Premios del Departamento de Matemáticas de la Universidad Autónoma de Madrid para Estudiantes de Secundaria

Primera Edición, 2006/2007

TRABAJO: Análisis del cumplimiento de la ley de Benford en el censo de la población española

MENCIÓN ESPECIAL (POR LA ORIGINALIDAD DEL TEMA TRATADO)

AUTORES:

- Sonsoles Vicente Ayuso
- Camino Vicente Ayuso

TUTORES:

- Enrique Vilches

CENTRO:

Análisis del cumplimiento de la Ley de Bendford en el censo de la población española

Bernoulli

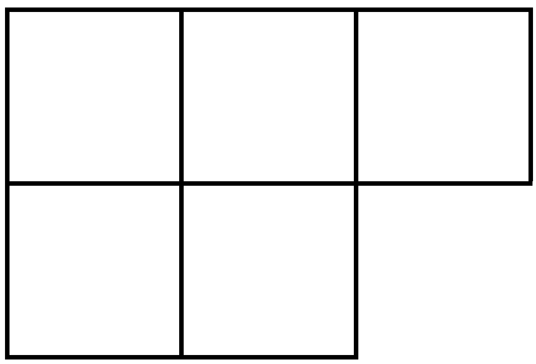
¡Por fin era Navidad! Y como todos los años, estábamos allí reunidos todos los primos en torno al abuelito que como siempre había sacado su libro.

Era un libro de color rojo, apaisado, con el papel ya amarillento, pero que siempre nos había parecido que contenía todo el saber de la tierra.

Estábamos mi primo Javier que es el mayor de todos tiene 23 años y es ingeniero de Telecomunicaciones, mi primo Víctor que está estudiando cuarto de Derecho y Económicas, mi hermana Carmen que estudia primero de Medicina, mis dos primas pequeñas Pilar y Ana que tienen 8 y 10 años y nosotras dos Sonsoles y Camino, y como no en el centro de todos nuestro abuelo y a su lado nuestra abuela, que ya tienen más de 85 años. Y en medio de la mesa de formica amarilla del comedor, estaba el libro de nuestro abuelo.

Según comentó nuestra abuela, no estaba escrito con bolígrafo, ni a pluma "Parker" de las que conoció nuestro padre, estaba escrito con "plumines" y "tintero", de los que todavía conserva algunos en un estuche de cuando él estudiaba. Y fue escrito cuando tenía la edad nuestra y estudiaba en el Colegio de San José de Valladolid. Lo cierto es que tiene una letra muy bonita y clara y que mi abuela dice que se llama "redondilla", y que la verdad nosotros no hemos visto nunca en el ordenador y sólo la hemos visto en ese libro.

Nuestro abuelo miró a nuestras primas pequeñas y les dijo: un problema de geometría. Cogió una caja de cerillas y dibujó lo siguiente:



Les dijo: con 15 cerillas hemos creado 5 cuadrados, tenéis que quitar 3 cerillas y que nos queden sólo 3 cuadrados.

Nos miró a los demás y nos dijo: ahora un poco de Matemáticas y nos preguntó: ¿cuál es la probabilidad de que si lanzamos una moneda al aire salga cara? Rápidamente contestamos ½, a lo que mi primo puntualizó pero eso sólo sucede si lanzamos muchas veces la moneda, ya que hay que emplear la "Ley de los Grandes Números" que dice que "cuando un fenómeno aleatorio se repite un elevado número de veces, la frecuencia relativa con que aparece un suceso se acerca a un determinado número, que depende de cada suceso y que se denomina probabilidad del mismo y que es:

$$p = \frac{\textit{Número de casos favorables}}{\textit{Número de casos posibles}}$$

Nuestro abuelo volvió a leer ¿y la probabilidad de que si lanzamos un dado salga un uno?

Rápidamente contestamos 1/6.

Correcto contestó nuestro abuelo si el dado tiene 6 caras y no está trucado. Y prosiguió: ahora un poco más difícil: tenemos tres cartas una es amarilla por los dos lados otra verde por los dos lados y la tercera es amarilla por un lado y verde por el otro, las guardamos en una bolsa y sacamos una y vemos que es de color amarillo de ¿qué color será la otra cara de la carta? ¿cuál es la probabilidad de que la otra cara sea también amarilla?

Nosotras nos miramos sin saber que decir, nuestra hermana Carmen dijo: amarillo, mi primo Javier la interrumpió y dijo hay un 50% de probabilidades de que sea verde y otro 50% de que sea amarilla porque hay dos cartas que pueden estar en estas condiciones, una la amarillo-amarillo y otra la amarillo-verde. Por lo tanto, la probabilidad es un 50%.

Nuestro abuelo continuó leyendo: la probabilidad de que sea amarilla la otra cara de una carta que sabemos que es amarilla es $\frac{2}{3}$ y de que sea verde un $\frac{1}{3}$.

Nuestro primo le cortó diciendo: no puede ser, ese libro está mal, en la carrera le habían enseñado casos favorables/casos posibles y que era un caso favorable frente a dos posibles por lo tanto un $\frac{1}{2}$.

Mi abuelo continuó: estás en un error, la probabilidad de que la carta tenga el mismo color en ambas caras es $\frac{2}{3}$ (dos casos frente a las tres cartas) por lo tanto la probabilidad de que, sabiendo que una carta tiene como color de una cara el amarillo, la otra sea amarilla es $\frac{2}{3}$, y de que siendo verde una cara la otra también sea verde es también de $\frac{2}{3}$.

Nuestro primo no se quedó muy conforme y prosiguió un rato la discusión. Y nuestro abuelo intentó explicarle que efectivamente la probabilidad de que salga cada carta es $\frac{1}{3}$, pero deja de ser así cuando disponemos de determinada información como es que sabemos uno de los colores de la carta, en ese momento la probabilidades cambian ya que hay dos cartas que cumplan esta condición y cuatro caras posibles y tres son de un color y sólo una del otro color, o lo que es lo mismo la carta con el mismo color tiene más posibilidades de ser la elegida ya que ambas caras son ganadoras, mientras que la carta con las dos caras de distinto color sólo tiene una solución ganadora, y otra perdedora. Nuestro primo no quedó nada convencido.

Bueno ahora todavía un poco más difícil, volvió a decir nuestro abuelo: imaginaros ahora que nosotros conocemos la longitud de los ríos del mundo, mi prima pequeña ponía cada vez más cara de asombro, creemos que pensaba: pero si yo sólo conozco los ríos de mi Comunidad Autónoma, ¿cómo es que mi abuelo conoce los de todo el mundo?. ¿os lo imagináis?, imaginaros que miden 10.235, 5.840, 124.543 etc. ¿cuál es la probabilidad de que el primer dígito sea un 1?. Mi prima Ana preguntó: ¿abuelo qué es el primer dígito?.

