

# 04

## **APLICACIÓN**

**DE LA LEY DE BENFORD PARA LA DETECCIÓN DE FRAUDES  
EN ENCUESTAS DE SATISFACCIÓN UNIVERSITARIA**

# APLICACIÓN

## DE LA LEY DE BENFORD PARA LA DETECCIÓN DE FRAUDES EN ENCUESTAS DE SATISFACCIÓN UNIVERSITARIA

### APPLICATION OF BENFORD'S LAW FOR THE DETECTION OF FRAUD IN UNIVERSITY SATISFACTION SURVEYS

Pedro Manuel Cabeza-García<sup>1</sup>

E-mail: [pedroca07@yahoo.es](mailto:pedroca07@yahoo.es)

ORCID: <https://orcid.org/0000-0002-0748-906X>

<sup>1</sup> Universidad San Francisco de Quito. Ecuador.

Cita sugerida (APA, séptima edición)

Cabeza-García, P. M. (2025). Aplicación de la Ley de Benford para la detección de fraudes en encuestas de satisfacción universitaria. *Revista Metropolitana de Ciencias Aplicadas*, 8(1), 35-41.

#### RESUMEN

Este estudio aplica la Ley de Benford para analizar la posible existencia de fraude en las respuestas de una encuesta de satisfacción realizada a 54 estudiantes de la Universidad Iberoamericana del Ecuador sede Quito, del área de Salud. Se aplicó un cuestionario de 12 preguntas con respuestas en una escala Likert del 1 al 5. Los resultados de la prueba de Chi-cuadrado no muestran evidencia significativa de fraude, sugiriendo que las respuestas se alinean con la distribución esperada de la Ley de Benford. Quiere decir que los estudiantes respondieron el cuestionario y no fueron manipuladas las respuestas por otra persona, lo cual da la veracidad de las respuestas.

#### Palabras clave:

Encuestas de satisfacción universitaria, Ley de Benford, fraude.

#### ABSTRACT

This study applies Benford's Law to analyze the possible existence of fraud in the responses to a satisfaction survey conducted with 54 students from the Universidad Iberoamericana del Ecuador, Quito, in the Health area. A questionnaire of 12 questions was applied with answers on a Likert scale from 1 to 5. The results of the Chi-square test do not show significant evidence of fraud, suggesting that the answers align with the expected distribution of Benford's Law. It means that the students answered the questionnaire and the answers were not manipulated by another person, which gives the truthfulness of the answers.

#### Keywords:

University satisfaction surveys, Benford's Law, fraud.

## INTRODUCCIÓN

En la actualidad, la calidad y veracidad de los datos recolectados a través de encuestas y cuestionarios es de vital importancia para la toma de decisiones informadas en diversos campos, incluyendo la educación, la administración pública y la investigación científica. Sin embargo, la posibilidad de fraude o manipulación de los datos es una preocupación constante. La detección de estas irregularidades es crucial para garantizar la integridad de los resultados y la confianza en los procesos de toma de decisiones basados en dichos datos.

Una herramienta estadística útil para detectar anomalías en conjuntos de datos es la Ley de Benford. Esta ley, también conocida como la Ley del Primer Dígito, establece que, en muchos conjuntos de datos numéricos, el dígito 1 aparece como primer dígito aproximadamente el 30% de las veces, mientras que los dígitos más altos aparecen con menor frecuencia. La distribución de los primeros dígitos se describe mediante una función logarítmica y es sorprendentemente consistente en una amplia variedad de contextos, desde datos financieros hasta resultados electorales.

La Ley de Benford ha demostrado ser particularmente eficaz en la detección de fraudes y manipulación de datos. Su aplicación en auditorías contables, por ejemplo, ha permitido identificar discrepancias y anomalías que pueden indicar actividades fraudulentas. Del mismo modo, se ha utilizado para analizar datos electorales y detectar posibles manipulaciones en los resultados de votaciones. En el ámbito académico, la ley ha sido aplicada para evaluar la veracidad de datos reportados en investigaciones científicas y encuestas.

