

Pruebas estadísticas de normalidad: un análisis comparativo entre Kolmogórov- Smirnov, Shapiro-Wilk, Anderson-Darling, Ryan-Joiner Y Jarque-Bera

Pablo Alexander Molina Panchi¹; Diego Fabián Molina Panchi²

Resumen

El objetivo de este artículo es aplicar las pruebas de normalidad Kolmogórov-Smirnov, Shapiro-Wilk, Anderson-Darling, Ryan-Joiner y Jarque-Bera a partir de la revisión de la literatura existente en el análisis de una serie de tiempo. La investigación adopta un enfoque mixto que combina el análisis cuantitativo y cualitativo de las variables, permitiendo evaluar el grado de desviación de los datos respecto a una distribución normal. El estudio es de tipo explicativo, proporcionando una comparación detallada de cada prueba y su aplicación en diversos contextos estadísticos aplicados a una serie de datos. Estas pruebas son fundamentales para verificar el supuesto de normalidad en los datos antes de aplicar métodos estadísticos paramétricos. Los principales resultados señalan que la capacidad para determinar si una muestra proviene de una población con distribución normal es crucial para seleccionar técnicas estadísticas adecuadas, asegurando la validez y confiabilidad. Además, facilita la identificación de desviaciones de la normalidad que pueden requerir transformaciones de datos o métodos no paramétricos alternativos. Ligado a ello, los resultados demuestran que estas herramientas son aplicables en diversos ámbitos económicos y sociales entre otros, lo que asegura que los análisis estadísticos sean precisos y adecuados, especialmente en contextos donde la toma de decisiones rápida es esencial.

Palabras clave: Anderson-Darling, Jarque-Bera, Kolmogórov-Smirnov, Pruebas de normalidad, Ryan-Joiner, Shapiro-Wilk.

Statistical tests of normality: a comparative analysis of Kolmogorov-Smirnov, Shapiro-Wilk, Anderson-Darling, Ryan-Joiner, And Jarque-Bera

Abstract

The objective of this article is to apply the normality tests Kolmogorov-Smirnov, Shapiro-Wilk, Anderson-Darling, Ryan-Joiner, and Jarque-Bera based on a review of the existing literature, in the analysis of a time series. The research adopts a mixed-methods approach, combining quantitative and qualitative analysis of the variables, allowing for the assessment of the degree of deviation from a normal distribution. This is an explanatory study that provides a detailed comparison of each test and its application in various statistical contexts applied to a data series. These tests are essential for verifying the assumption of normality in data before applying parametric statistical methods. The main findings indicate that the ability to determine whether a sample comes from a normally distributed population is crucial for selecting appropriate statistical techniques, ensuring both validity and reliability. Furthermore, it facilitates the identification of deviations from normality that may require data transformations or the use of alternative non-parametric methods. In this regard, the results show that these tools are applicable in various economic and social fields, among others, ensuring that statistical analyses are accurate and appropriate especially in contexts where rapid decision-making is essential.

Keywords: Anderson-Darling, Jarque-Bera, Kolmogorov-Smirnov, Normality tests, Ryan-Joiner, Shapiro-Wilk.

Recibido: 12 de noviembre de 2024

Aceptado: 15 de agosto de 2025

¹ Investigador Independiente, Quito-Ecuador, paalmopa92@hotmail.com, <https://orcid.org/0000-0002-1708-8446>

² Investigador Independiente, Latacunga-Ecuador, diegofabian1@hotmail.es, <https://orcid.org/0000-0002-5803-1876>

I. INTRODUCCIÓN

En el análisis estadístico, es esencial asumir que los datos siguen una distribución normal para aplicar correctamente diversas técnicas y métodos paramétricos. La distribución normal con su curva simétrica en forma de campana es fundamental en la teoría estadística y respalda diversos modelos y pruebas inferenciales (Flores y Flores, 2021).

Esta suposición de normalidad permite el uso de herramientas como la prueba t, el análisis de varianza (ANOVA) y la regresión lineal, esenciales en disciplinas como la psicología y la economía. La normalidad facilita la interpretación de los resultados al proporcionar un marco teórico sólido para la estimación de parámetros y la construcción de intervalos de confianza (Molina M. , 2022).

Sin embargo, es importante tener en cuenta que la normalidad no siempre se cumple en los datos del mundo real, lo que puede conducir a conclusiones incorrectas si se emplean métodos paramétricos sin precaución. Por ello, se han desarrollado pruebas de normalidad, como la prueba de Shapiro-Wilk y la de Kolmogórov-Smirnov, para verificar esta suposición. Cuando los datos se desvían significativamente de la normalidad, se pueden emplear métodos no paramétricos, como la prueba de Mann-Whitney o el coeficiente de correlación de Spearman (Gandica, 2021).