De todas formas, aunque ya hablaremos otro día más sobre los dígitos, quiero que busquéis su significado en el Diccionario uno de vosotros y nos lea la definición y durante unos días penséis y observéis noticias o ideas relacionadas con los dígitos porque vosotros vais a vivir en la era digital. ¡Ah! Exclamó uno de los nietos pequeños ¿eso de digital tiene que ver con la tele? ¿O con los relojes? ¿O con las huellas digitales? La que había buscado en el diccionario nos leyó: Deriva de la palabra latina “digitum” = dedo, y se dice del número que el sistema de numeración decimal se expresa con una sola cifra.

Nuestro abuelo le miró con ojos dulces y le dijo: por lo tanto, el primer dígito es el primer número que tiene la cifra, por ejemplo cuando hemos dicho 10.235, el primer dígito es el 1, cuando hemos dicho 5.840, es el 5, el segundo es el 8 y así sucesivamente.

Nosotras, mientras estaba dando la explicación, estábamos pensando la solución de la adivinanza, pero no estábamos seguras, no sabemos la longitud de casi ningún río, como para saber cual el la probabilidad de que el primer dígito sea un 1.

No había terminado la contestación cuando mi primo dijo: abuelo la probabilidad de que el primer dígito sea un 1 es $\frac{1}{9}$, casos favorables entre casos posibles.

Nuestro abuelo le miró y dijo error.

Mi primo se sorprendió y dijo: tu libro, abuelo, está equivocado, no puede ser. Según las leyes de la estadística la probabilidad de que la lotería termine en 7 como este año es un 10%, la probabilidad de que termine en “67” es un 1%, y por tanto la probabilidad de que el primer dígito de la longitud de los ríos del mundo sea un 1 es un $\frac{1}{9}$.

Nuestro abuelo con la paciencia que tienen los abuelos le dijo, la probabilidad de que el primer dígito de la longitud de los ríos del mundo sea un 1 es aproximadamente un 30%.

Mi primo le replicó: imposible abuelo, tu libro está equivocado. Tienes que rehacer tu libro, yo no estoy de acuerdo.

Nuestra abuela miraba a nuestro primo con la misma cara que nuestras primas pequeñas miraban al abuelo, con esos ojos de admiración de la abuela que tiene un nieto Ingeniero, y con cara de duda hacia el abuelo, tal vez pensando, llevará razón mi nieto, antes no se sabían tantas cosas como ahora.

El abuelo prosiguió mirando a nuestras primas pequeñas, hace muchos, muchos años hubo un señor que se llamaba Benford. Frank Benford era muy pero que muy observador y era profesor. En 1938 se dio cuenta que en un libro que se llamaba las tablas de logaritmos, que ellos utilizaban mucho pues en aquellos años no había ordenadores, estaban mucho más desgastadas las primeras páginas y que cuando las hojas avanzaban estaban menos usadas. Benford que vivió a principios del siglo XX cuando no existían calculadoras, ni móviles, ni esas máquinas para jugar que tenéis ahora, como era tan observador pensó ¿por qué la gente joven necesita encontrar con más frecuencia los logaritmos de los números que empiezan por 1 o por 2 y menos los que empiezan por 8 o por 9?

Estudió mucho y mucho, y llegó a la conclusión de que la probabilidad de que el primer dígito en las series de números de la naturaleza sea un 1 es $P(1) = \text{Log}_{10}(1+1/1) = \text{Log}_{10}(2) = 0{,}3010$, de que sea un 2 es $P(2) = \text{Log}_{10}(1+\frac{1}{2}) = \text{Log}_{10}(1{,}5) = 0{,}1761$ y así sucesivamente, y nos leyó la formula de su libro:

La probabilidad de que el primer dígito de una serie de números sea (n) no es 1/número de dígitos, sino que es $P(n) = \text{Log}_{10}(1+\frac{1}{n})$; con $n \in \{1, 2, \ldots, 9\}$. Mi primo Javier le comentó al abuelo entonces $\sum_{n=1}^{9} Log_{10}(1+\frac{1}{n}) = 1$?. El abuelo dijo si sabes bien las propiedades de los logaritmos te será fácil hacer la comprobación. Cogió enseguida papel y lápiz y escribió:

$$\sum_{n=1}^{9} Log_{10}(1+\frac{1}{n}) = Log_{10}(\prod_{n=1}^{9}(1+\frac{1}{n})) = Log_{10}(\prod_{n=1}^{9}\frac{n+1}{n}) = Log_{10}(\frac{2}{1}\cdot\frac{3}{2}\cdot....\cdot\frac{9}{8}\cdot\frac{10}{9}) = Log_{10}10 = 1$$

Y mi primo Víctor que en Económicas también estudia Matemáticas le dijo al abuelo: entonces si la P(n) = Log_{10} (1+$\frac{1}{n}$) y Benford observó que los números que empezaban por 1 eran más frecuentes que los que empezaban por 2 y así sucesivamente la función $y = Log_{10}(1+\frac{1}{x})$ debe ser decreciente para x $\in\{1, 2, \dots , 9\}$. El abuelo le animó a que lo comprobara y cogiendo también papel y lápiz empezó a escribir:

$$y' = \frac{-1/x^2}{1+1/x} Log_{10}e = -\frac{1}{x(x+1)} Log_{10}e$$

Dudó un poco de la fórmula de la derivada por eso de ser logaritmos en base 10 en lugar de logaritmos neperianos que esa la sabía bien, enseguida pensó en buscarla en un libro pero como eran Navidades y estaban en casa de los abuelos Carmen le confirmó que si estaba bien. Efectivamente la derivada era negativa para x $\in\{1, 2, \dots , 9\}$.

Mis primas Ana y Pilar miraban con cara de asombro, como pensando cuanto saben mis abuelos y los primos.

¿Qué son las series de números de la naturaleza? Le preguntamos nosotras.

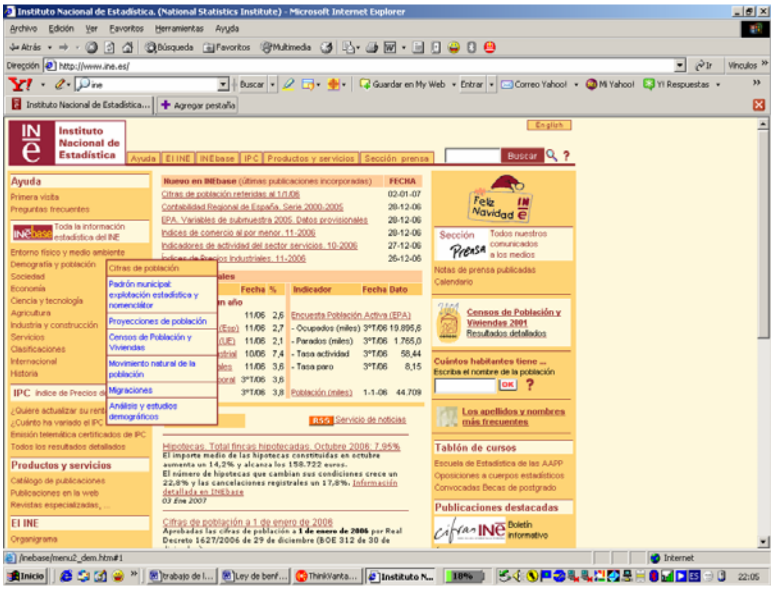
Nuestro abuelo nos miró y nos dijo: En la vida hay serie de números aleatorios, como por ejemplo si cogemos los números de la lotería, y mirando a nuestras primas dijo, esos de ese anuncio que vosotros cantáis cada vez que sale en la televisión, o la serie de números cuando tiramos un dado al aire cuando jugamos al parchís, o a la oca. Hay otra serie de números que no son aleatorios, como por ejemplo los números de los teléfonos, ya que van creados por el hombre y siguen una serie de reglas como por ejemplo los de Ávila empiezan por 920, los de Madrid empiezan por 91, los de los móviles, generalmente empiezan por 6 y así sucesivamente. Tampoco son aleatorios los precios de los distintos productos, ya que es más frecuente que un vestido valga 10,99€ que 11€. Por último existen otra serie de números que son las series de números de la naturaleza, como por ejemplo la longitud de los ríos del mundo, o la longitud de los ferrocarriles de la tierra, o la población de las ciudades o el número de hombres y mujeres de todos los municipios de España.