El presente estudio se centra en la aplicación de la Ley de Benford para evaluar la veracidad de las respuestas obtenidas de una encuesta de satisfacción aplicada a estudiantes universitarios del área de salud de la universidad Iberoamericana del Ecuador sede Quito La Query. La encuesta, realizada a 54 estudiantes, incluyó 12 preguntas con respuestas en una escala Likert del 1 al 5. El objetivo principal de este estudio es determinar si las frecuencias observadas de los primeros dígitos en las respuestas de la encuesta se desvían significativamente de la distribución esperada según la Ley de Benford, lo que podría indicar la presencia de fraude o manipulación.

La importancia de este estudio radica en su aplicación práctica para evaluar la integridad de los datos en un contexto educativo. Las encuestas de satisfacción son herramientas comunes utilizadas por las instituciones educativas para medir la percepción de los estudiantes sobre diversos aspectos de su experiencia académica. La posibilidad de que los datos de estas encuestas sean manipulados puede tener implicaciones significativas para la toma de decisiones y la implementación de mejoras en los servicios ofrecidos (Cruz Collin et al., 2024)

en las prácticas institucionales coadyuvando a proponer acciones que contribuyan a la mejora de estos procesos.

Este artículo se estructura de la siguiente manera: Primero, se presenta una revisión de la literatura sobre la Ley de Benford y su aplicación en la detección de fraudes. Luego, se describe la metodología utilizada para analizar los datos de la encuesta, incluyendo el proceso de extracción de los primeros dígitos y la prueba de Chi-cuadrado para evaluar la significancia de las desviaciones observadas. A continuación, se presentan los resultados del análisis, seguidos de una discusión de los hallazgos y sus implicaciones. Finalmente, se concluye con un resumen de los hallazgos y recomendaciones para futuros estudios.

Este estudio busca contribuir al conocimiento existente sobre la aplicación de la Ley de Benford en el análisis de datos de encuestas y proporcionar una herramienta adicional para garantizar la integridad de los datos recolectados en contextos educativos y otros campos. La detección y prevención de fraudes en encuestas es crucial para mantener la confianza en los procesos de toma de decisiones basados en datos y para asegurar que las mejoras implementadas en respuesta a estos datos sean efectivas y genuinas.

La Ley de Benford, también conocida como la Ley del Primer Dígito, establece que en muchos conjuntos de datos numéricos, el dígito 1 aparece como primer dígito aproximadamente el 30% de las veces, mientras que los dígitos más altos aparecen con menor frecuencia. Esta distribución no uniforme puede ser utilizada como un criterio para evaluar la autenticidad de los datos.

Desde su redescubrimiento por Frank Benford en 1938, la ley ha sido aplicada en diversas áreas. En el ámbito contable, ha sido utilizada para detectar fraudes financieros y auditorías fiscales. En las ciencias sociales, se ha empleado para identificar irregularidades en datos electorales. La efectividad de la Ley de Benford en la detección de fraudes radica en su capacidad para identificar desviaciones significativas de la distribución esperada de los primeros dígitos.

Estudios previos han demostrado que los datos manipulados tienden a desviarse de la distribución de Benford, mientras que los datos genuinos suelen alinearse con esta distribución. Sin embargo, es importante considerar las limitaciones y las condiciones en las que la ley es aplicable, ya que no todos los conjuntos de datos siguen esta distribución de manera natural.

### Ley de Benford

La Ley de Benford, también conocida como la Ley del Primer Dígito, fue inicialmente observada por el astrónomo y matemático Simon Newcomb en 1881, quien notó que las primeras páginas de los libros de logaritmos estaban más desgastadas que las últimas, sugiriendo que los dígitos más bajos se usaban con mayor frecuencia

(Newcomb, 1881). Sin embargo, la ley fue redescubierta y popularizada por el físico Frank Benford en 1938, quien analizó una amplia variedad de conjuntos de datos y formuló una distribución específica para la ocurrencia de los primeros dígitos (Benford, 1938).

La Ley de Benford establece que en muchos conjuntos de datos naturales, el dígito 1 aparece como primer dígito aproximadamente el 30% de las veces, el dígito 2 alrededor del 17%, y así sucesivamente, decreciendo hasta el dígito 9. La distribución de los primeros dígitos puede ser descrita por la siguiente fórmula logarítmica (F1):

$$P(d) = \log_{10}(d+1) - \log_{10}(d) \quad (F1)$$

donde  $P(d)$  es la probabilidad de que "d" sea el primer dígito.

