

Tipo de artículo: Artículo original

## Modelo matemático para la validación de encuestas por medio de la Ley de Benford

### *Mathematical model for the validation of surveys using Benford's Law*

Santiago Xavier Valarezo Cofre<sup>1\*</sup> , <https://orcid.org/0000-0003-4518-6710>

<sup>1</sup>Instituto de Ciencias Exactas FERMAT, Ecuador. [sant\\_valarezo23@hotmail.com](mailto:sant_valarezo23@hotmail.com)

\* Autor para correspondencia: [sant\\_valarezo23@hotmail.com](mailto:sant_valarezo23@hotmail.com)

#### Resumen

La investigación busca solucionar la falencia en la toma de decisiones que provocan los datos erróneos o falsos que se pueden presentar en una investigación. De esta forma, se pretende mejorar los resultados y que las investigaciones tengan mayor validez al momento de que sean utilizadas en forma práctica. Muchas veces el investigador utiliza datos erróneos como consecuencia de la aplicación de una encuesta de manera inadecuada. Esta situación impacta en los resultados obtenidos y en las decisiones que se van a tomar. Es por esta razón que la Ley de Benfordes indispensable para validar la calidad en la recopilación de los datos, antes de comenzar con el proceso de análisis de los resultados. El objetivo es formular un modelo matemático para la validación de encuestas por medio de la Ley de Benford, que contribuya a mejorar las decisiones en el proceso de obtención de datos durante las investigaciones. El alcance es descriptivo, con enfoque cuantitativo. Se emplean dos encuestas como técnicas de recolección de datos. Como solución al objetivo general definido, se formuló un modelo matemático basado en la Ley de Benford y la fiabilidad. El modelo fue aplicado a los datos obtenidos en una encuesta realizada al sector empleador de la ciudad de Quito. Con la aplicación de la encuesta se conoció el ajuste y correspondencia de los datos a la Ley de Benford. De esta forma se determinó cuán alejado están los datos de la realidad o si en verdad representan a la misma.

**Palabras clave:** calidad del dato; confiabilidad; encuesta; Ley de Benford; modelo matemático.

#### Abstract

*The investigation seeks to solve the failure in decision-making that causes erroneous or false data that can be presented in an investigation. In this way, it is intended to improve the results and that the investigations have greater validity when they are used in a practical way. Many times the researcher uses erroneous data as a result of the inappropriate application of a survey. This situation has an impact on the results obtained and the decisions to be made. It is for this reason that Benford's Law is essential to validate the quality of data collection, before starting the process of analyzing the results. The objective is to formulate a mathematical model for the validation of surveys through the Law Benford, to help improve decisions in the data collection process during investigations. The scope is descriptive, with a quantitative approach. Two surveys are used as data collection techniques. As a solution to the general objective defined, a mathematical model based on Benford's Law and reliability was formulated. The model was applied to the data obtained in a survey carried out in the employer sector of the city of Quito. With the application of the survey, the adjustment and correspondence of the data to Benford's law was known. In this way, it was determined how far the data is from reality or if it really represents it.*

**Keywords:** data quality; reliability; poll; Benford's Law; mathematical model.

**Recibido:** 8/08/2021

**Aceptado:** 23/11/2021



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

## Introducción

El desarrollo científico y tecnológico va de la mano de las investigaciones básica y aplicada que generan soluciones a los problemas de la sociedad (Betancourt et al., 2017; Pérez & Téllez, 2015). El esfuerzo que se realice por apoyar la investigación debe ser continuo y ascendente. El trabajo del investigador es arduo, comienza con interrogantes o problemas y terminan con hallazgos y respuestas a esas interrogantes. Toda institución o persona que agregue un aporte a mejorar el proceso investigativo vale la pena que reciba el apoyo y se cristalicen sus esfuerzos. En este mismo orden de ideas, las organizaciones públicas y privadas requieren de información que esté validada para la toma de decisiones (Bermúdez & Fernández, 2021; Cristo et al., 2020; Valencia-Moreno et al., 2020). Es allí donde juega un rol importante el buen uso de los instrumentos para la recolección de datos que sea una representación de la realidad, a fin de que no se pierdan los esfuerzos que impliquen pérdidas de recursos humanos, físicos y financieros.

Como una contribución a lo antes expuesto, la presente investigación se basará en la aplicación de un modelo matemático, fundamentado en la Ley de Benford y la fiabilidad. Esta tiene el propósito de disminuir los resultados erróneos al analizar los datos de encuestas durante una investigación (Kaiser, 2019; Macías & Igua, 2018). De esta manera se mejora la toma de decisiones de los investigadores y asistentes de investigación, que los lleve a resultados y conclusiones efectivas. La implementación de la Ley de Benfordes de elevada ayuda como una técnica de análisis en ámbitos como el análisis geográfico, las finanzas, la auditoría, el análisis geográfico y para detectar fraudes de carácter electoral (Azevedo et al., 2021; García, 2021; Moreno-Montoya, 2020; Visbal et al., 2021). Es allí donde se quiere extender su uso para ayudar al investigador en sus decisiones en el desarrollo de las investigaciones científicas y las que requieren las organizaciones.

El empleo de la Ley de Benford para el análisis de fraudes en distintos sectores ha tenido mucho auge en los actuales momentos. Todos estos eventos han dado pie a diferentes investigaciones donde se estudia la aplicación de esta ley en casos prácticos, ayudando así a la toma de decisiones, como también a tomar previsiones y revisar las causas que la originaron (Castañeda, 2011; La casa, 2019; Matute & Zúñiga, 2010). Zamora (2015) llevó a cabo una investigación para evaluar la relevancia de la Ley de Benford para auditar las organizaciones por medio de difundir y aplicar la declaración de los impuestos a la renta en la Compañía COINFRA en el 2014. Encontró que los estados financieros para ese lapso presentaban diferencias entre las cuentas del impuesto sobre las rentas en comparación con el porcentaje de Benford del 21% de manera aproximada mensualmente en febrero y un 8% al año. Seguidamente concluyó sobre la importancia en las decisiones de la organización, donde la utilización de la Ley de Benfordes una herramienta muy



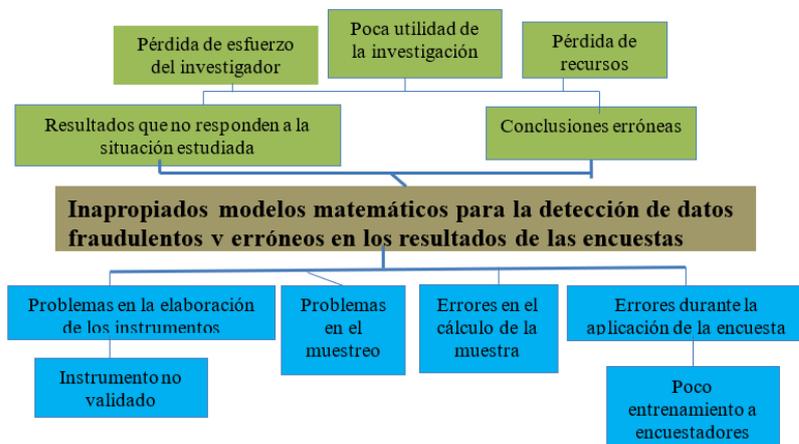
útil para las auditorías, por cuanto sirve como herramienta para la evaluación de la misma (Hashemi & Hariri, 2017; Nigrini, 2017; Singh & Peter, 2020).