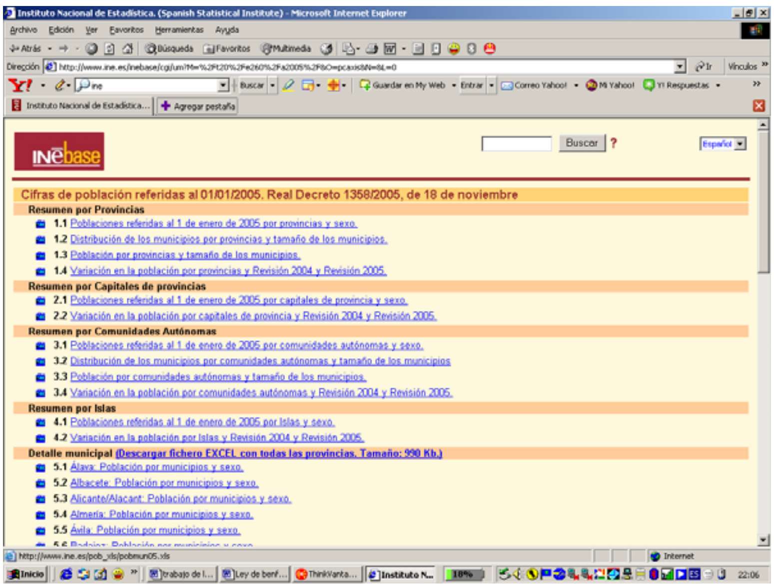
No cumplen la ley de Benford continuó nuestro abuelo los números aleatorios, ni tampoco, por ejemplo, la edad de las personas que tienen derecho a votar en las elecciones ya que sólo se puede votar a partir de 18 años, y además es muy difícil llegar a viejo, o la edad de los reclusos, ya que estos no pueden ser menores de 18 ni mayores de 70 años, ni lo cumplirían los ríos si exigiéramos que para ser un río fuera preciso que tuviera más de 5 kilómetros y menos de 80, ya que con ello estamos distorsionando los datos.

No puede ser replicó rápidamente nuestro primo Javier, eso es imposible, no puede ser que si cogemos los municipios de España y contamos los que el primer dígito sea un 1 en hombres ese número coincida con el de municipios en los que el primer dígito de mujeres sea también un 1 y además eso tiene que suponer un 11,11% un caso favorable frente a 9 posibles y no un 30,1% como tu dices.

Nuestro abuelo nos miró a nosotras y nos dijo tenéis que demostrar a vuestro primo que la población española cumple la ley de Benford. Así que nos dispusimos a ello:

Lo primero que hemos hecho, siguiendo los consejos de nuestro primo Víctor ha sido acudir a Internet y buscar la Página Web del Instituto Nacional de Estadística.





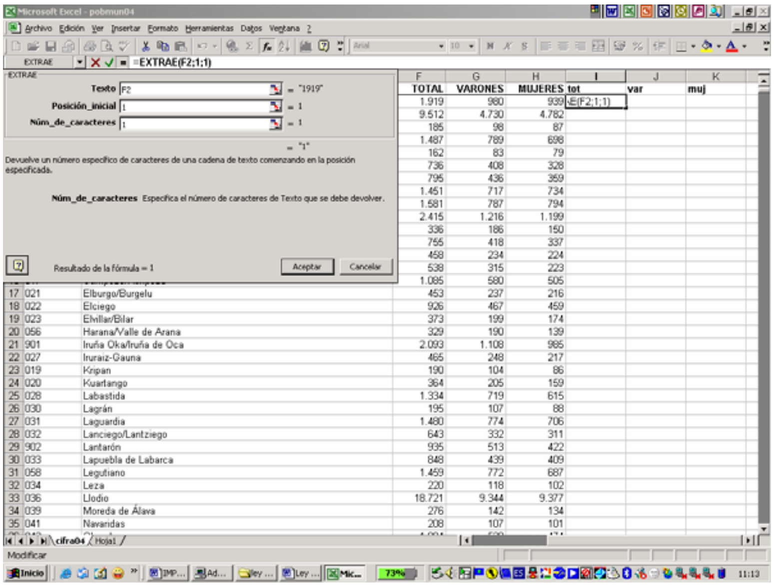
Entre las opciones que permite está bajar los datos como hoja de Excel, por lo que hemos conseguido la siguiente hoja de cálculo (se adjunta una muestra)

Cifras de población resultantes de la Revisión del Padrón municipal a 1 de enero de 2004						
Cód. Prov	**PROVINCIA**	**Cód Mun**	**NOMBRE**	**TOTAL**	**VARONES**	**MUJERES**
01	Álava	001	Alegría-Dulantzi	1.919	980	939
01	Álava	002	Amurrio	9.512	4.730	4.782
01	Álava	049	Añana	85	98	87
01	Álava	003	Aramaio	1487	789	698
01	Álava	006	Armiñón	162	83	79
01	Álava	037	Arraia-Maeztu	736	408	328
01	Álava	008	ArrazuaUbarrundia	795	436	359
01	Álava	004	Artziniega	1.451	717	734
01	Álava	009	Asparrena	1.581	787	794
01	Álava	010	Ayala/Aiara	2.415	1.216	1.199
01	Álava	011	Baños deEbro/Mañueta	336	186	150
01	Álava	013	Barrundia	755	418	337
01	Álava	014	Berantevilla	458	234	224
01	Álava	016	Bernedo	538	315	223
01	Álava	017	Campezo/Kanezu	1.085	580	505
01	Álava	021	Elburgo/Burgelu	453	237	216
01	Álava	022	Elciego	926	467	459
01	Álava	023	Elvillar/Bilar	373	199	174
01	Álava	056	Harana/Valle de Arana	329	190	139
01	Álava	901	Iruña Oka/Iruña de Oca	2.093	1.108	985
01	Álava	027	Iruraiz-Gauna	465	248	217
01	Álava	019	Kripan	190	104	86
01	Álava	020	Kuartango	364	205	159
01	Álava	028	Labastida	1.334	719	615
01	Álava	030	Lagrán	195	107	88