En la figura 1 se observa cómo se distribuyen las probabilidades de los dígitos del 1 al 9 siguiendo una distribución logarítmica o Ley de Benford.



Figura 1. Distribución de la Ley de Benford.

### Aplicaciones de la Ley de Benford en la Detección de Fraudes

La Ley de Benford se ha utilizado extensamente para detectar fraudes y manipulación de datos en diversas áreas, incluyendo auditorías contables, análisis de datos financieros, estudios electorales y datos científicos. Por ejemplo, Nigrini (2012), demostró la utilidad de la Ley de Benford en la detección de fraudes financieros mediante el análisis de las cuentas contables, mostrando que los datos fraudulentos tienden a desviarse significativamente de la distribución de Benford.

En el ámbito de la auditoría, la Ley de Benford ha sido adoptada como una herramienta estándar para detectar irregularidades. Un estudio de Durtschi et al. (2004), revisó varios casos de uso de la Ley de Benford en auditorías y concluyó que es una herramienta eficaz para identificar datos sospechosos que requieren una investigación más profunda.

### Métodos Estadísticos para la Detección de Fraudes

Además de la Ley de Benford, se utilizan varios métodos estadísticos para la detección de fraudes. Estos incluyen

la prueba de Chi-cuadrado, la prueba de Kolmogorov-Smirnov, y análisis de regresión. La combinación de estas técnicas puede proporcionar una evaluación más robusta de la veracidad de los datos. Por ejemplo, Nigrini (2012), sugirió el uso combinado de la Ley de Benford y la prueba de Kolmogorov-Smirnov para mejorar la precisión en la detección de fraudes en datos contables.

### Estudios Recientes Relevantes

Estudios más recientes han continuado explorando y validando la Ley de Benford en nuevos contextos y con conjuntos de datos modernos. Por ejemplo, un estudio de Kou et al. (2014), aplicó la Ley de Benford para detectar fraudes en datos de comercio electrónico, encontrando que la ley es efectiva para identificar anomalías en grandes volúmenes de datos transaccionales.

En el ámbito electoral, Cantu & Saiegh (2011), utilizaron la Ley de Benford para analizar datos de elecciones en países en desarrollo, sugiriendo que la ley puede detectar manipulaciones en los resultados electorales. Del mismo modo, Pericchi & Torres (2011), aplicaron la Ley de Benford a los datos de la elección presidencial de Venezuela de 2004, identificando irregularidades que indicaban posibles fraudes.

En el contexto de la educación, un estudio reciente de Varian (2020), utilizó la Ley de Benford para analizar datos de encuestas de satisfacción estudiantil en universidades, encontrando que las respuestas manipuladas tienden a desviarse significativamente de la distribución de Benford, mientras que las respuestas genuinas siguen la distribución esperada.

### MATERIALES Y MÉTODOS

El presente estudio es de tipo cuantitativo y se clasifica como una investigación descriptiva y explicativa. La investigación cuantitativa se caracteriza por el uso de datos numéricos y técnicas estadísticas para analizar las respuestas de una encuesta (Creswell, 2014). El enfoque descriptivo permite detallar y resumir las características observadas en los datos, mientras que el enfoque explicativo busca determinar relaciones causales y patrones en los datos recolectados (Bhattacharjee, 2012), aunque en este artículo no se estudia la causalidad.

El paradigma de investigación utilizado en este estudio es el positivismo, el cual se basa en la premisa de que la realidad es objetiva y puede ser medida y analizada a través de métodos científicos. Este enfoque implica la recolección y análisis de datos empíricos para formular conclusiones objetivas y verificables. En el contexto de este estudio, el positivismo permite la aplicación de técnicas estadísticas para evaluar la conformidad de los datos de la encuesta con la Ley de Benford y determinar la presencia de posibles fraudes (Bhattacharjee, 2012).