En cambio, Matute & Zúñiga (2010) implementaron un sistema por medio de la utilización de la Ley de Benford que permite la detección de fraudes de carácter electoral en el Ecuador. Los autores afirman que esta ley puede ser utilizada como una herramienta para evidenciar manipulaciones o situaciones anómalas con los datos, para detectar problemas en los resultados finales. A tales resultados arribaron otros investigadores como Cerioli et al. (2019) y Lacasa & Fernández-Gracia (2019).

Por otro lado, Díaz (2004) indica que los errores en las medidas en una investigación social mediante encuesta, se deben a cuatro factores: errores producidos por los entrevistadores, por los entrevistados, debido al cuestionario y los producidos por el modo de recogida de los datos y la situación de la entrevista. A estos factores, Díaz (2004) le anexa, un quinto error, el producido durante la preparación y el tratamiento de datos. Indica también, que para minimizar los errores producidos por los entrevistadores se debe implementar supervisión de las entrevistas, que debe intensificarse al comienzo de cualquier investigación (Berger & Hill, 2021).

De todo lo anterior expuesto, cobra importancia este trabajo investigativo en el cual se formula un modelo matemático para validar encuestas por medio de la Ley de Benford y la fiabilidad. Su justificación teórica radica en que manejan fundamentos teóricos de la ley que pueden ser aplicados a casos particulares. Algunos de ellos son el análisis de datos proveniente de investigaciones científicas, de mercado, clientes y otros requerimientos que necesiten las organizaciones. Desde lo práctico, ayudará a la toma de decisiones de los investigadores respecto a información obtenida, incidiendo en mejores resultados que se acerquen más a la realidad estudiada y se evite la pérdida de recursos. Esta situación problemática detectada es mostrada en la figura 1, a partir de un árbol de problemas desarrollado.





**Figura 1.** Árbol de problemas que refleja el fenómeno objeto de estudio. Fuente: elaboración propia.

La presente investigación se circunscribe a la aplicación de la fiabilidad del instrumento utilizado para las encuestas, los aspectos teóricos de la Ley de Benford y los procesos estadísticos para su prueba de ajuste. El objetivo general de la investigación es formular un modelo matemático para la validación de encuestas por medio de la Ley de Benford a fin de no tomar malas decisiones por datos obtenidos durante las investigaciones.

## Materiales y métodos

Esta investigación se enfoca bajo el paradigma cuantitativo, ya que se utiliza en esta investigación un método de análisis causal y correlacional. El método que sigue la presente investigación es el método deductivo, ya que mediante este método se puede obtener una conclusión basado en el método científico, como parte del proceso de deducción. En el caso particular de este estudio, se partirá de la Ley de Benford y las fórmulas de fiabilidad de instrumentos, para obtener conclusiones sobre los datos obtenidos de las encuestas realizadas al sector empleador de la ciudad de Quito. Del mismo modo, la investigación tiene un alcance descriptivo.

### Técnica para recolección de datos

Se utilizó la encuesta para el estudio de mercado en el lanzamiento de un producto. Además, se aplicó otra encuesta, desarrollada por la Universidad Metropolitana de Ecuador (UMET), destinada a los empleadores del sector empresarial, con el propósito de determinar estrategias y actividades que mejorarían el quehacer de la carrera Ingeniería en Gestión Empresarial.

### Procesamiento de datos



obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

Se recogieron los datos en la aplicación de la encuesta al sector empresarial y laboral de la ciudad de Quito por la UMET. Se determinó el coeficiente de correlación interna y se le aplicó la Ley de Benford utilizando hoja de cálculo Excel. Seguidamente los resultados se representaron en un gráfico de barras donde se compararon los resultados con los establecidos por la ley. Seguidamente se le aplicó Chi-Cuadrado, el cual sirvió para calcular la bondad de ajuste de todos los dígitos respecto a lo esperado por la Ley de Benford.

Para la comparación de los resultados se utiliza una tabla de frecuencias, con base en la matriz en la que se ordenan con los resultados. De esta se extrae el primer dígito para posteriormente formar una tabla de frecuencias. Se calcula la frecuencia de cada dígito del uno al nueve. Con estos datos se procede al cálculo de la probabilidad de cada número, por medio de la división de la frecuencia de cada dígito para el total de datos. En la tabla de frecuencias se calcula la probabilidad de cada dígito por medio de la Ley de Benford para finalmente calcular el test de ajuste mediante la prueba de bondad chi cuadrada. Este procedimiento fue automatizado en la investigación mediante una macro de Excel.

### **Población y muestra**

La población se conformó por los empleados del sector empresarial y laboral de la ciudad de Quito, para un total de 90.034 sujetos. En cambio, la muestra se integró por dos grupos de sujetos: trabajadores que participan en el lanzamiento de un producto y trabajadores del sector empresarial y laboral de la ciudad de Quito. La muestra calculada fue de 383 sujetos como se precisa en la fórmula seguidamente. Para la implementación de la encuesta fue utilizado el muestreo por conglomerados. Se realizó las encuestas a los estudiantes de la UMET y por medio de los estudiantes se efectuó las encuestas a sus empleadores.

$$n = \frac{N\sigma^2 Z^2}{(N - 1)e^2 + \sigma^2 Z^2}$$

Donde:

n = es el tamaño esperado de la muestra

N = es el tamaño del universo o población

$\sigma$  = es la desviación estándar (0.5)

Z = responde al nivel de confianza del 95% (1.96)

e = error muestral (0.05)

### **Generalidades de la Ley de Benford**



obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional**  
(CC BY 4.0)

La Ley de Benford, enunciada en 1938, se conoce como la ley de los números anómalos afirma que, en una serie de números de la vida real, los dígitos iniciales de los mismos no tienen la misma probabilidad, es decir los números que empiezan por el dígito 1 tienen mayor frecuencia de aparición, seguidos de los que empiezan por el 2, así sucesivamente hasta llegar al 9, el cual es el que tiene menos probabilidad de aparecer. A medida que aumente el dígito, va a haber menos frecuencia que aparezca en el número (García, 2021).