También, técnicas de transformación de datos y métodos de remuestreo como el bootstrap pueden ser útiles en situaciones donde la normalidad no se cumple rigurosamente. En última instancia, comprender la importancia y las limitaciones de la suposición de normalidad es crucial para realizar análisis estadísticos precisos y confiables (Flores et al., 2019).

Más allá de su utilidad práctica, la distribución normal es fundamental en la teoría estadística debido a su estructura matemática y propiedades únicas. Es decir, la simetría de su curva alrededor de la media, junto con sus colas infinitas, permite modelar fenómenos naturales y sociales con gran precisión (Dagnino, 2014).

El objetivo de este artículo es aplicar las pruebas de normalidad Kolmogórov-Smirnov, Shapiro-Wilk, Anderson-Darling, Ryan-Joiner y Jarque-Bera a partir de la revisión de la literatura existente en el análisis de una serie de tiempo.

La pregunta que responde la investigación es: ¿Cuáles son las características, ventajas, limitaciones y efectividad comparativa de las diferentes pruebas estadísticas de normalidad Kolmogórov-Smirnov, Shapiro-Wilk, Anderson-Darling, Ryan-Joiner y Jarque-Bera cuando se aplican a una serie de tiempo?

Además, el Teorema del Límite Central, uno de los resultados más relevantes en estadística, establece que, bajo ciertas condiciones, la distribución de la media muestral de muchas variables aleatorias independientes se aproxima a una distribución normal, independientemente de la distribución original de estas variables. Esto resalta por qué la distribución normal es predominante en el análisis de datos en diversas disciplinas científicas (Delgado, 2022).

El concepto de normalidad fue realizado por Abraham De Moivre en 1733 y posteriormente ampliado por Laplace en 1812, conocido como el Teorema De Moivre-Laplace. En 1805, Legendre introdujo el método de mínimos cuadrados, que Gauss justificó en 1809 al asumir que los errores seguían una distribución normal. El término "*bell surface*" fue usado por Esprit Jouret en 1872 para describir una distribución normal bivalente y el nombre de distribución normal comenzó a utilizar hacia 1875 (Platas, 2021).

En efecto, el autor De Moivre desarrolló la curva normal como una aproximación a la distribución binomial, lo que muestra la conexión entre conceptos estadísticos. Luego, Gauss aplicó la distribución normal al análisis de errores en mediciones astronómicas, estableciendo su utilidad en las ciencias naturales. Esta base histórica facilitó que la distribución normal se convirtiera en un componente esencial de la estadística moderna, dado que sus propiedades matemáticas facilitan el cálculo de probabilidades y la estimación de parámetros (Soto y Herrera, 2024).

En la práctica, los datos a menudo no se ajustan perfectamente a una distribución normal, lo que subraya la necesidad de utilizar herramientas fiables para evaluar la normalidad antes de realizar análisis complejos. Dentro de las mencionadas pruebas de normalidad aplicadas en este estudio tenemos la de Kolmogórov-Smirnov, Shapiro-Wilk, Anderson-Darling, Jarque-Bera y Ryan-Joiner.

De igual forma, se debe mencionar que la falta de ajuste perfecto de los datos a la distribución normal plantea desafíos en la práctica estadística, lo que ha llevado al desarrollo de diversos métodos para evaluar la normalidad (Lind et al., 2012). Frente a ello, la prueba de Kolmogórov-Smirnov compara la distribución acumulativa empírica con la teórica, mientras que la prueba de Shapiro-Wilk se basa en la correlación entre los datos y las puntuaciones normales correspondientes.

En este caso, la elección de la prueba depende de factores como el tamaño de la muestra y la naturaleza de los datos, siendo crucial entender estas diferencias para una evaluación precisa de la normalidad (Schaum, 1991). La verificación de la normalidad es esencial no solo para asegurar la validez de las pruebas estadísticas, sino también para evitar conclusiones incorrectas y malas interpretaciones (Pek et al., 2017).

Al mismo tiempo, la infracción de los supuestos de normalidad puede resultar en estimaciones sesgadas y en intervalos de confianza inexactos, llevando a conclusiones erróneas sobre relaciones entre variables. Las limitaciones de estos tests subrayan la necesidad de un enfoque crítico y matizado en la evaluación de la normalidad, combinando pruebas formales con métodos gráficos y consideraciones teóricas (Webster, 2000).

En la práctica para evaluar la normalidad de los datos, se combinan técnicas gráficas, pruebas estadísticas y el juicio de expertos. La verificación de la normalidad es fundamental para asegurar que las pruebas estadísticas sean válidas y también para entender mejor la estructura de los datos. Si los datos no siguen una distribución normal, se pueden aplicar transformaciones o utilizar métodos estadísticos no paramétricos como alternativas (Molina et al., 2023).

Esta combinación de métodos gráficos, como gráficos Q-Q e histogramas, y pruebas formales ofrece una evaluación más completa de la normalidad. Los métodos gráficos proporcionan una representación visual intuitiva de la distribución de los datos, mientras que las pruebas formales ofrecen una evaluación cuantitativa rigurosa (Lemus y Huincahue, 2019).