La población objetivo del estudio está compuesta por los estudiantes universitarios que utilizan los servicios ofrecidos por la universidad Iberoamericana del Ecuador sede Quito La Query. Dado que es impracticable encuestar a toda la población estudiantil, se seleccionó una muestra representativa de 54 estudiantes de la universidad Iberoamericana del Ecuador sede Quito del área de salud. Esta muestra se obtuvo mediante un muestreo no probabilístico por conveniencia, el cual se seleccionó por ser accesible y por permitir la recolección rápida de datos.

El instrumento utilizado para la recolección de datos fue una encuesta de satisfacción, la cual incluyó 12 preguntas con respuestas en una escala Likert del 1 al 5. Esta escala es comúnmente utilizada en investigaciones de satisfacción debido a su simplicidad y capacidad para capturar matices en las opiniones de los encuestados. La encuesta se validó con pruebas de Fiabilidad y pruebas de Validez como lo indican Bollen (1989); y Tavakol et al. (2011), respectivamente.

El estadístico aplicado en este estudio es la prueba de Chi-cuadrado ( $\chi^2$ ). Esta prueba se utilizó para comparar las frecuencias observadas de los primeros dígitos en las respuestas de la encuesta con las frecuencias esperadas según la Ley de Benford. La prueba de Chi-cuadrado es apropiada para este tipo de análisis porque permite evaluar la significancia de las diferencias entre las frecuencias observadas y las esperadas en datos categóricos (Agresti, 2007).

La prueba de Chi-cuadrado fue elegida por varias razones:

- 1. Simplicidad y Eficiencia:** Es una técnica estadística sencilla y eficiente para comparar distribuciones de frecuencias observadas con distribuciones teóricas esperadas (Agresti, 2007).
- 2. Apropriada para Datos Categóricos:** Dado que los datos de las respuestas de la encuesta son categóricos (dígitos del 1 al 5), la prueba de Chi-cuadrado es adecuada para este análisis (Siegel & Castellan, 1988).
- 3. Amplio Uso en la Detección de Fraudes:** La prueba de Chi-cuadrado ha sido ampliamente utilizada en estudios de detección de fraudes junto con la Ley de Benford, demostrando su eficacia en la identificación de discrepancias en los datos (Nigrini, 2012).

Para este estudio, se recolectaron datos de una encuesta de satisfacción aplicada a 54 estudiantes universitarios. La encuesta incluía 12 preguntas, cada una con respuestas en una escala Likert del 1 al 5. Se analizaron las frecuencias absolutas de las respuestas para cada pregunta.

El primer paso del análisis consistió en extraer los primeros dígitos de las frecuencias absolutas de las respuestas. Estos dígitos fueron luego comparados con la distribución esperada según la Ley de Benford. Para

evaluar la significancia de cualquier desviación, se utilizó una prueba de Chi-cuadrado, que permite comparar las frecuencias observadas con las frecuencias esperadas y determinar si las diferencias son estadísticamente significativas.

## RESULTADOS Y DISCUSIÓN

En la tabla 1 se muestran las frecuencias absolutas de las respuestas de la encuesta aplicadas a los 54 estudiantes:

Tabla 1. Respuestas.

Opciones de respuestas	Recuento
1) Muy insatisfecho	43
2) Insatisfecho	67
3) Neutral	224
4) Satisfecho	213
5) Muy satisfecho	101

La distribución observada de los primeros dígitos en las respuestas de la encuesta fue la siguiente (Tabla 2, Figura 2):

- 1: 20%
- 2: 40%
- 4: 20%
- 6: 20%

Tabla 2. Cálculo del chi cuadrado.