Una de las principales aplicaciones de esta ley es la detección de fraudes. En 1992, se determinó que los estados financieros cumplen plenamente con la ley de Benford, que es importante para detectar el fraude financiero, incluidas las declaraciones de impuestos. Usarlo también es importante para detectar cambios en las cifras reportadas por empresas o personas, durante años consecutivos, y esto es muy útil porque indicará que algo anda mal, si no, esta detección temprana ahorrará tiempo y dinero.

En algunos países de América Latina se ha utilizado para detectar fraudes electorales. Por ejemplo, en algunos resultados electorales, la asignación de los primeros números, especialmente el segundo, no sigue la ley de Benford, lo que resulta en un recuento de votos para verificar el resultado. Por otro lado, los científicos belgas han utilizado esta ley para detectar anomalías en casos clínicos y verificar datos demográficos, entre otros usos.

## Resultados y discusión

Para la formulación y programación del modelo matemático se utilizó la herramienta Excel. Esta aplicación informática permite llevar a cabo actividades financieras y contables, mediante las funciones implementadas para permitir la creación y trabajo con hojas de cálculo.

Se procedió a crear una macro la cual permitirá que con simples pasos se puedan hacer los análisis de los datos aplicando la Ley de Benford, los mismos que darán como resultado las tablas de frecuencias, el test de ajuste chi-cuadrada y la gráfica de ajuste que comparara las frecuencias de los datos con la frecuencia de Benford. En las siguientes figuras se muestra de manera detallada el modo de utilización del programa en cuatro pasos (figuras 2, 3, 4 y 5):



	A	B	C	D	E	F	G	I
1								
2								
3			Análisis de Benfort					
4								
5								
6								
7		Datos						
8		7	2	1				
9		98	5	27				
10		2	13	45				
11		27	1	4				
12		3	4	2				
13		1	8	3				
14								
15								

Figura 2. Cargar el archivo en Excel (paso 1). Fuente: elaboración propia.

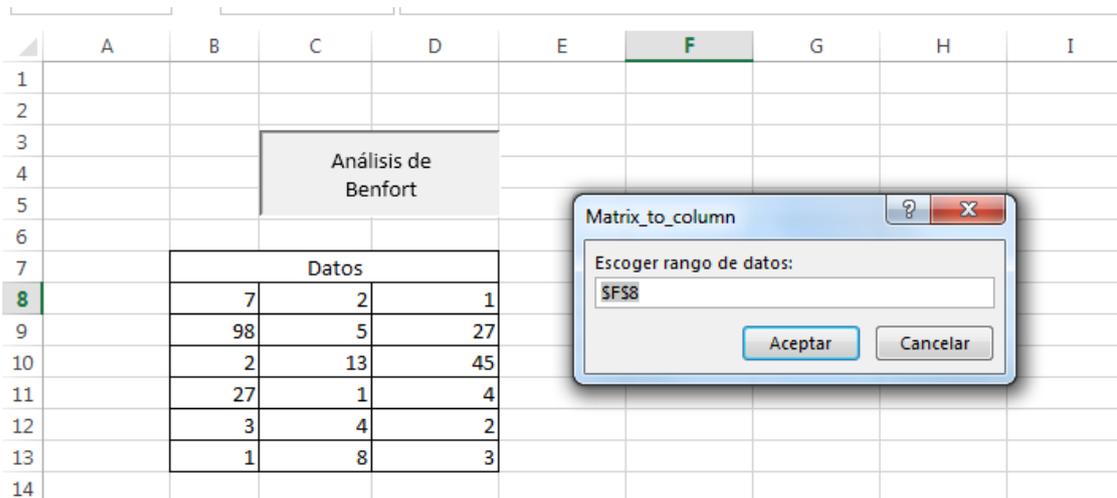


Figura 3. Dar un clic en la tecla análisis Benford (paso 2). Fuente: elaboración propia.



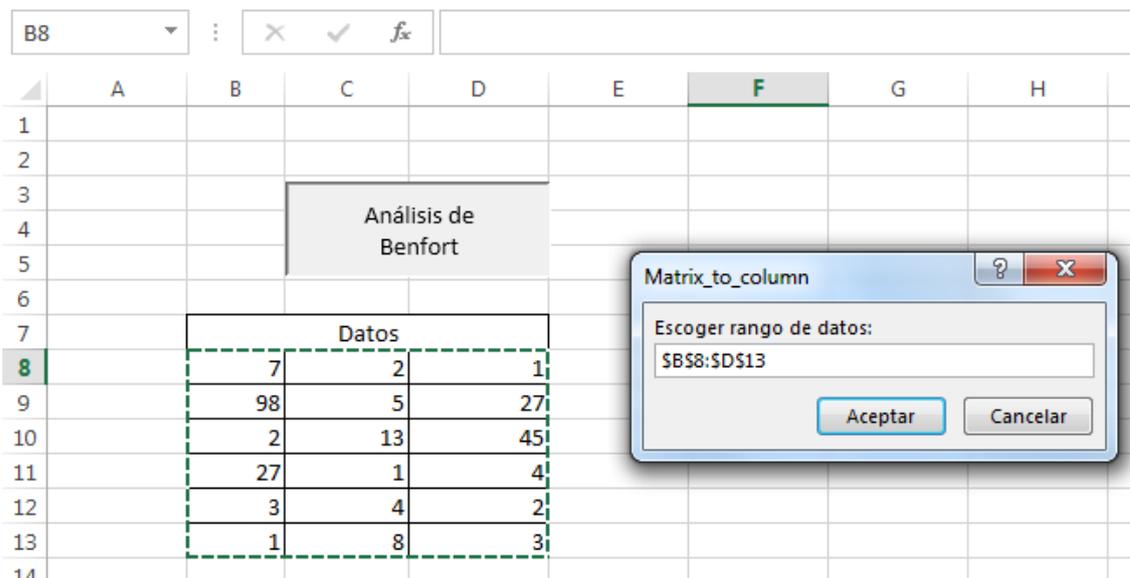


Figura 4. Escoger el rango de datos (paso 3). Fuente: elaboración propia.

Finalmente, en el paso 4 se debe dar clic en aceptar y se obtienen los resultados que son mostrados en la figura 5, específicamente: dígito, frecuencia, frecuencia relativa de cada dígito, la frecuencia según Benford y el estadístico chi-cuadrado. Este último determinará si los resultados se ajustan a la Ley de Benford. En la figura 5 aparece resaltado en color amarillo en el test de ajuste chi-cuadrada, el cual se debe comparar con el nivel de significación seleccionado en la tabla de distribución de chi-cuadrada.