Los diferentes tests de normalidad abordan fundamentos teóricos, aplicaciones prácticas y limitaciones. En este contexto la prueba de Shapiro-

Wilk se basa en la regresión de las estadísticas de orden sobre sus valores esperados, mientras que la prueba de Anderson-Darling da más peso a las colas de la distribución (Molina et al., 2023). Por último, la estructura de este artículo de investigación se encuentra compuesta por la introducción, metodología, resultados, discusión y conclusiones.

II. METODOLOGÍA

La investigación presenta un enfoque mixto porque combina el análisis de variables cuantitativas, utilizando estadísticos descriptivos y pruebas de normalidad como Shapiro-Wilk, Kolmogórov-Smirnov, Anderson-Darling, Ryan-Joiner y Jarque-Bera, junto con la evaluación de variables cualitativas relacionadas con la interpretación visual de gráficos Q-Q e histogramas.

Este enfoque es caracterizado por una distribución de los datos que permite medir el grado de desviación de la normalidad (Romero et al., 2020). El estudio se basa en una revisión exhaustiva de la literatura pertinente al tema en cuestión. Para ello, se emplean técnicas de búsqueda y análisis documental que incluyen el uso de bases de datos académicas reconocidas como Scopus, Web of Science, Latinindex y Google Scholar, las cuales aseguran la obtención de información precisa y fiable (Hernández, 2010).

El estudio es de tipo explicativo, debido a que detalla los diferentes tipos de pruebas de normalidad con el objetivo de comparar sus fortalezas, debilidades y su aplicación en diversos contextos estadísticos (Bernal, 2010). En el presente estudio, se procederá a corroborar la teoría mediante un análisis riguroso de los datos recopilados, específicamente en relación con el número de casos de homicidios intencionales.

Para validar las hipótesis, se aplicaron pruebas de normalidad mediante los programas estadísticos EViews, Minitab y SPSS, asegurando la precisión de los resultados. Este enfoque permitirá verificar la consistencia teórica y proporcionar una base sólida para las conclusiones del estudio, asegurando que los hallazgos y conclusiones estén respaldados por evidencia empírica confiable.

III. RESULTADOS Y DISCUSIÓN

Las pruebas de normalidad se fundamentan en la comparación entre la distribución de los datos

observados y una distribución normal teórica. Este principio fundamental de las pruebas de normalidad en estadística funciona comparando la distribución de los datos reales recolectados con una distribución normal ideal o teórica (Muñoz et al., 2021).

Este proceso de cuantificación permite evaluar estadísticamente si es razonable asumir que los datos provienen de una población con distribución normal. También, la importancia de estas pruebas radica en que muchos análisis estadísticos asumen normalidad en los datos, por lo que confirmar o rechazar esta suposición es crucial para elegir los métodos analíticos apropiados y para la validez de las conclusiones derivadas de los análisis estadísticos subsecuentes (Padilla, 2022).

Generalmente, la hipótesis nula de estas pruebas estadísticas sostiene que la muestra proviene de una población que sigue una distribución normal. Otro aspecto crucial, en la aplicación de las pruebas de normalidad es la elección del nivel de significancia (α). El valor de α más comúnmente utilizado es 0,05; lo cual implica que se acepta un 5,00% de probabilidad de cometer un error de Tipo I (rechazar la hipótesis nula de normalidad cuando en realidad es verdadera).

Sin embargo, el valor no es fijo ni universal, sino que puede y debe ajustarse según las particularidades y necesidades específicas de cada investigación.

Además, esta flexibilidad en la elección del nivel de significancia refleja la importancia de considerar el contexto, los objetivos y las implicaciones del estudio al momento de tomar decisiones estadísticas (Lemus y Huincahue, 2019).

Métodos gráficos de las pruebas de normalidad

Los métodos gráficos, aunque son más subjetivos, proporcionan una visión valiosa sobre la forma de la distribución de los datos. A continuación, se analizan los Gráficos Q-Q (Quantile-Quantile) y los histogramas. En primer lugar, los gráficos Q-Q (Quantile-Quantile) comparan los cuantiles de los datos observados con los cuantiles teóricos de una distribución normal. Si los puntos del gráfico se alejan significativamente de una línea recta, esto indica que los datos podrían no ajustarse a una distribución normal (Molina et al., 2023).

Además, los gráficos Q-Q ofrecen una herramienta visual eficaz para evaluar la normalidad de los datos. Por último, la línea recta en estos gráficos representa la distribución normal teórica, y cualquier desviación notable de esta línea sugiere que los datos observados difieren de lo esperado en una distribución normal (Marques et al., 2023). La Figura 1 presenta el Gráfico Q-Q normal correspondiente a la serie histórica de homicidios intencionales.