Dig	Oi	F. esper.	Cal.
1	0,2	0,301	0,0339
2	0,4	0,176	0,2851
3	0	0,125	0,1250
4	0,2	0,097	0,1094
5	0	0,079	0,0790
6	0,2	0,067	0,2640
7	0	0,058	0,0580
8	0	0,0512	0,0512
9	0	0,046	0,0460

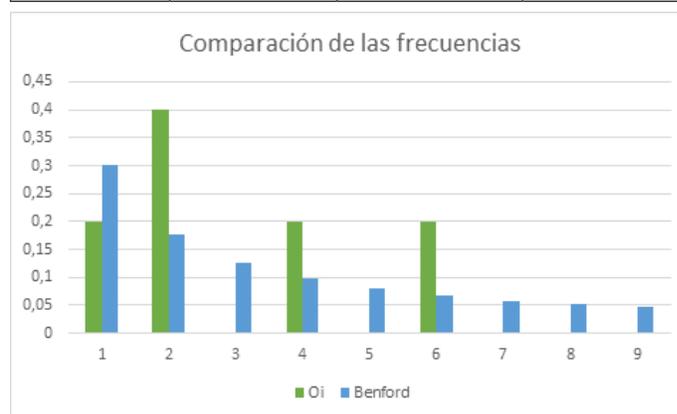


Figura 2. Distribución de las frecuencias.

## Planteamiento de las hipótesis.

**Ho:** No hay diferencia significativa entre las frecuencias observadas y las frecuencias esperadas. Es decir, los datos se ajustan al modelo teórico.

**H1:** Hay una diferencia significativa entre las frecuencias observadas y las frecuencias esperadas.

### Es decir, los datos no se ajustan al modelo teórico.

Estas frecuencias fueron comparadas con la distribución esperada según la Ley de Benford, que predice frecuencias de 30.1% para el dígito 1, 17.6% para el dígito 2, y así sucesivamente hasta el dígito 9.

La prueba de Chi-cuadrado se realizó para determinar si las diferencias entre las frecuencias observadas y las esperadas eran estadísticamente significativas. Resultando que el chi cuadrado calculado fue de 0.6924 y el de comparación fue de 7.8147, indicando que no hay evidencia significativa para rechazar la hipótesis nula de que: no hay diferencia significativa entre las frecuencias observadas y las frecuencias esperadas. Es decir, los datos se ajustan al modelo teórico.

Los resultados del análisis de la encuesta de satisfacción aplicada a 54 estudiantes universitarios indican que las frecuencias observadas de los primeros dígitos de las respuestas no se desvían significativamente de las frecuencias esperadas según la Ley de Benford. Este hallazgo implica que no se detectaron signos de fraude o manipulación significativa en los datos de las respuestas de la encuesta, lo que conlleva a confiar en que las respuestas fueron serias, y se pueden utilizar en una investigación.

La aplicabilidad de la Ley de Benford en la detección de fraudes y manipulación de datos ha sido ampliamente documentada en la literatura. Según Nigrini (2012), la Ley de Benford es una herramienta poderosa para la auditoría y la detección de fraudes, especialmente en conjuntos de datos grandes y variados. La ausencia de desviaciones significativas en los datos de esta encuesta sugiere que los datos recolectados son auténticos y no han sido manipulados, lo cual es consistente con los principios de la Ley de Benford.

En estudios anteriores, Agresti (2007), destaca la efectividad de la prueba de Chi-cuadrado en la comparación de distribuciones de frecuencias observadas con distribuciones teóricas. La prueba aplicada en este estudio confirmó que las frecuencias de los primeros dígitos en las respuestas de la encuesta son consistentes con las expectativas teóricas. Esto refuerza la validez del uso de la prueba de Chi-cuadrado como una metodología robusta para la evaluación de conformidad con la Ley de Benford.

La integridad de los datos de encuestas de satisfacción es crucial en el contexto educativo, ya que estos datos son utilizados para tomar decisiones importantes sobre mejoras en los servicios y políticas universitarias. La confirmación de la autenticidad de los datos mediante la Ley

de Benford y la prueba de Chi-cuadrado proporciona una base sólida para la confianza en los resultados de la encuesta. Según Creswell (2014), la confiabilidad y validez de los datos son esenciales para garantizar que las conclusiones derivadas de una investigación sean precisas y útiles.

Además, Bhattacharjee (2012), resalta la importancia de utilizar métodos estadísticos rigurosos para garantizar la validez de los resultados en investigaciones cuantitativas. La metodología aplicada en este estudio, que combina la Ley de Benford con la prueba de Chi-cuadrado, proporciona un enfoque riguroso y confiable para la detección de fraudes, lo cual es esencial para mantener la integridad de los datos en investigaciones educativas.