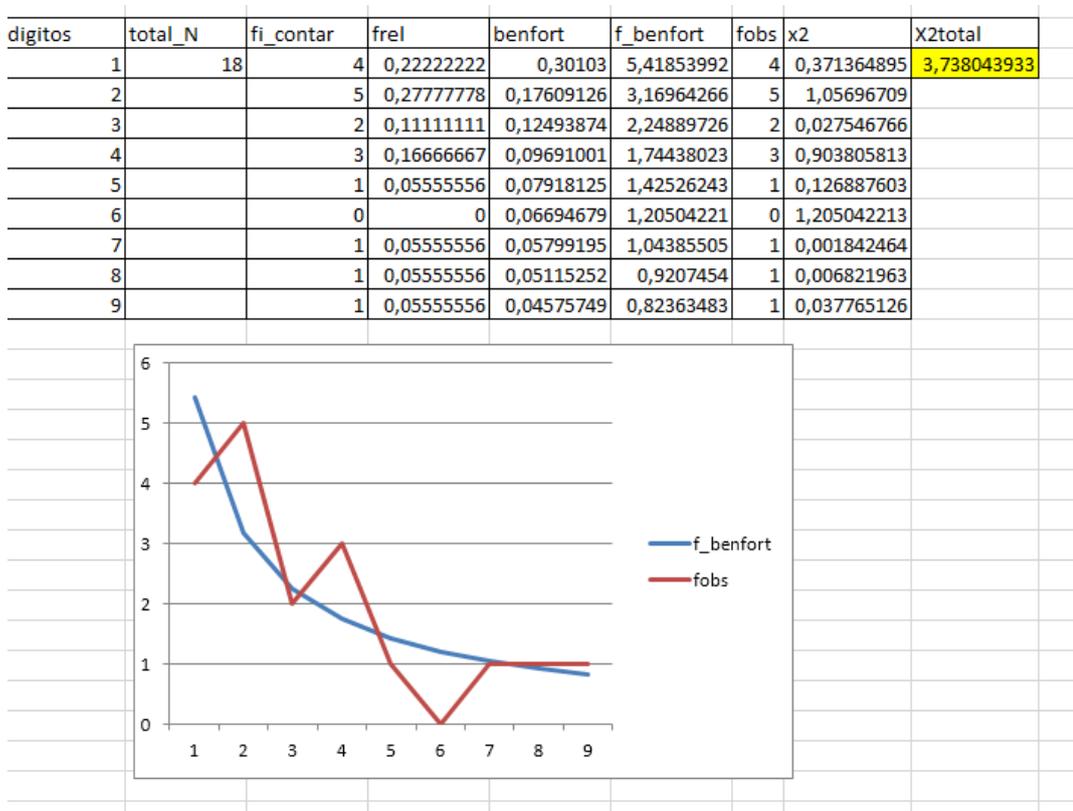


Figura 5. Dar clic en aceptar y se obtienen los resultados (paso 4). Fuente: elaboración propia.

### Aplicación del modelo matemático

Para la aplicación del modelo matemático y de fiabilidad se utilizaron los datos de los siguientes instrumentos para recolectarla información: una encuesta de investigación de mercado, la cual se aplicó para la prueba de producto, y una encuesta realizada a los empleadores. Ambas encuestas fueron entregadas a las estudiantes de la UMET que trabajan en relación de dependencia para que las aplicaran a sus jefes inmediatos.

Los primeros datos utilizados fueron los obtenidos en la encuesta de la investigación de mercados. Esta tuvo el objetivo de lanzar un producto a la venta. Con base en los datos, a continuación, se comprueba su validez por medio de la Ley de Benford, como es mostrado en la figura6. En este caso con 50, si solo se analizan los datos gráficamente se puede concluir que los datos siguen una Ley de Benford, pero si se le aplica el test de ajuste de datos chi cuadrado se comprueba que los datos no siguen una Ley de Benford ya que el valor del mismo es de  $X^2$  de 30,11 el cual es superior a 15,51 fijando con un nivel de significación del 5%.



obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

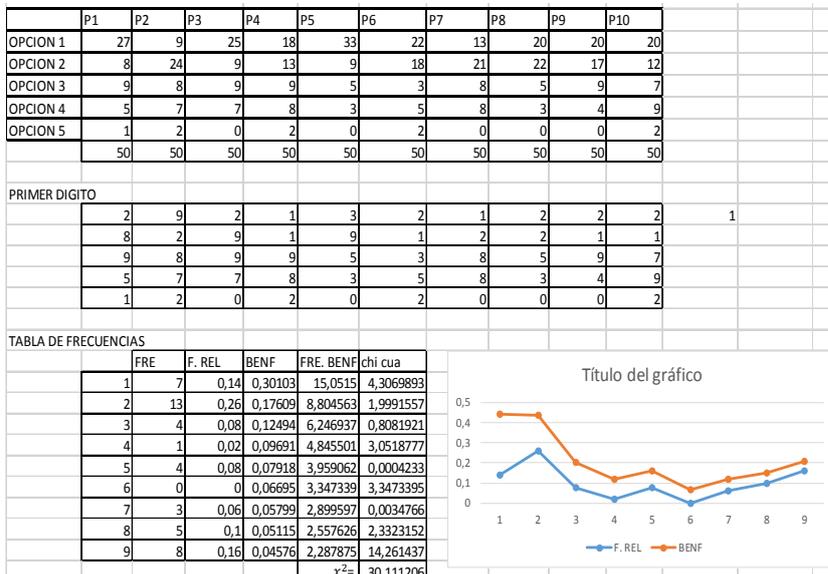


Figura 6. Análisis de los datos para una muestra de 50 individuos. Fuente: elaboración propia.

Posteriormente, como se muestra en la figura7, en este caso con 100 datos, si sólo se analiza gráficamente se puede concluir que los mismos siguen una Ley de Benford, pero al aplicarle el test de ajuste de datos denominada Chi-cuadrado ( $\chi^2$ ) se comprueba que los datos no siguen una Ley de Benford ya que el valor de  $\chi^2$  resultó de 16,05 el cual es superior a 15,51 con un nivel de significación del 5%.