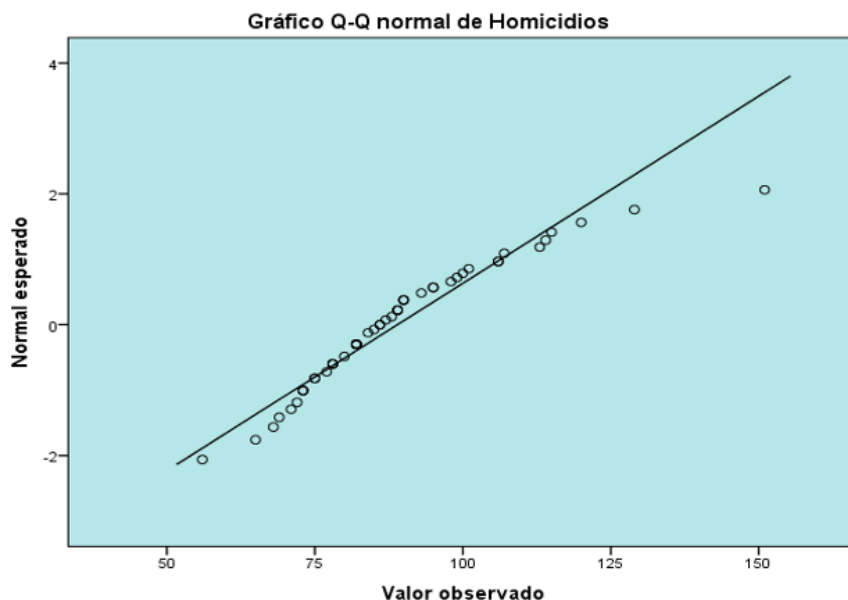


Figura 1. Nube de palabras

Nota. Los datos muestran el Gráfico Q-Q normal de la serie histórica de homicidios intencionales (enero 2014 - febrero 2018). Fuente: Ministerio del Interior, (2024).

La figura 1 presentada muestra un gráfico Q-Q (cuantil-cuantil) normal para los datos de homicidios intencionales. Este tipo de gráfico es utilizado para evaluar visualmente si un conjunto de datos sigue una distribución normal. En el gráfico, la línea diagonal recta representa la distribución normal teórica. Si los datos siguieran perfectamente esta distribución, todos los puntos se situarían exactamente sobre esta línea.

Cada punto en el gráfico corresponde a un dato observado, y su posición en relación con la línea de referencia indica qué tan cerca están los datos de ajustarse a una distribución normal. En general, los puntos siguen bastante de cerca la línea diagonal, lo que sugiere que los datos se aproximan a una distribución normal.

En cuanto a las desviaciones, en los valores más bajos (extremo inferior izquierdo), los puntos tienden a estar ligeramente por debajo de la línea. En los valores más altos (extremo superior derecho),

algunos puntos se desvían más notablemente de la línea, especialmente los dos puntos más altos. Además, la desviación de los puntos en el extremo superior derecho sugiere la presencia de algunos valores más altos de lo esperado en una distribución normal perfecta, lo que podría indicar una cola ligeramente más pesada en la parte superior de la distribución.

Aunque los datos parecen seguir aproximadamente una distribución normal, hay algunas desviaciones, especialmente en los extremos. En consecuencia, la distribución de homicidios se aproxima a la normal, aunque presenta desviaciones. En este contexto, las desviaciones en los extremos, particularmente en los valores más altos, podrían indicar la presencia de algunos valores atípicos o una ligera asimetría en la distribución. La Figura 2 presenta el Gráfico Q-Q normal sin tendencia correspondiente a la serie histórica de homicidios intencionales.

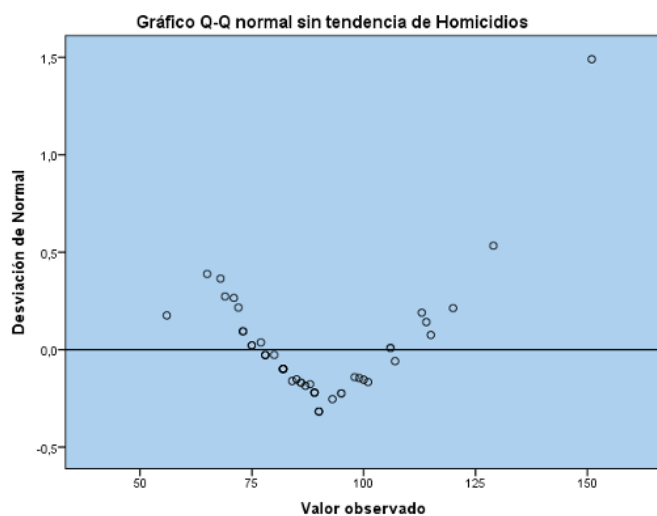


Figura 2. Gráfico Q-Q normal sin tendencia

Nota. Los datos muestran el Gráfico Q-Q normal sin tendencia de la serie histórica de homicidios intencionales (enero 2014 - febrero 2018). Fuente: Ministerio del Interior, (2024).