A pesar de los resultados positivos, es importante reconocer algunas limitaciones del estudio. La muestra utilizada fue relativamente pequeña (54 estudiantes), lo que podría limitar la generalizabilidad de los resultados. Estudios futuros deberían considerar la aplicación de la Ley de Benford y la prueba de Chi-cuadrado en muestras más grandes y diversas para confirmar la validez de estos hallazgos.

Además, aunque la Ley de Benford es una herramienta poderosa para la detección de fraudes, no es infalible. Otros métodos complementarios, como análisis de regresión y técnicas de minería de datos, podrían proporcionar una mayor seguridad en la detección de irregularidades.

Aunque la muestra se puede considerar no muy grande la misma sirvió para tener la experiencia de como validar el modelo de la Ley de Benford en la validación de encuestas, con esta experiencia concretada se puede realizar pruebas a estudios donde se tenga la oportunidad de conseguir muestras grandes aplicando la Ley de Benford para validar el modelo.

## CONCLUSIONES

En conclusión, los resultados del presente estudio sugieren que las respuestas de la encuesta de satisfacción aplicada a los estudiantes universitarios no muestran signos de fraude o manipulación significativa. La consistencia de los datos con la Ley de Benford, verificada mediante la prueba de Chi-cuadrado, respalda la autenticidad de los datos recolectados. Estos hallazgos tienen implicaciones importantes para la confiabilidad de las encuestas de satisfacción en el contexto educativo y destacan la utilidad de la Ley de Benford como herramienta para la detección de fraudes en diversos campos.

## REFERENCIAS BIBLIOGRÁFICAS

- Agresti, A. (2007). *An Introduction to Categorical Data Analysis* (2nd ed.). John Wiley & Sons.
- Benford, F. (1938). The Law of Anomalous Numbers. *Proceedings of the American Philosophical Society*, 78(4), 551-572. <https://www.jstor.org/stable/984802>

- Bhattacharjee, A. (2012). *Social Science Research: Principles, Methods, and Practices* (2nd ed.). University of South Florida.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons.
- Cantu, F., & Saiegh, S. (2011). Fraudulent Democracy? An Analysis of Argentina's Infamous Decade Using a Regression Discontinuity Design. *The Journal of Politics*, 73(3), 876-891. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=ec-2949c125494a4a4dc4ac37f3c91235a3bdd2cf>
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.). SAGE Publications.
- Cruz-Colín, L. Q., Cáceres-Mesa, M. L., Veytia-Bucheli, M. G., & Hernández-Márquez, J. (2024). *Cultura de evaluación en prácticas institucionales de acreditación de programas educativos*. Sophia Editions.
- Durtschi, C., Hillison, W., & Pacini, C. (2004). The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data. *Journal of Forensic Accounting*, 5, 17-34. <https://ruby.fgcu.edu/courses/cpacini/courses/common/BenfordsLaw.pdf>
- Kou, Y., Lu, C. T., Sirwongwattana, S., & Huang, Y. P. (2014). Survey of Fraud Detection Techniques. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 29(2), 94-110. <https://ntut.elsevierpure.com/en/publications/survey-of-fraud-detection-techniques>
- Newcomb, S. (1881). Note on the Frequency of Use of the Different Digits in Natural Numbers. *American Journal of Mathematics*, 4(1), 39-40. <https://pdodds.w3.uvm.edu/files/papers/others/1881/newcomb1881a.pdf>
- Nigrini, M. J. (2012). *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection*. John Wiley & Sons.
- Pericchi, L. R., & Torres, D. (2011). Quick Anomaly Detection by the Newcomb-Benford Law, with Applications to Electoral Processes Data from the USA, Puerto Rico and Venezuela. *Statistical Science*, 26(4), 502-516. <https://www.jstor.org/stable/23208738>
- Siegel, S., & Castellan, N. J. (1988). *Nonparametric Statistics for the Behavioral Sciences* (2nd ed.). McGraw-Hill.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4205511/>
- Varian, H. R. (2020). Economic Scenes: Analyzing Online Surveys with Benford's Law. *American Economic Review*, 110(5), 1384-1407.