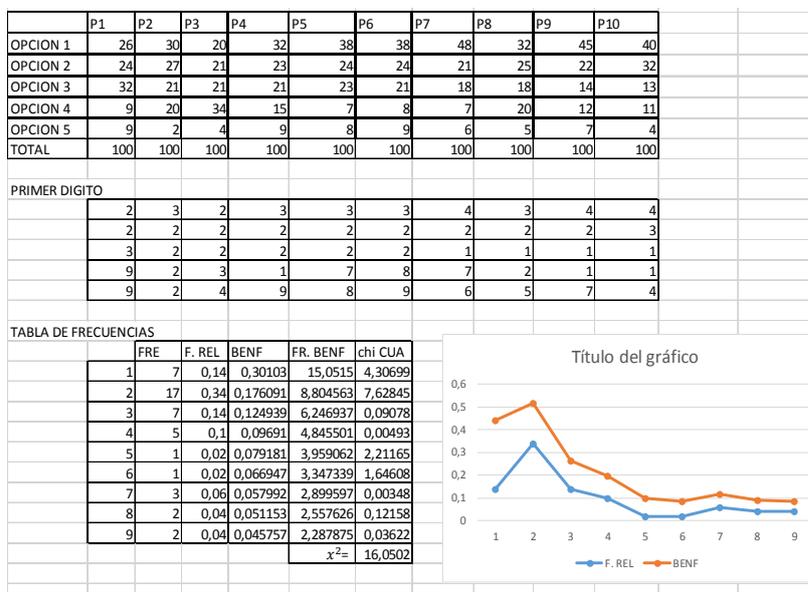
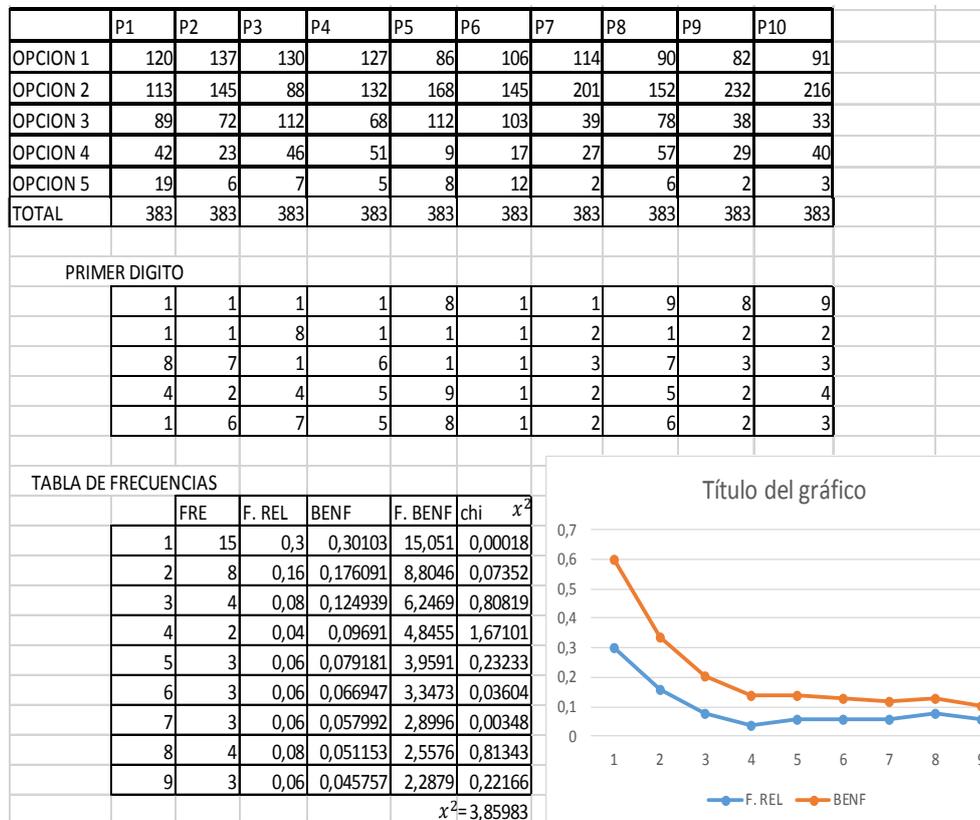


Figura 7. Análisis de los datos para una muestra de 100 individuos. Fuente: elaboración propia.



obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)



**Figura 8.** Análisis de los datos para una muestra de 383 individuos. Fuente: elaboración propia.

Por último, en la figura 8 mostrada anteriormente, se pueden observar las estadísticas, en este caso con 383 datos, si sólo se analiza gráficamente se puede observar que los datos siguen la Ley de Benford, sin embargo, por medio del test de ajuste de datos denominada chi cuadrado ( $X^2$ ) se comprueba que los datos siguen una Ley de Benford ya que el valor obtenido es de 3,86 el cual es inferior a 15,51 establecido con un nivel de significación del 5%.

### Aplicación del modelo matemático para la encuesta de la demanda laboral para los empleadores

Para el análisis de los datos de la encuesta, la UMET proporcionó los datos tabulados los cuales se procedió a analizarlos. El modelo matemático fue probado con estos resultados. Con los datos proporcionados por la UMET, se procedió con el análisis partiendo de la tabulación de la información, que son los resultados de las encuestas aplicadas a una muestra de 155 individuos, como se muestra en la figura 9.



obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

DATOS DE LA ENCUESTA DEMANDA LABORAL PARA EMPLEADORES													
131	2	70	0	13	16	121	30	0	0	5	26	20	10
24	8	31	7	1	46	77	1	1	0	128	21	7	114
0	9	0	4	27	38	89	51	0	0	22	129	0	30
0	1	2	9	114	38	55	22	1	21	0	1	1	
0	33	3	40	0	34	62	8	4	98	0	17	114	
0	16	53	95	1	1	74	25	34	36	7	9	35	
8	46	62	0	1	56	80	111	107	0	97	15	69	
1	33	35	0	11	66	53	38	9	0	51	108	27	
1	23	0	12	21	52	44	10	0	2	0	21	37	
3	53	1	14	121	99	86	47	1	9	0	22	35	
6	0	3	25	30	63	113	7	6	111	7	7	10	
7	0	67	107	19	91	35	20	28	33	114	8	27	
3	0	62	0	54	87	22	18	110	0	33	7	32	
13	30	22	11	71	25	73	39	10	0	0	19	65	
2	78	0	0	82	73	62	26	0	2	0	6	51	
2	46	1	8	70	47	29	37	0	14	5	20	39	
2	0	7	36	80	65	48	19	0	110	119	19	21	
30	1	28	100	101	43	113	41	16	28	31	39	0	
2	8	68	0	53	8	52	39	91	0	37	21	0	
7	54	51	0	42	0	92	35	48	0	9	11	1	

Figura 9. Tabulando los datos para una muestra de 155 individuos. Fuente: elaboración propia.

Con estos datos se procedió al análisis de los mismos aplicando la Ley de Benford, para lo cual se extrae el primer dígito de cada número, estos se presentan a continuación en la figura 10.