El gráfico presentado es un gráfico Q-Q normal sin tendencia de homicidios. Este tipo de gráfico es una variación del gráfico Q-Q estándar y se utiliza para evaluar la normalidad de los datos de una manera más detallada, eliminando la tendencia lineal para resaltar las desviaciones de la normalidad.

En este gráfico, el eje X representa los valores observados de homicidios, mientras que el eje Y muestra la desviación de la normal. La línea horizontal en el valor 0 del eje Y representa la

distribución normal perfecta. Los puntos por encima de esta línea indican valores mayores a los esperados en una distribución normal, mientras que los puntos por debajo indican valores menores.

Analizando la distribución de los puntos, se observa una forma de "S" invertida. Esta forma es característica y sugiere que los datos tienen colas más pesadas que una distribución normal perfecta. En los valores más bajos (izquierda del gráfico) y más altos (derecha del gráfico), los puntos tienden a estar por

encima de la línea del cero, indicando valores más extremos de lo esperado en una distribución normal.

En la parte central del gráfico, aproximadamente entre los valores observados de 75 y 125, los puntos tienden a estar por debajo de la línea del cero. Esto sugiere que hay una concentración de valores en torno a la media mayor de lo que se esperaría en una distribución normal perfecta. Se observan algunos valores atípicos, particularmente en el extremo superior derecho del gráfico, donde hay un punto que se desvía considerablemente de la tendencia general. Este punto podría representar un caso de homicidio excepcionalmente alto en comparación con el resto de los datos.

Este gráfico Q-Q sin tendencia refuerza la idea de que los datos de homicidios se desvían de una distribución normal perfecta. Las desviaciones son más pronunciadas en los extremos, lo que indica colas más pesadas, y hay una concentración ligeramente mayor de valores cerca de la media. Estos hallazgos son consistentes con los resultados de la prueba

de Shapiro-Wilk y el gráfico Q-Q normal anterior, proporcionando una visión más detallada de cómo los datos se apartan de la normalidad.

Esta consideración es crucial porque pone de relieve la posible subjetividad en la evaluación visual de la normalidad a través de histogramas. Al mismo tiempo, un número inadecuado de intervalos puede alterar la percepción de la forma de la distribución: un exceso de intervalos puede dar lugar a una apariencia fragmentada y ruidosa, mientras que un número insuficiente puede ocultar características importantes de la distribución.

Por lo tanto, se enfatiza la importancia de elegir cuidadosamente el número de intervalos en un histograma y se sugiere que los investigadores utilicen múltiples representaciones o métodos complementarios para evaluar la normalidad, en lugar de depender únicamente de una visualización del histograma. La Figura 3 presenta el Gráfico de histograma correspondiente a la serie histórica de homicidios intencionales.

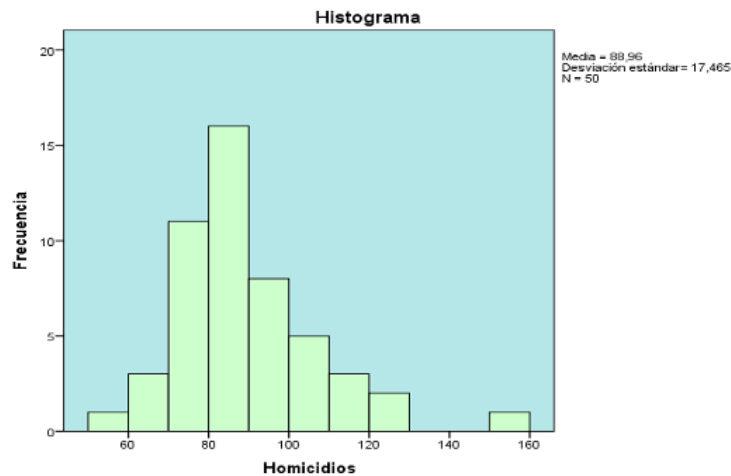


Figura 3. Gráfico de histograma

Nota. Los datos muestran el Gráfico histograma de la serie histórica de homicidios intencionales (enero 2014 - febrero 2018). Fuente: Ministerio del Interior, (2024)

El histograma presentado muestra la distribución de frecuencias de homicidios, ofreciendo una visión general de los datos analizados. La forma de la distribución se asemeja a una campana, pero con una ligera asimetría hacia la derecha, lo que indica una cola más larga en los valores más altos de homicidios.

La tendencia central de los datos se puede observar claramente, con la barra más alta del histograma situada aproximadamente entre 80 y 90 homicidios. Esto se alinea con la media reportada de

88.96, que aparece en la información proporcionada en el gráfico. La desviación estándar de 17.465 sugiere una variabilidad considerable en los datos, lo que se refleja en la amplitud de la distribución en el histograma.