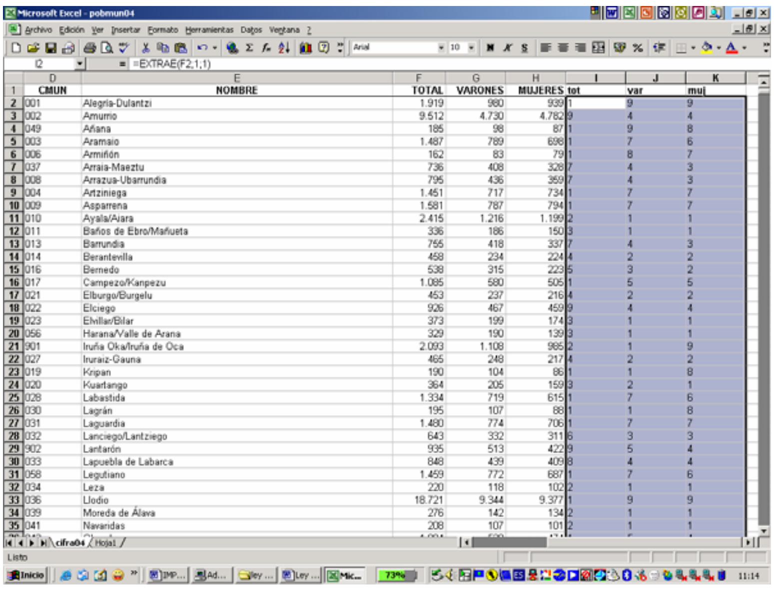
A continuación hemos extraído el primer dígito de la serie empleando la función de Excel Extrae:

EXTRAE (texto; posición_inicial; núm_de_caracteres)

En nuestro caso en cada fila tenemos que extraer del texto de las columnas 5, 6 y 7 la primera posición y un carácter, nos queda:

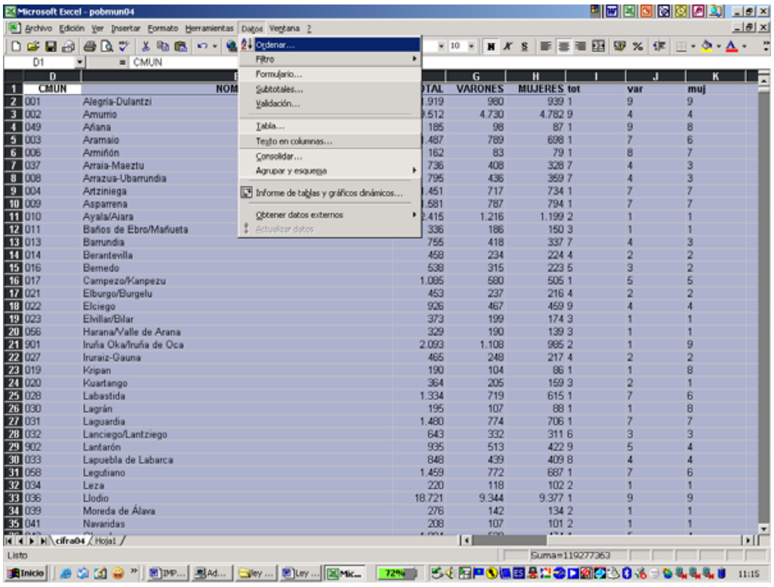


y el resultado es:

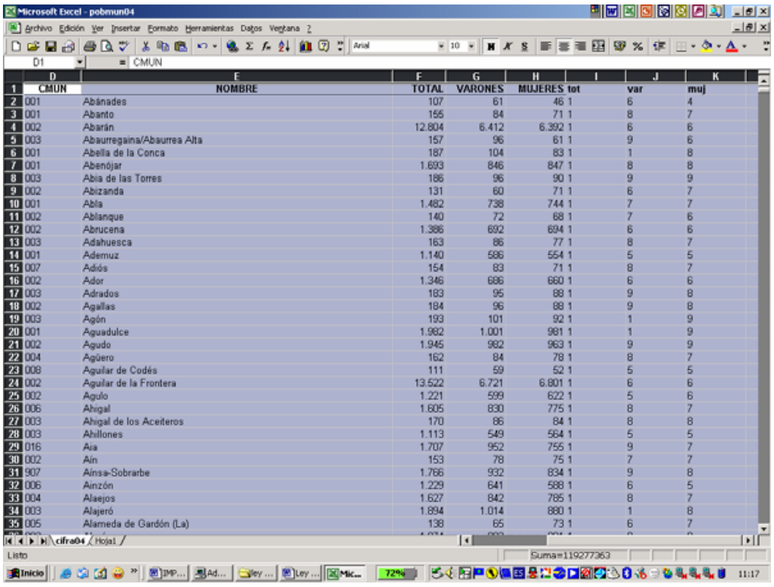


El siguiente paso es contar los casos en los que empieza por 1, por 2, por 3, etc. para ello lo primero que tenemos que hacer es ordenar los datos.

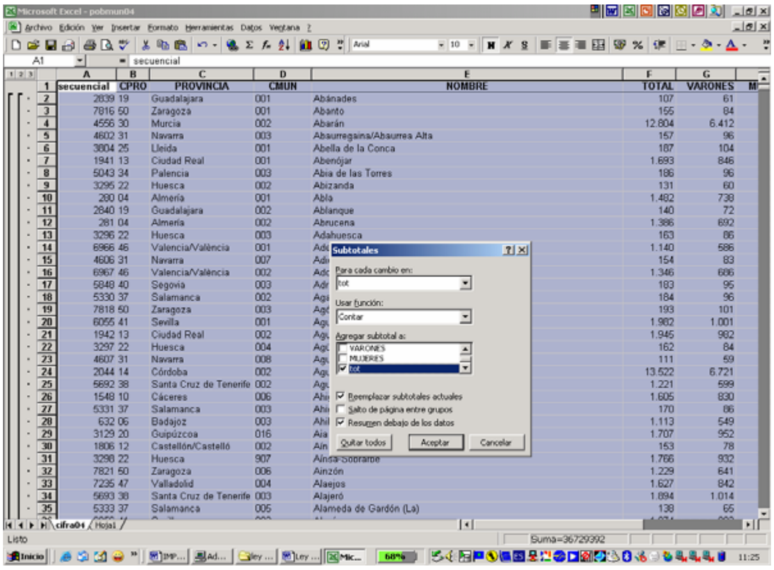
Usamos la opción de ordenar del menú datos:



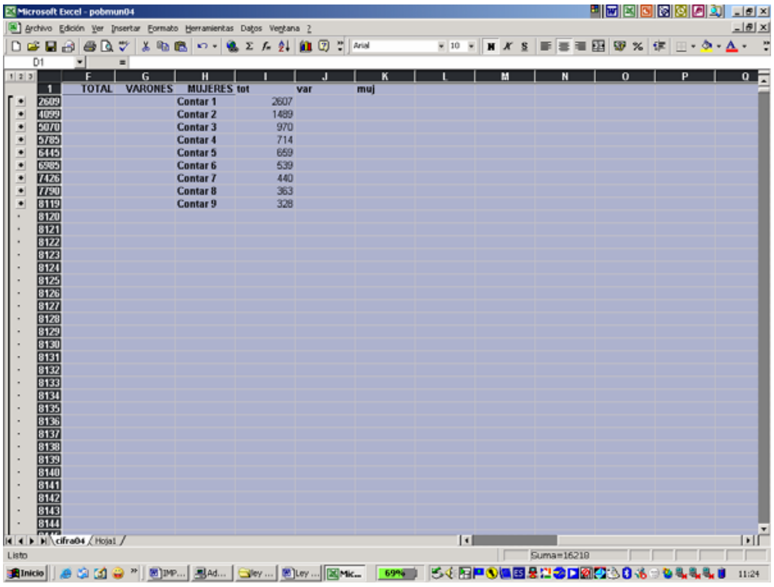
Seleccionamos primero la columna I y nos quedará:



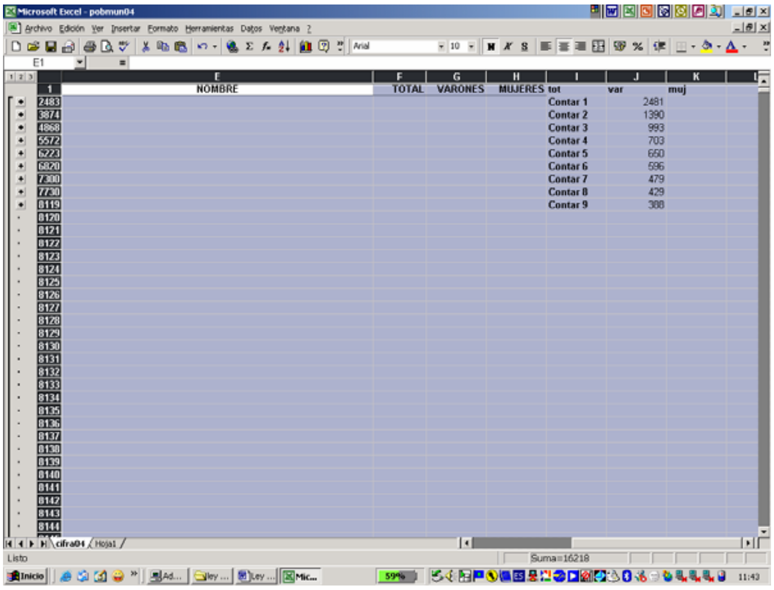
A continuación mediante el menú Datos, Subtotalización por cada cambio en la columna I, eligiendo la opción de contar:



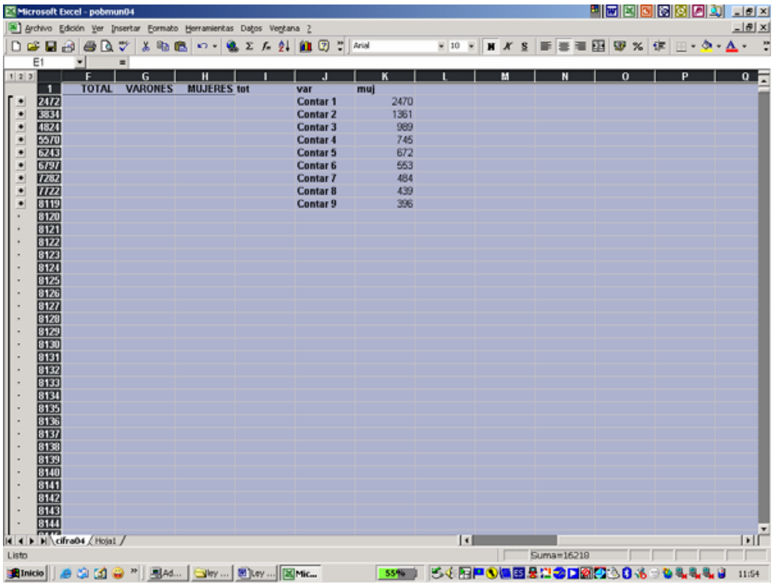
Si seleccionamos en los números que nos aparecen debajo de la casilla activa el número 2 nos aparecerá:



Estos datos nos los llevamos a una hoja nueva, y hacemos lo mismo con la columna J: es decir: datos, ordenar (por la columna J). Y después datos, subtotalizar (por cada cambio de la columna J, con la opción de contar), y luego seleccionamos el número 2 de las columnas que aparecen a la izquierda de la hoja nos resulta:



Llevamos este resultado a la hoja nueva y volvemos a realizar las mismas operaciones para la columna K que contiene el total de la población de cada municipio y nos da como resultado:



Si agrupamos en una hoja todos los resultados nos resulta:

Dígito	Total	Varones	Mujeres
1	2607	2481	2470
2	1489	1390	1361
3	970	993	989

4	714	703	745
5	659	650	672
6	539	596	553
7	440	479	484
8	363	429	439
9	328	388	396
Total	**8109**	**8109**	**8109**

Viendo estos resultados nos dimos rápidamente cuenta que nuestro abuelo tenía razón, que resulta que de los 8.109 municipios de España, hay 2.607 que tienen como primer dígito el 1, hay 1.489 municipios en los que el primer dígito es un 2 y tan sólo hay 328 municipios en los que el primer dígito sea un 9.

Pero necesitábamos calcular los porcentajes que representaban estos datos para ello emplearíamos las fórmulas de calcular el cociente de cada celda respecto al total de la columna:

=B2/B$11

	Dígito	Total	Varones	Mujeres	porcentaje sobre el total	porcentaje sobre varones	porcentaje sobre mujeres
Contar 1	1	2607	2481	2470	0,321494636	0,305956345	0,304599827
Contar 2	2	1489	1390	1361	0,183623135	0,171414478	0,16783204
Contar 3	3	970	993	989	0,119620175	0,12245653	0,121963251
Contar 4	4	714	703	745	0,088050314	0,086693797	0,091873227
Contar 5	5	659	650	672	0,081267727	0,080157849	0,082870884
Contar 6	6	539	596	553	0,066469355	0,073498582	0,068195832
Contar 7	7	440	479	484	0,054260698	0,059070169	0,059686768
Contar 8	8	363	429	439	0,044765076	0,052904181	0,054137378
Contar 9	9	328	388	396	0,040448884	0,04784807	0,048834628
Total		**8109**	**8109**	**8109**	**1**	**1**	**1**

Nuevamente nos sorprendieron los resultados y nos acordamos de lo que el abuelo nos dijo:

La probabilidad de que el primer dígito de la longitud de los ríos del mundo sea un 1 es un 30,1%.

Y a nosotras nos salía un 32,14; 30,59 y 30,45%.