PRIEMR DÍGITO													
1	2	7	0	1	1	1	3	0	0	5	2	2	1
2	8	3	7	1	4	7	1	1	0	1	2	7	1
0	9	0	4	2	3	8	5	0	0	2	1	0	3
0	1	2	9	1	3	5	2	1	2	0	1	1	
0	3	3	4	0	3	6	8	4	9	0	1	1	
0	1	5	9	1	1	7	2	3	3	7	9	3	
8	4	6	0	1	5	8	1	1	0	9	1	6	
1	3	3	0	1	6	5	3	9	0	5	1	2	
1	2	0	1	2	5	4	1	0	2	0	2	3	
3	5	1	1	1	9	8	4	1	9	0	2	3	
6	0	3	2	3	6	1	7	6	1	7	7	1	
7	0	6	1	1	9	3	2	2	3	1	8	2	
3	0	6	0	5	8	2	1	1	0	3	7	3	
1	3	2	1	7	2	7	3	1	0	0	1	6	
2	7	0	0	8	7	6	2	0	2	0	6	5	
2	4	1	8	7	4	2	3	0	1	5	2	3	
2	0	7	3	8	6	4	1	0	1	1	1	2	
3	1	2	1	1	4	1	4	1	2	3	3	0	
2	8	6	0	5	8	5	3	9	0	3	2	0	
7	5	5	0	4	0	9	3	4	0	9	1	1	

Figura 10. Primer dígito para una muestra de 155 individuos. Fuente: elaboración propia.



obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

Con los datos del primer dígito antes mostrado, se procedió a realizar la tabla de frecuencia, como se muestra en la figura 11, la cual está conformada por:

- Los  $X_i$ , que son los números del 1 al 9.
- Frecuencia absoluta ( $f$ ) que nos indica con qué frecuencia aparecen los números del 1 al 9.
- Frecuencia relativa ( $fr$ ) que resulta de la división de cada frecuencia absoluta para el total, cuya fórmula es  $fr = \frac{f_i}{n}$
- La Ley de Benford (LB) que en la probabilidad con que aparecen cada número del uno al nueve que está dada por la formula,  $p(x_i) = \log(x_i + 1) - \log(x_i)$
- La frecuencia de Benford que se determina multiplicando cada valor de la Ley de Benford por el número de datos de la muestra que es  $n=219$ .
- Se calcula la sumatoria del cuadrado de la diferencia entre la frecuencia y la frecuencia de Benford a esto se divide para la frecuencia de Benford con la finalidad de obtener el test de ajuste de datos denomina chi-cuadrada.

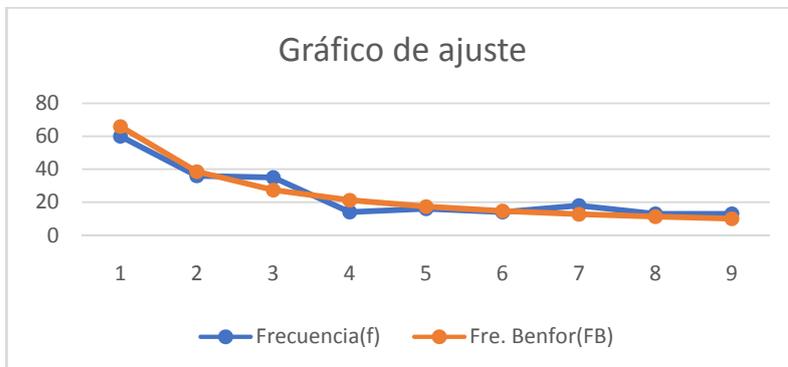
TABLA DE FRECUENCIAS					
Xi	Frecuencia(f)	Fre. Relativa(fr)	ley Benford(LB)	Fre. Benfor(FB)	$\frac{(f - FB)^2}{FB}$
1	60	0,273972603	0,301029996	65,92556905	0,53260623
2	36	0,164383562	0,176091259	38,56398573	0,17047052
3	35	0,159817352	0,124938737	27,36158332	2,13238425
4	14	0,063926941	0,096910013	21,22329285	2,45842905
5	16	0,073059361	0,079181246	17,34069288	0,10365545
6	14	0,063926941	0,06694679	14,66134693	0,02983217
7	18	0,082191781	0,057991947	12,70023639	2,21157256
8	13	0,059360731	0,051152522	11,20240242	0,28845215
9	13	0,059360731	0,045757491	10,02089043	0,8856592
Total	219	1	1	219	8,81306159
				$X^2 =$	8,81306159

**Figura 11.** Tabla frecuencias y prueba de Chi-Cuadrado. Fuente: elaboración propia.

En este caso se obtiene que los datos siguen la Ley de Benford, ya que por medio del test de ajuste de datos denominada chi cuadrado ( $X^2$ ) se comprueba la afirmación anterior por cuanto el valor obtenido de  $X^2$  es de 8,81, el cual es inferior a 15,51 con un nivel de significación del 5%. A continuación, se muestra la figura 12, donde se realiza una comparación de frecuencias para determinar si se ajustan o no a la Ley de Benford.



obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)



**Figura 12.** Comparación de frecuencias absolutas y Ley de Benford. Fuente: elaboración propia.

En este gráfico se puede observar que las frecuencias absolutas se ajustan a la Ley de Benford, tomando en cuenta que las frecuencias absolutas corresponden al color azul y las frecuencias dadas por la Ley de Benford al color rojo.

### Datos alterados

Para la comprobación de la Ley de Benford, y con la finalidad de detectar datos erróneos y fraudulentos, se realizaron las mismas encuestas a tres individuos, los cuales llenaron los cuestionarios al azar y esto lo hicieron con las 155 encuestas. Este procedimiento arrojó los siguientes resultados:

DATOS DE LA ENCUESTA DEMANDA LABORAL PARA EMPLEADORES													
39	16	0	12	71	30	37	11	12	49	0	15	8	0
92	11	0	76	0	69	9	20	3	87	0	61	9	0
14	13	13	56	0	41	3	34	6	0	3	27	6	15
0	9	66	0	0	56	45	14	3	0	79	38	0	32
79	7	66	0	8	42	25	5	0	0	62	55	0	98
59	9	0	0	68	41	94	6	0	5	0	0	0	
3	3	0	15	69	19	25	0	0	63	1	6	37	
9	1	0	76	0	7	11	13	6	77	3	21	22	
29	0	12	54	0	5	9	8	56	0	76	12	50	
36	43	71	0	0	79	16	31	83	0	65	29	28	
9	63	62	0	10	63	18	52	0	0	0	58	8	
10	39	0	0	67	88	4	62	0	9	1	19	5	
5	0	0	17	67	62	40	22	0	67	5	11	2	
7	0	0	67	50	68	36	11	8	69	61	9	20	
5	0	10	61	61	51	18	5	58	0	78	6	26	
5	0	74	0	67	25	35	4	79	0	0	8	32	
5	13	60	0	64	63	34	1	0	1	15	16	11	
3	58	0	0	70	67	17	9	1	16	11	14	6	
7	74	0	7	63	61	20	8	0	58	40	17	0	
11	0	0	67	39	49	5	9	8	70	79	12	1	

**Figura 13.** Datos de encuesta alterados. Fuente: elaboración propia.

Con estos datos se procede a su análisis. Para ello, se aplicó el modelo matemático propuesto basado en la Ley de Benford, para lo cual se extrae el primer dígito de cada número, lo cual se presenta en la figura 13, así como la distribución de frecuencias y Chi- Cuadrado para datos alterados, que se muestra en la figura 14.