Se debe agregar que, el rango de los datos se extiende desde aproximadamente 60 hasta 160 homicidios, con la mayoría de los casos concentrados entre 70 y 110. Esta distribución muestra una ligera asimetría positiva, evidenciada por la presencia

de algunos casos que se extienden más allá de 120 homicidios, formando una cola más larga en el extremo derecho del gráfico. La categoría con la frecuencia más alta, que podría considerarse la moda de la distribución, parece estar en el rango de 80-90 homicidios.

El gráfico también indica que se están analizando 50 casos de homicidios mensuales ($N = 50$), lo cual proporciona un contexto importante sobre el tamaño de la muestra. Se observan algunos casos en el extremo derecho, cerca de 160 homicidios, que podrían considerarse como valores atípicos o outliers. Estos casos contribuyen a la asimetría de la distribución y podrían representar situaciones excepcionales o extremas en los datos de homicidios.

Aunque la distribución se aproxima a una forma normal, la ligera asimetría y la presencia de posibles outliers sugieren una desviación de la normalidad

perfecta. Esta observación es consistente con los análisis previos realizados mediante los gráficos Q-Q, reforzando la conclusión de que los datos de homicidios, si bien se acercan a una distribución normal, presentan ciertas características que los alejan de una normalidad estricta.

Estadísticos descriptivos

Las pruebas estadísticas formales para evaluar la normalidad son esenciales en el análisis de datos, debido a que permiten a los investigadores determinar con precisión si una muestra proviene de una población con distribución normal (Lugo y Pino, 2022). Sin embargo, antes de utilizar estas pruebas de normalidad, se recomienda analizar los estadísticos descriptivos. A continuación, en la Tabla 1 se presentan los estadísticos descriptivos:

Tabla 1. Estadísticos descriptivos

Variable	Descriptivos	Estadístico	Error estándar
Homicidios	Media	88,96	2,470
	95% de intervalo de confianza para la media	Límite inferior	84,00
		Límite superior	93,92
	Media recortada al 5%	87,84	
	Mediana	86,00	
	Varianza	305,019	
	Desviación estándar	17,465	
	Mínimo	56	
	Máximo	151	
	Rango	95	
	Rango intercuartil	21	
	Asimetría	1,188	,337
	Curtosis	2,320	,662

Nota. Los datos muestran los estadísticos descriptivos de la serie histórica de homicidios intencionales (enero 2014 - febrero 2018). Fuente: Ministerio del Interior, (2024).

Los estadísticos descriptivos ofrecen una visión detallada de la distribución de los homicidios intencionales. La media de homicidios es de 88,96; con un error estándar de 2,47. Esto nos da una estimación del promedio de homicidios en la población estudiada. El intervalo de confianza del 95% para la media se sitúa entre 84,00 y 93,92, lo que significa que podemos estar 95% seguros de que la verdadera media poblacional se encuentra dentro de este rango.

La media recortada al 5% es de 87.84, ligeramente

inferior a la media aritmética, lo que sugiere que algunos valores extremos están inflando la media general. La mediana es de 86.00, también algo menor que la media, lo que refuerza la idea de una ligera asimetría positiva en la distribución.

La varianza de 305.019 y la desviación estándar de 17.465 indican una dispersión considerable en los datos. El rango de los datos es amplio, con un mínimo de 56 y un máximo de 151 homicidios, abarcando un total de 95 unidades. El rango intercuartil de 21 nos da una idea de la dispersión del 50% central de los datos.

La asimetría positiva de 1.188 confirma que la distribución está sesgada hacia la derecha, con una cola más larga en los valores altos. Esto es consistente con las observaciones previas del histograma. La curtosis de 2.320 sugiere que la distribución es leptocúrtica, es decir, tiene colas más pesadas y un pico más pronunciado que una distribución normal perfecta.

Estos estadísticos descriptivos corroboran y cuantifican las observaciones realizadas en los análisis gráficos anteriores (Álvarez y Barreda, 2020). La distribución de homicidios, aunque cercana a la normal, presenta características de asimetría positiva y mayor concentración de datos alrededor de la media de lo que se esperaría en una distribución normal perfecta. La presencia de valores extremos en el lado derecho de la distribución está influyendo en las medidas de tendencia central y dispersión.

Sobre la base de lo antes expuesto, se debe mencionar que esta evaluación es vital porque muchos métodos estadísticos paramétricos presuponen que los datos siguen una distribución normal. Entre las pruebas más utilizados están Shapiro-Wilk, Kolmogórov-Smirnov, Anderson-Darling y Jarque-Bera, cada uno con sus propias características, ventajas y limitaciones. A continuación, se describe cada uno:

La prueba de Shapiro-Wilk

La prueba de Shapiro-Wilk es una herramienta estadística esencial para evaluar la normalidad de un conjunto de datos, lo cual es fundamental en muchos análisis estadísticos. Esta prueba se basa en la comparación entre los cuantiles de una distribución normal teórica y los cuantiles observados en los datos, lo que permite detectar desviaciones de la normalidad en diferentes partes de la distribución.

