja	Mariscala	1.626
Canelones	Aguas Corrientes	1.047	Lavalleja	Minas	38.446
Canelones	Atlántida	5.562	Lavalleja	Solís de Mataojo	2.825
Canelones	Barra de Carrasco	5.410	Maldonado	Aiguá	2.465
Canelones	Barros Blancos	31.650	Maldonado	Balneario Buenos Aires	1.551
Canelones	Canelones	19.865	Maldonado	Barrio Hipódromo	1.973
Canelones	Cerrillos	2.508	Maldonado	Cerro Pelado	8.177
Canelones	City Golf	1.104	Maldonado	El Tesoro	1.396
Canelones	Colinas de Solymar	2.813	Maldonado	La Capuera	2.838
Canelones	Colonia Nicolich	9.624	Maldonado	La Sonrisa	1.562
Canelones	Dr. Francisco Soca	1.797	Maldonado	Maldonado	62.590
Canelones	El Pinar	21.091	Maldonado	Pan de Azúcar	6.597
Canelones	Empalme Olmos	4.199	Maldonado	Pinares - Las Delicias	9.819
Canelones	Estación Atlántida	2.274	Maldonado	Piriápolis	8.830
Canelones	Fracc. Cno.del Andaluz y R.84	9.295	Maldonado	Playa Grande	1.031
Canelones	Fracc. sobre R.74	1.513	Maldonado	Punta del Este	9.277
Canelones	Joaquín Suárez	6.570	Maldonado	San Carlos	27.471
Canelones	Juanicó	1.305	Maldonado	San Rafael - El Placer	3.146
Canelones	La Floresta – Estación	1.313	Maldonado	Villa Delia	1.703
Canelones	La Floresta	1.595	Montevideo	Montevideo	1.304.729
Canelones	La Paz	20.524	Paysandú	Chacras de Paysandú	3.965
Canelones	Lagomar	8.066	Paysandú	Guichón	5.039
Canelones	Las Piedras	71.258	Paysandú	Nuevo Paysandú	8.578
Canelones	Las Toscas	3.146	Paysandú	Paysandú	76.412
Canelones	Lomas de Solymar	19.124	Paysandú	Piedras Coloradas	1.094
Canelones	Marindia	3.543	Paysandú	Porvenir	1.159
Canelones	Migues	2.109	Paysandú	Quebracho	2.853
Canelones	Montes	1.760	Paysandú	San Félix	1.718
Canelones	Neptunia	4.774	Paysandú	Tambores	1.561
Canelones	Pando	25.947	Rivera	La Pedrera	3.363
Canelones	Parque Carrasco	8.628	Rivera	Lagunón	2.376
Canelones	Parque del Plata	7.896	Rivera	Mandubí	6.019
Canelones	Paso de Carrasco	15.908	Rivera	Minas de Corrales	3.788
Canelones	Pinamar – Pinepark	4.724	Rivera	Rivera	64.465
Canelones	Progreso	16.244	Rivera	Santa Teresa	2.657
Canelones	Salinas	8.626	Rivera	Tranqueras	7.235