PRIMER DÍGITO													
3	1	0	1	7	3	3	1	1	4	0	1	8	0
9	1	0	7	0	6	9	2	3	8	0	6	9	0
1	1	1	5	0	4	3	3	6	0	3	2	6	1
0	9	6	0	0	5	4	1	3	0	7	3	0	3
7	7	6	0	8	4	2	5	0	0	6	5	0	9
5	9	0	0	6	4	9	6	0	5	0	0	0	0
3	3	0	1	6	1	2	0	0	6	1	6	3	0
9	1	0	7	0	7	1	1	6	7	3	2	2	0
2	0	1	5	0	5	9	8	5	0	7	1	5	0
3	4	7	0	0	7	1	3	8	0	6	2	2	0
9	6	6	0	1	6	1	5	0	0	0	5	8	0
1	3	0	0	6	8	4	6	0	9	1	1	5	0
5	0	0	1	6	6	4	2	0	6	5	1	2	0
7	0	0	6	5	6	3	1	8	6	6	9	2	0
5	0	1	6	6	5	1	5	5	0	7	6	2	0
5	0	7	0	6	2	3	4	7	0	0	8	3	0
5	1	6	0	6	6	3	1	0	1	1	1	1	0
3	5	0	0	7	6	1	9	1	1	1	1	6	0
7	7	0	7	6	6	2	8	0	5	4	1	0	0
1	0	0	6	3	4	5	9	8	7	7	1	1	0

Figura 14. Primer dígito de encuesta con datos alterados. Fuente: elaboración propia.

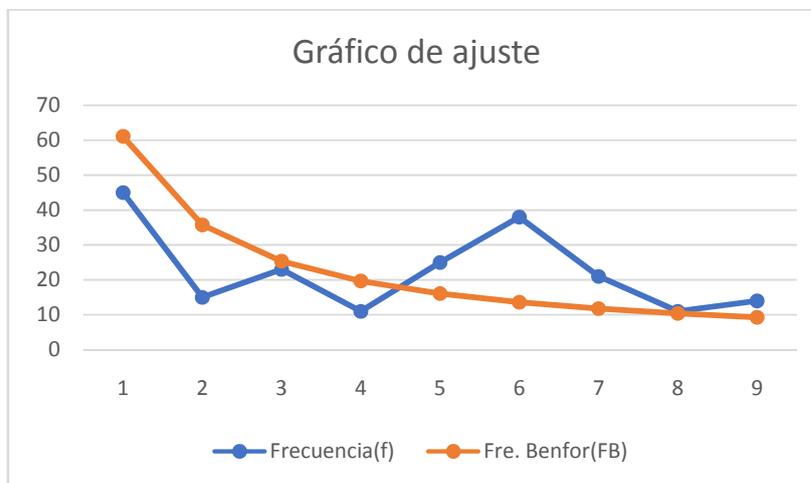
TABLA DE FRECUENCIAS					
Xi	Frecuencia(f)	Fre. Relativa(fr)	ley Benford(LB)	Fre. Benfor(FB)	$\frac{(f - FB)^2}{FB}$
1	45	0,221674877	0,301029996	61,10908912	4,24654918
2	15	0,073891626	0,176091259	35,74652559	12,0408436
3	23	0,113300493	0,124938737	25,36256353	0,220076588
4	11	0,054187192	0,096910013	19,67273264	3,823377912
5	25	0,123152709	0,079181246	16,07379295	4,956961471
6	38	0,187192118	0,06694679	13,5901983	43,8432469
7	21	0,103448276	0,057991947	11,77236524	7,232976689
8	11	0,054187192	0,051152522	10,38396206	0,036547008
9	14	0,068965517	0,045757491	9,288770584	2,389517796
Total	203	1	1	203	78,79009715
				$\chi^2 =$	78,79009715

Figura 15. Distribución de frecuencias y Chi- Cuadrado para datos alterados. Fuente: elaboración propia.

En este caso, se obtiene que los datos no siguen la Ley de Benford, ya por medio del test de ajuste de datos denominada chi cuadrado ( $\chi^2$ ) arroja un valor de 78.79, siendo este valor superior a 15,51 con un nivel de significación del 5%. En la figura 16 se muestra las frecuencias absolutas no se ajustan a la Ley de Benford, tomando en cuenta que las frecuencias absolutas están representadas en color azul y las frecuencias dadas por la Ley de Benford en color rojo.



obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)



**Figura 16.** Comparación de frecuencias absolutas con la Ley de Benford. Fuente: elaboración propia.

## Conclusiones

Luego de realizada la investigación se puede afirmar que el modelo matemático formulado es aplicable para la validación de encuestas. El modelo matemático fue aplicado a dos instrumentos diferentes y con un número de muestras variadas. Los resultados arrojados en la aplicación del test de ajuste chi-cuadrada indicaron que mientras la aplicación de las encuestas se acercó a la muestra, los datos se ajustaron a lo referido en la Ley de Benford.

El modelo matemático basado en la Ley de Benford es aplicable para validación de encuestas. Ello se debe a que cuando se aplica el método a encuestas reales, los datos se ajustan a la Ley de Benford y cuando se aplican a datos alterados o erróneos no se ajustan a la Ley de Benford. En tal caso, la importancia de la validación de los datos en las encuestas permitió constatar la precisión del modelo matemático formulado, así como la obtención de resultados ajustados y satisfactorios.

Como resultado de la investigación, igualmente se puede afirmar que el hecho de que los resultados se ajusten a lo referido en la Ley de Benford indica que las encuestas fueron aplicadas con un procedimiento adecuado y a la cantidad correcta de la muestra definida. De igual manera, el hecho de que una encuesta tenga una consistencia interna adecuada no significa necesariamente que esta se realizó a toda la muestra.

Para trabajos futuros se recomienda aplicar este modelo matemático en las encuestas que realizan los estudiantes en todos los niveles, tanto para un producto integrador como para sus trabajos de titulación. Del mismo modo se recomienda aplicar este modelo matemático para la investigación de mercados para darle veracidad a los resultados.



## Conflictos de intereses

Los autores de la investigación no presentan conflictos de intereses.

## Contribución de los autores

1. Conceptualización: Santiago Xavier Valarezo Cofre
2. Curación de datos: Santiago Xavier Valarezo Cofre
3. Análisis formal: Santiago Xavier Valarezo Cofre
4. Adquisición de fondos: Santiago Xavier Valarezo Cofre
5. Investigación: Santiago Xavier Valarezo Cofre
6. Metodología: Santiago Xavier Valarezo Cofre
7. Administración del proyecto: Santiago Xavier Valarezo Cofre
8. Recursos: Santiago Xavier Valarezo Cofre
9. Software: Santiago Xavier Valarezo Cofre
10. Supervisión: Santiago Xavier Valarezo Cofre
11. Validación: Santiago Xavier Valarezo Cofre
12. Visualización: Santiago Xavier Valarezo Cofre
13. Redacción – borrador original: Santiago Xavier Valarezo Cofre
14. Redacción – revisión y edición: Santiago Xavier Valarezo Cofre

## Financiamiento

La investigación no requirió fuente de financiamiento.