Se calcula como una razón entre dos estimaciones de la varianza, y valores de W cercanos a 1 sugieren una mayor probabilidad de que los datos sigan una distribución normal. El p-valor asociado ayuda a decidir sobre la hipótesis nula de normalidad (Romero M. , 2016).

Aunque inicialmente fue diseñado para muestras de hasta 50 observaciones, versiones posteriores han ampliado su aplicabilidad a muestras más grandes. Sin embargo, sigue siendo particularmente efectivo en muestras pequeñas y medianas, donde otras

pruebas pueden ser menos confiables. Comparado con otros tests de normalidad, como Kolmogórov-Smirnov o Anderson-Darling, la prueba de Shapiro-Wilk suele mostrar mayor potencia estadística, lo que significa que es más eficaz para detectar desviaciones de la normalidad cuando estas existen realmente.

Sin embargo, su potencia disminuye con muestras muy pequeñas y grandes, puede detectar desviaciones mínimas de la normalidad que podrían no ser significativas en la práctica (Pedrosa et al., 2015).

Este análisis detallado de la prueba de Shapiro-Wilk destaca su importancia como una herramienta estadística sofisticada y versátil, cuyo valor trasciende la simple evaluación de la normalidad, convirtiéndolo en un recurso fundamental para cualquier analista de datos o investigador que trabaje con estadísticas inferenciales. A continuación, en la ecuación 1 se muestra la fórmula:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

Donde:

n= Es el tamaño de la muestra

$x_{(i)}$ = Representa los datos ordenados de menor a mayor

$\bar{x}$ = Es la media de los datos observados

a_i = Son los coeficientes que dependen de los valores de los cuantiles de la distribución normal estándar y se calculan mediante una fórmula específica que involucra la matriz de covarianza de estos cuantiles.

Su alta potencia lo convierte en una opción ideal para validar los supuestos de normalidad en análisis paramétricos con muestras de hasta 50 observaciones. En la tabla 2, se aplica la prueba de normalidad de Shapiro-Wilk.

Tabla 2. Prueba de normalidad Shapiro-Wilk

Variable	Shapiro-Wilk	gl	Sig.
Homicidios	,929	50	,005

Nota. Los datos muestran la prueba de normalidad de la serie histórica de homicidios intencionales (enero 2014 - febrero 2018). Fuente: Ministerio del Interior, (2024).

La tabla 2 muestra el estadístico de Shapiro-Wilk (0,929) está cerca de 1, lo que generalmente sugiere una aproximación a la normalidad. Los grados de

libertad (50) indican el tamaño de la muestra. El valor de significancia (0,005) es menor que el nivel de significancia típico de 0,05.

Además, esto sugiere que debemos rechazar la hipótesis nula de normalidad. Basándonos en estos resultados, podemos concluir que la distribución de los datos de homicidios probablemente no sigue una distribución normal. Esto se debe a que el valor p (0,005) es menor que 0,05, lo que indica una desviación significativa de la normalidad.

Kolmogórov-Smirnov

La prueba de Kolmogórov-Smirnov (K-S) es una herramienta estadística ampliamente utilizada para comparar una distribución de datos observados con una distribución teórica, siendo la distribución normal una de las más comunes en estas comparaciones. El fundamento de la prueba K-S se basa en la comparación entre la función de distribución acumulativa empírica (ECDF) de los datos observados y la función de distribución acumulativa teórica (CDF) de la distribución de referencia (Moraguez et al., 2017).

El estadístico D de la prueba K-S se define como la máxima diferencia absoluta entre estas dos funciones de distribución acumulativas, y se expresa matemáticamente con la siguiente fórmula:

$$D = \max |F_0(x) - S_n(x)| \quad (2)$$

Donde:

$F_0(x)$ = Es la función de distribución acumulativa teórica.

$S_n(x)$ = Es la función de distribución acumulativa empírica, calculada a partir de la muestra de datos.

D = Es la función de distribución acumulativa teórica.

En este contexto, una de las principales ventajas de la prueba K-S es su capacidad para evaluar el ajuste de los datos frente a una variedad de distribuciones

teóricas, no limitándose solo a la distribución normal. Esto lo hace especialmente útil en áreas donde las distribuciones de datos pueden presentar patrones diversos, como en las ciencias sociales, económicas y naturales.

Además, la prueba K-S es particularmente efectiva en muestras grandes, donde puede detectar pequeñas desviaciones respecto a la distribución teórica. Sin embargo, en muestras muy grandes, incluso desviaciones mínimas pueden resultar estadísticamente significativas, lo que requiere una interpretación cuidadosa según el contexto del problema analizado (Moraguez et al., 2017).

En comparación con la prueba de Shapiro-Wilk, el K-S suele ser menos sensible para detectar desviaciones de la normalidad en muestras pequeñas, ya que el Shapiro-Wilk está diseñado específicamente para evaluar la normalidad y utiliza información más detallada sobre la forma de la distribución. Sin embargo, la flexibilidad del K-S para probar múltiples distribuciones compensa esta limitación en muchos casos prácticos.