Canelones	San Antonio	1.489	Rivera	Vichadero	3.698
Canelones	San Bautista	1.973	Río Negro	Fray Bentos	24.406
Canelones	San Jacinto	4.510	Río Negro	Nuevo Berlín	2.450
Canelones	San José de Carrasco	7.288	Río Negro	San Javier	1.781
Canelones	San Luis	1.878	Río Negro	Young	16.756
Canelones	San Ramón	7.133	Rocha	Castillos	7.541
Canelones	Santa Lucía	16.742	Rocha	Cebollatí	1.609
Canelones	Santa Rosa	3.727	Rocha	Chuy	9.675
Canelones	Sauce	6.132	Rocha	La Aguada - Costa Azul	1.090
Canelones	Shangrilá	3.195	Rocha	La Paloma	3.495
Canelones	Solymar	18.573	Rocha	Lascano	7.645
Canelones	Tala	5.089	Rocha	Rocha	25.422
Canelones	Toledo	4.397	Rocha	Velázquez	1.022
Canelones	Villa Aeroparque	4.307	San José	Delta del Tigre y Villas	20.239
Canelones	Villa Crespo y San Andrés	9.813	San José	Ecilda Paullier	2.585
Canelones	Villa Felicidad	1.344	San José	Libertad	10.166
Canelones	Villa San José	1.419	San José	Monte Grande	1.287
Cerro Largo	Aceguá	1.511	San José	Playa Pascual	6.870
Cerro Largo	Fraile Muerto	3.168	San José	Puntas de Valdez	1.491
Cerro Largo	Isidoro Noblía	2.331	San José	Rafael Perazza	1.277
Cerro Largo	Melo	51.830	San José	Rodríguez	2.604
Cerro Largo	Río Branco	14.604	San José	Safici (Parque Postel)	1.087
Cerro Largo	Tupambaé	1.122	San José	San José de Mayo	36.743
Colonia	Carmelo	18.041	San José	Santa Mónica	1.662
Colonia	Colonia del Sacramento	26.231	Salto	Belén	1.926
Colonia	Colonia Valdense	3.235	Salto	Constitución	2.762
Colonia	Florencio Sánchez	3.716	Salto	Salto	104.011
Colonia	Juan Lacaze	12.816	Soriano	Cardona	4.600
Colonia	Nueva Helvecia	10.630	Soriano	Chacras de Dolores	1.961
Colonia	Nueva Palmira	9.857	Soriano	Dolores	17.174
Colonia	Ombúes de Lavalle	3.390	Soriano	José Enrique Rodó	2.120
Colonia	Rosario	10.085	Soriano	Mercedes	41.974
Colonia	Tarariras	6.632	Soriano	Palmitas	2.123
Durazno	Blanquillo	1.084	Soriano	Villa Soriano	1.124
Durazno	Carmen	2.692	Tacuarembó	Ansina	2.712
Durazno	Centenario	1.136	Tacuarembó	Curtina	1.037
Durazno	Cerro Chato	2.818	Tacuarembó	Las Toscas	1.142
Durazno	Durazno	34.368	Tacuarembó	Paso de los Toros	12.985
Durazno	La Paloma	1.443	Tacuarembó	San Gregorio de Polanco	3.415
Durazno	Santa Bernardina	1.094	Tacuarembó	Tacuarembó	54.755
Durazno	Sarandí del Yí	7.176	Treinta y Tres	Ejido de Treinta y Tres	6.782
Florida	Alejandro Gallinal	1.357	Treinta y Tres	Gral. Enrique Martínez	1.430
Florida	Casupá	2.402	Treinta y Tres	Santa Clara de Olimar	2.341
Florida	Florida	33.639	Treinta y Tres	Treinta y Tres	25.477
Florida	Fray Marcos	2.398	Treinta y Tres	Vergara	3.810
Florida	Nico Pérez	1.030	Treinta y Tres	Villa Sara	1.199

Bibliografía y referencias

[1] Benford, F. (1938) *The law of anomalous numbers*, Proc. Amer. Phil. Soc. 78, 551-572.

[2] Hill, TP (1995). *A Statistical Derivation of the Significant-Digit Law*. Statistical Science 10(4), 354-363.

[3] Hill, TP (1995). *Base-Invariance Implies Benford's Law*. Proceedings of the American Mathematical Society 123(3), 887-895

[4] Hill, TP (1995). *The Significant-Digit Phenomenon*. American Mathematical Monthly 102(4), 322-327.

[5] Hill, TP (1998). *The First-Digit Phenomenon*. American Scientist 86 (4), 358-363.

[6] Instituto Nacional de Estadística. (s.f.). *Instituto Nacional de Estadística Uruguay*. Recuperado el 14 de Septiembre de 2015, de http://www.ine.gub.uy/

[7] Nectoux, A. (12 de febrero de 2012). *Blog Proyecto Klein*. Recuperado el 11 de junio de 2015, de Ley de Benford:¿aprender a defraudar o a detectar fraude?: http://blog.kleinproject.org/?p=1634&lang=es

[8] Newcomb, S. (1881). *Note on the frecuency of use of the diferent digits in natural numbers.* American Journal of Mathematics , 4 (1/4), 39-40.

[9] Nigrini, M. (2012). *Benford's Law Applications for Forensic Accounting, Auditing, and Fraud Detection* . Wiley.

[10] Nigrini, M. (2011). *Forensic Analytics Methods and Techniques for Forensic Accounting Investigations.* Wiley.

[11] Perera, M., & Ayllón, J. (1999). *El primer dígito significativo*. EPSILON. SAEM Thales, 339-349.

[12] Petrov, V., & Mordecki, E. (2003). *Teoría de Probabilidades*. Moscú: Editorial URSS.

[13] Rousseau, C. (2010). Apprendre à frauder ou à détecter les fraudes? Accromath. 2-7.

[14] Spiegel, M. (1991). *Estadística*. Madrid: McGraw Hill.