Deseábamos calcular el dato exacto de la Ley de Benford tal y como estaba en el libro del abuelo:

La probabilidad de que el primer dígito (n) de una serie no es 1/número de dígitos, sino que es $P(n)=Log_{10}(1+1/n)$; $n \in \{1,2,...,9\}$

Nos acordamos de nuestros abuelos cuando nos dijeron que ellos usaban el libro de "Tablas logarítmicas", pero nosotros lo tenemos mucho más fácil, vamos a Excel y usamos una de las funciones que tiene y de la que además podemos usar la ayuda.

LOG

Devuelve el logaritmo de un número en la base especificada.

Sintaxis

LOG(**número**;base)

Número es el número real positivo cuyo logaritmo desea obtener.

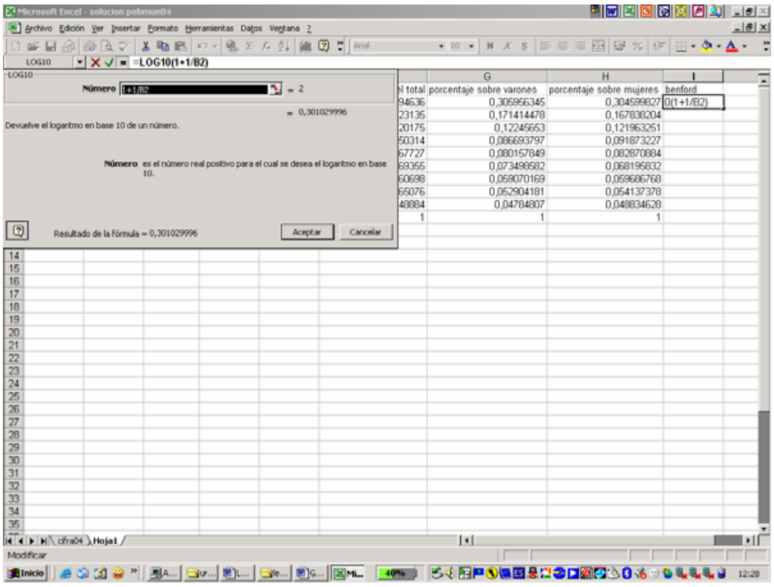
Base es la base del logaritmo. Si base se omite, el valor predeterminado es 10.

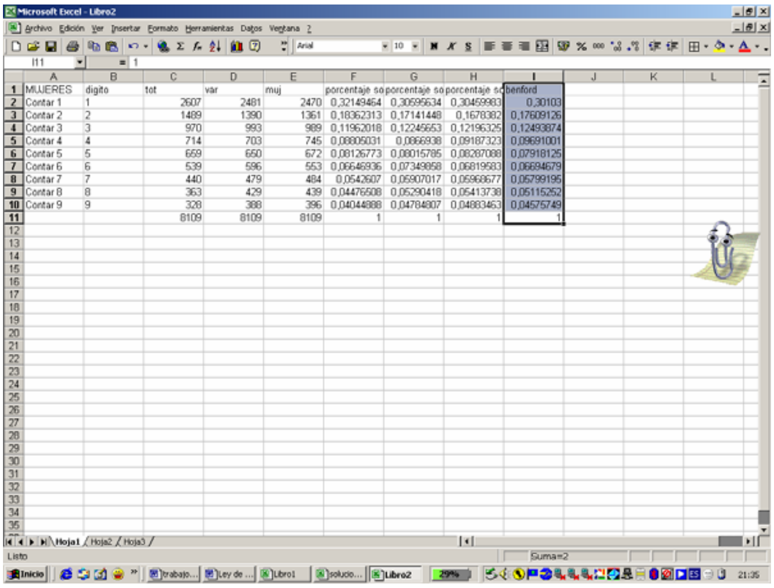
Ejemplos

LOG(10) es igual a 1

LOG(8; 2) es igual a 3

LOG(86; 2,7182818) es igual a 4,454347





Veamos los resultados obtenidos:

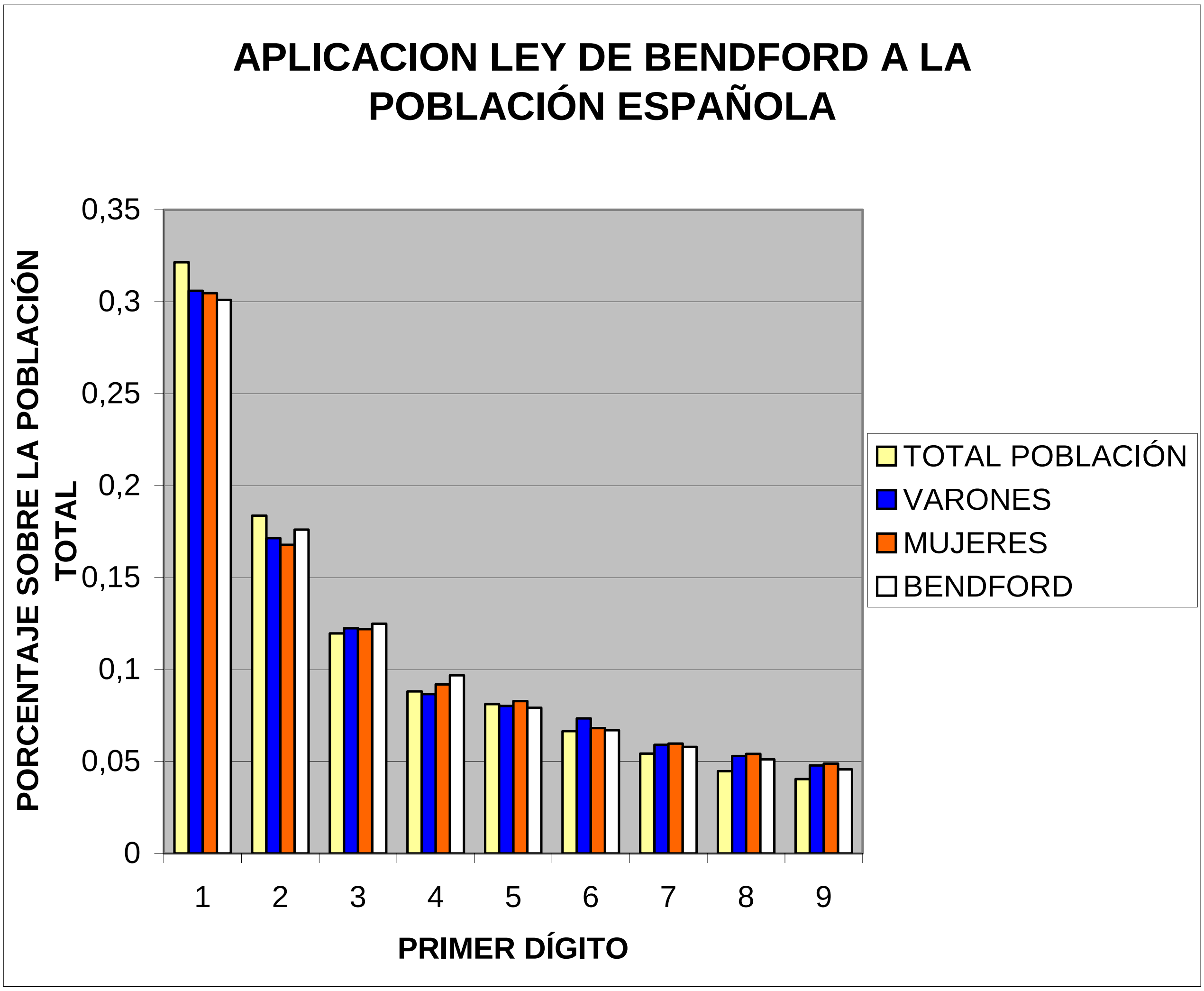
	Dígito	Total	Varones	Mujeres	porcentaje sobre el total	Porcentaje sobre varones	porcentaje sobre mujeres	Benford
Contar 1	1	2607	2481	2470	0,321494636	0,305956345	0,304599827	0,301029996
Contar 2	2	1489	1390	1361	0,183623135	0,171414478	0,167838204	0,176091259
Contar 3	3	970	993	989	0,119620175	0,12245653	0,121963251	0,124938737