## Referencias

- AZEVEDO, C. D. S., GONÇALVES, R. F., GAVA, V. L., DE MESQUITA SPINOLA, M. A Benford's Law based methodology for fraud detection in social welfare programs: Bolsa Familia analysis. *Physica A: Statistical Mechanics and its Applications*. 2021, 567, 125626. <https://doi.org/10.1016/j.mex.2021.101575>
- BERGER, A., HILL, T. P. The mathematics of Benford's law: a primer. *Statistical Methods Applications*. 2021, 30(3), 779-795. <https://doi.org/10.1007/s10260-020-00532-8>



obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional**  
(CC BY 4.0)

- BERMÚDEZ, M. M. D. C., FERNÁNDEZ, M. D. Gestión del gobierno orientado a la innovación: Contexto y caracterización del Modelo. *Universidad y Sociedad*. 2021, 13(1), 6-16. <https://rus.ucf.edu.cu/index.php/rus/article/view/1892>
- BETANCOURT, O. B., TORRES, D. C., PÉREZ, N. D. L. C. S. Evaluación del estado de la gestión de la información científica y tecnológica: dimensiones e indicadores. *Revista Cubana de Información en Ciencias de la Salud*. 2017, 28(4), 1-19. <https://www.redalyc.org/pdf/3776/377654835002.pdf>
- CASTAÑEDA, G. La ley de Benford y su aplicabilidad en el análisis forense de resultados electorales. *Política y gobierno*. 2011, 18(2), 297-329. [http://www.scielo.org.mx/scielo.php?script=sci\\_arttextpid=S1665-20372011000200004](http://www.scielo.org.mx/scielo.php?script=sci_arttextpid=S1665-20372011000200004)
- CERIOLO, A., BARABESI, L., CERASA, A., MENEGATTI, M., PERROTTA, D. Newcomb–Benford law and the detection of frauds in international trade. *Proceedings of the National Academy of Sciences*. 2019 116(1), 106-115. <https://doi.org/10.1073/pnas.1806617115>
- CRISTO, N. M., PÉREZ, J. F. R., IZAGUIRRE, L. V. Estrategia de entrenamiento y acompañamiento a usuarios para el Sistema de Información Hospitalaria XAVIA HIS. *Revista Cubana de Informática Médica*. 2020, 12(1), 76-91. <https://pesquisa.bvsalud.org/portal/resource/fr/biblio-1126556>
- DÍAZ, R. Problemas de representatividad en las encuestas con muestreos probabilísticos. *Papers: Revista de Sociología*. 2004, 45-66. <https://www.raco.cat/index.php/Papers/article/download/25793/25627>
- GARCÍA, P. M. C. Aplicabilidad de la ley de Benford a la detección de fraudes. *Universidad y Sociedad*. 2021, 3(4), 461-467. <https://rus.ucf.edu.cu/index.php/rus/article/view/2186>
- HASHEMI, S. A., HARIRI, A. The analysis of benford's law ability to identify and predict financial fraud detection. *Accounting and Auditing Review*. 2017, 24(2), 283-302. <https://doi.org/10.22059/acctgrev.2017.220346.1007490>
- KAISER, M. Benford'S Law As An Indicator Of Survey Reliability—Can We Trust Our Data?. *Journal of Economic Surveys*. 2019, 33(5), 1602-1618. <https://doi.org/10.1111/joes.12338>
- LACASA, L. Newcomb–Benford law helps customs officers to detect fraud in international trade. *Proceedings of the National Academy of Sciences*. 2019, 116(1), 11-13. <https://doi.org/10.1073/pnas.1819470116>
- LACASA, L., FERNÁNDEZ-GRACIA, J. Election forensics: Quantitative methods for electoral fraud detection. *Forensic science international*. 2019, 294, e19-e22. <https://doi.org/10.1016/j.forsciint.2018.11.010>



- MACÍAS, A. L. O., IGUA, S. T. Encontrando datos anómalos en la tributación.: Aplicación de la Ley de Benford en el Impuesto a la Renta en Ecuador. *SaberEs*. 2018, 10(2), 173-188. <https://dialnet.unirioja.es/servlet/articulo?codigo=6759823>
- MATUTE, L., ZÚÑIGA, M. Sistema utilizando la Ley de Benford para detectar posibles fraudes electorales en las elecciones convocadas en Ecuador. Quito: Escuela Politécnica Nacional. 2010. <http://bibdigital.epn.edu.ec/handle/15000/2547>
- MORENO-MONTOYA, J. Ley de Benford con muestras pequeñas: una prueba exacta nueva útil en ciencias de la salud durante epidemias. *Salud UIS*. 2020, 52(2), 153-155. <https://revistas.uis.edu.co/index.php/revistasaluduis/article/view/10656>
- NIGRINI, M. J. Audit sampling using Benford's Law: A review of the literature with some new perspectives. *Journal of emerging technologies in accounting*. 2017, 14(2), 29-46. <https://doi.org/10.2308/jeta-51783>
- PÉREZ, J. F. R., TÉLLEZ, R. B. Propuesta de red cubana Aurora para la colaboración médica a través de Infomed utilizando un enfoque de redes sociales. La Habana: Memorias Convención Internacional de Salud Cuba Salud. 2015. <http://actasdecongreso.sld.cu/index.php?P=FullRecord&ID=719>
- SINGH, K., PETER, B. Implementing Benford's law in continuous monitoring applications. *Accounting and Management Information Systems*. 2020, 19(2), 379-404. <https://www.ceeol.com/search/article-detail?id=879225>
- VALENCIA-MORENO, J. M., LÓPEZ, E. G., PÉREZ, J. F. R., RODRÍGUEZ, J. P. F., XOCHIHUA, O. Á. Exploring Breast Cancer Prediction for Cuban Women. In *International Conference on Information Technology Systems* (pp. 480-489). Springer, Cham. 2020. [https://link.springer.com/chapter/10.1007/978-3-030-40690-5\\_47](https://link.springer.com/chapter/10.1007/978-3-030-40690-5_47)
- VISBAL, J. H. W., PEDRAZA, M. C. C., ANAYA, O. P. Evaluación preliminar de un sistema de vigilancia en salud pública usando la ley de Benford. *Revista Cubana de Medicina General Integral*. 2021, 37. [www.revmgisld.cu/index.php/mgi/article/download/1816/451](http://www.revmgisld.cu/index.php/mgi/article/download/1816/451)
- ZAMORA, M. Estudio y análisis de la Ley de Benford aplicada en los estados financieros de COINFRA SA (Disertación doctoral). Guayaquil: Universidad de Guayaquil. 2015. <http://repositorio.ug.edu.ec/bitstream/redug/11169/1/TESIS%20LEY%20DE%20BENFORD%20MARIA%20ZAMORA%20L.pdf>