Contar 4	4	714	703	745	0,088050314	0,086693797	0,091873227	0,096910013
Contar 5	5	659	650	672	0,081267727	0,080157849	0,082870884	0,079181246
Contar 6	6	539	596	553	0,066469355	0,073498582	0,068195832	0,06694679
Contar 7	7	440	479	484	0,054260698	0,059070169	0,059686768	0,057991947
Contar 8	8	363	329	339	0,044765076	0,052904181	0,054137378	0,051152522
Contar 9	9	328	388	396	0,040448884	0,04784807	0,048834628	0,045757491
Total		**8109**	**8109**	**8109**	**1**	**1**	**1**	**1**

Pues va a ser verdad que nuestro abuelo llevaba razón. Casi sale exacto, un 30% de los pueblos de España tienen como primer dígito un 1 en el caso de hombres y mujeres y un 2 puntos mas en el porcentaje total, lo que representa un 6% de diferencia , en este caso parece como si hubiera algunos municipios que en el total de habitantes varían los datos, parece que como 150 municipios han cambiado las cifras para tener un 1 como primer dígito en detrimento del 9 y sobre todo del 8 donde queda ligeramente por debajo de lo que indica la ley de Bendfor .

En general cumplen mejor la Ley de Bendfor los hombres y las mujeres por separado que en el total.

Para enviárselo a nuestro primo le hemos realizado el siguiente gráfico:


APLICACION LEY DE BENDFORD A LA POBLACIÓN ESPAÑOLA
PORCENTAJE SOBRE LA POBLACIÓN TOTAL
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
1
2
3
4
5
6
7
8
9
PRIMER DÍGITO
TOTAL POBLACIÓN
VARONES
MUJERES
BENDFORD

Preparando este trabajo hemos aprendido más cosas sobre la ley de Benford de las que nos indicó nuestro abuelo, por ejemplo que la misma ley se cumple para los dos primeros dígitos, y para el los tres primeros dígitos. Es decir la

$$P(n)=Log_{10}(1+1/n) \quad ; \quad n \in \{10,11,...,20,21,...,99\}$$

Así la probabilidad de que los dos primeros dígitos sean un 10 es

$$P(10)=Log_{10}(1+1/10)=0{,}04139269$$

La probabilidad de que sea un 20 es

$$P(20)=Log_{10}(1+1/20)= 0{,}0211892990$$

Y así sucesivamente. A esta ley se le denomina test de los dos primeros dígitos

Por la misma razón se cumple para tres dígitos:

$$P(n)=Log_{10}(1+1/n) \quad ; \quad n \in \{100,101,...,200,...210,...,999\}$$

A esta ley se le denomina test de los tres primeros dígitos.

El problema es que las muestras tienen que ser más numerosas ya que al existir más posibilidades, si la población es pequeña pueden existir desviaciones que invalidan los resultados

También hemos aprendido que como consecuencia de la ley de los dos primeros dígitos existe otra ley que se denomina del segundo dígito que dice que la probabilidad de que el segundo dígito sea un 0 es de un 12%, y de que sea un 9 es de sólo un 8,5% es decir casi la tercera parte menos.

Las probabilidades del test del segundo dígito son las siguientes

PROBABILIDAD DEL TEST DEL SEGUNDO DÍGITO	
Digito	Probabilidad
0	0,11967927
1	0,1138901
2	0,1088215
3	0,10432956
4	0,1003082
5	0,09667724
6	0,09337474
7	0,09035199

8	0,08757005
9	0,08499735

Este test como decimos es consecuencia del test de los dos primeros dígitos y así la probabilidad de que el segundo dígito sea un 0 es la suma de las probabilidades de que los dos primeros sean un 10, mas la de que sean un 20, mas la de que sean un 30, y así sucesivamente, en definitiva:

$$P(D\ 2=0)=Log_{10}(1+1/10)+ Log_{10}(1+1/20)+ Log_{10}(1+1/30)+\dots +Log_{10}(1+1/90)$$

Nos queda ver si la población española cumple esta ley.

Realizamos el mismo cálculo que anteriormente. Pero vemos que existen pueblos en los que o bien el número de varones, o el de mujeres, o ambos no llegan a 10 personas, por lo tanto en estos municipios no existe el segundo dígito y por lo tanto estos casos no hay que tenerlos en cuenta.

El resultado es el siguiente:

	total	varones	mujeres
Contar 0	996	1002	968
Contar 1	874	950	931
Contar 2	943	891	881
Contar 3	871	844	883
Contar 4	765	803	789
Contar 5	761	747	772
Contar 6	785	760	739
Contar 7	707	686	716
Contar 8	700	715	715
Contar 9	703	690	659
Cuenta general	8105	8088	8053

Calculamos los porcentajes:

	total	varones	mujeres	porcentaje	porcentaje	porcentaje

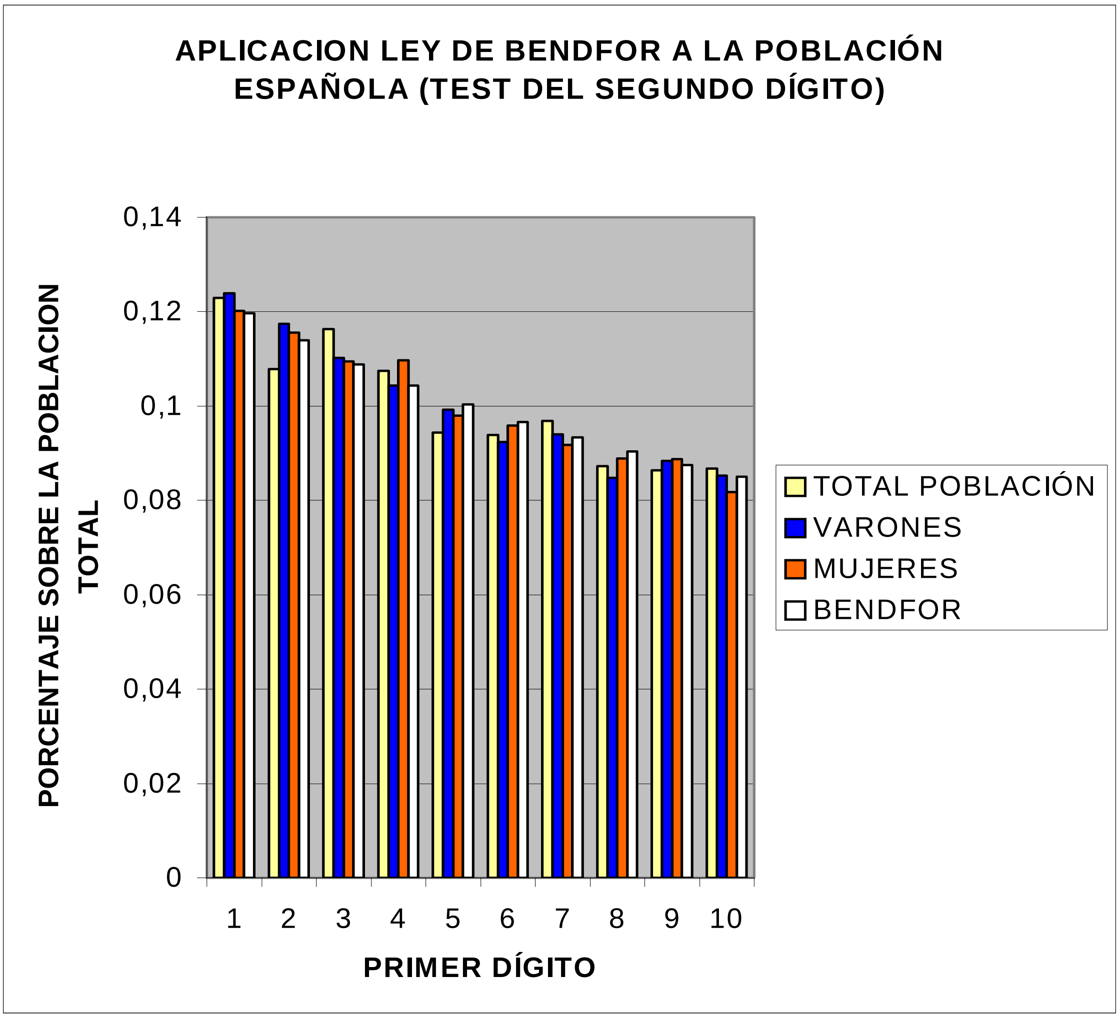
				sobre el total	sobre varones	sobre mujeres
Contar 0	996	1002	968	0,122887107	0,12388724	0,120203651
Contar 1	874	950	931	0,10783467	0,117457962	0,11560909
Contar 2	943	891	881	0,116347933	0,110163205	0,109400224
Contar 3	871	844	883	0,107464528	0,104352127	0,109648578
Contar 4	765	803	789	0,094386181	0,099282888	0,09797591
Contar 5	761	747	772	0,093892659	0,09235905	0,095864895
Contar 6	785	760	739	0,096853794	0,09396637	0,091767043
Contar 7	707	686	716	0,087230105	0,084817013	0,088910965
Contar 8	700	715	715	0,08636644	0,088402572	0,088786788
Contar 9	703	690	659	0,086736582	0,085311573	0,081832857
Cuenta general	8105	8088	8053	1	1	1

Si al cuadro anterior le añadimos la Ley de Benford

Dígito	Total	Varones	Mujeres	porcentaje sobre el total	Porcentaje sobre varones	porcentaje sobre mujeres	Benford
0	996	1002	968	0,122887107	0,12388724	0,120203651	0,11967927
1	874	950	931	0,10783467	0,117457962	0,11560909	0,1138901
2	943	891	881	0,116347933	0,110163205	0,109400224	0,1088215
3	871	844	883	0,107464528	0,104352127	0,109648578	0,10432956
4	765	803	789	0,094386181	0,099282888	0,09797591	0,1003082
5	761	747	772	0,093892659	0,09235905	0,095864895	0,09667724
6	785	760	739	0,096853794	0,09396637	0,091767043	0,09337474
7	707	686	716	0,087230105	0,084817013	0,088910965	0,09035199

8	700	715	715	0,08636644	0,088402572	0,088786788	0,08757005
9	703	690	659	0,086736582	0,085311573	0,081832857	0,08499735
total	8105	8088	8053	1	1	1	

Gráficamente los resultados son los siguientes:



Vemos que cumple La Ley de Benford aunque existen algunas variaciones que haría falta estudiar, como el total de la población en el dígito 2 y en el 3, pero en general los resultados son correctos.

A lo largo del estudio hemos visto alguna de las aplicaciones prácticas de la Ley de Bendfor como son

. Explicación de fraudes en los procesos electorales. Así hay varios estudios de las anomalías en las elecciones de México de 2006 realizado por la Comunidad Académica de la Universidad de México y la Universidad de Cornell, Estados Unidos de Norteamérica en las direcciones http://www.fisica.unam.mx/octavio/

http://macht.arts.cornell.edu/wrm1/pm06.pdf

La ley de Bendford se utiliza para investigación de fraude contable y fraude fiscal en la mayoría de las auditoras del mundo:

http://inza.wordpress.com/2006/07/20/ley-de-benford-newcomb/

http://www.nigrini.com.

También existen explicaciones matemáticas a la ley en diversas páginas como son:

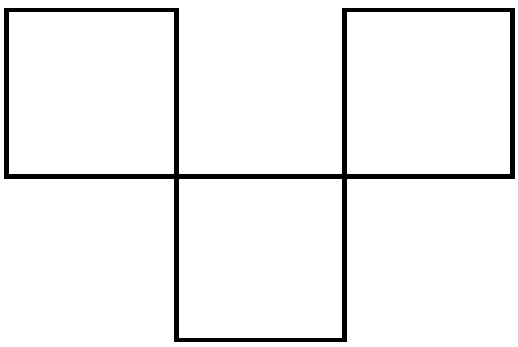
http://mathworld.wolfram.com/BenfordsLaw.html

http://www.mathpages.com/home/kmath302/kmath302.htm

http://plus.maths.org/issue9/features/benford/

http://www.nigrini.com/Benford's_law.htm

Por cierto, la solución al problema de geometría de nuestras primas pequeñas es:



BIBLIOGRAFÍA:

- JUEGOS DE INGENIO. Javier Vicente Cuadrillero

- MATEMÁTICAS Y JUEGOS DE AZAR. JUGAR CON LA PROBABILIDAD. John Haigh. Ed Tusquets editores. Barcelona 2003.

-DIGITAL ANALYSIS USING. BENDFORD´S LAW. Mark J. Nigrini, Ph.D. Ed. Global Audit Publications. Vancouver (Canadá) 2000

-EXCEL 5 GUIA DE INICIACIÓN. Susana Vazquez Jiménez. Ed Anaya multimedia. Madrid 1995

-Internet:

http://www.ine.es

http://es.wikipedia.org/wiki/Ley_de_Benford