La implementación de la prueba K-S en la mayoría de los paquetes estadísticos modernos ha facilitado su uso generalizado. Además, existen variantes como la prueba de Lilliefors (una adaptación del K-S cuando los parámetros de la distribución teórica se estiman a partir de la muestra) y la prueba de Cramér-von Mises, que ofrecen alternativas para situaciones específicas (Martínez y Carleos, 2011).

En la práctica, es común y recomendable combinar la prueba K-S con métodos gráficos, como los gráficos Q-Q y P-P, para obtener una visión más completa de cómo los datos observados se desvían de la distribución teórica. Esta combinación permite no solo identificar si existe una diferencia significativa, sino también visualizar dónde y cómo se presentan estas diferencias a lo largo de la distribución. En la tabla 3, se aplica la prueba de normalidad de Kolmogórov-Smirnov.

Tabla 3. Prueba de normalidad Kolmogórov-Smirnov

		Homicidios
N		120
Parámetros normales ^{a,b}	Media	194,82
	Desviación estándar	193,536
Máximas diferencias extremas	Absoluta	,313
	Positivo	,313
	Negativo	-,243
Estadístico de prueba		,313
Sig. asintótica (bilateral)		,000 ^c

a. La distribución de prueba es normal.

b. Se calcula a partir de datos.

c. Corrección de significación de Lilliefors.

Nota. Los datos muestran la prueba de normalidad de la serie histórica de homicidios intencionales (2014 - 2023).

Fuente: Ministerio del Interior, (2024).

El análisis de la tabla de la prueba de Kolmogórov-Smirnov para la variable "Homicidios" revela información crucial sobre la distribución de los datos. En primer lugar, la muestra consta de 120 observaciones, lo que proporciona una base sólida para el análisis. Al examinar los parámetros de la distribución normal, se observa una media de 194,82 y una desviación estándar de 193,536. Es importante señalar que estos valores sugieren una dispersión considerable en los datos, dado que la desviación estándar es prácticamente igual a la media.

En cuanto a las máximas diferencias extremas, el valor absoluto más alto (0,313) se establece como el estadístico de prueba D. Este valor es fundamental para la interpretación de los resultados. La significación asintótica (bilateral) de 0,000 indica un p-valor extremadamente bajo ($< 0,001$), lo que conlleva al rechazo de la hipótesis nula.

Es relevante mencionar que la prueba se realizó contra una distribución normal, utilizando parámetros calculados a partir de los datos de la muestra. Además, se aplicó la corrección de Lilliefors, lo cual es apropiado en casos donde los parámetros de la distribución se estiman a partir de la muestra misma.

La interpretación de estos resultados conduce a varias conclusiones significativas. En primer lugar, el p-valor extremadamente bajo proporciona evidencia estadística robusta para rechazar la hipótesis de que los datos siguen una distribución normal. La diferencia entre los valores positivo (0,313) y negativo (-0,243) sugiere una posible asimetría positiva en la

distribución de los datos.

Otro aspecto es la alta variabilidad en los datos, evidenciada por la desviación estándar casi igual a la media. Esta característica podría indicar una distribución con cola larga hacia la derecha, típica en datos de conteo como los homicidios.

La falta de normalidad en los datos sugiere la necesidad de considerar métodos no paramétricos o transformaciones de datos para análisis posteriores. Además, sería pertinente explorar si los datos se ajustan mejor a otras distribuciones teóricas más apropiadas para conteos, como la distribución de Poisson o la binomial negativa (Arguello et al., 2023).

No obstante, es importante reconocer las limitaciones de este análisis. Si bien la prueba de Kolmogórov-Smirnov es útil, sería recomendable complementar estos resultados con métodos gráficos, como histogramas o gráficos Q-Q, para obtener una comprensión más completa y matizada de la distribución de los datos. Esta prueba de Kolmogórov-Smirnov proporciona evidencia estadística sólida de que los datos de homicidios no siguen una distribución normal (Cruz et al., 2014).

Este hallazgo tiene implicaciones significativas para cualquier análisis estadístico posterior que se realice con estos datos, y subraya la importancia de seleccionar cuidadosamente los métodos estadísticos apropiados para su tratamiento y análisis (Cruz et al., 2014).

Anderson-Darling

La prueba de Anderson-Darling representa un

avance significativo en el ámbito de las pruebas de bondad de ajuste, evolucionando notablemente desde la prueba de Kolmogórov-Smirnov. Su principal distinción radica en la ponderación diferencial que aplica a las colas de la distribución, lo que se refleja en una fórmula matemática más compleja (Mares y Domínguez, 2022):

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\ln F(Y_i) + \ln(1 - F(Y_{n+1-i}))] \quad (3)$$

Donde:

n = El tamaño de la muestra.

Y_i = es el valor ordenado de la i -ésima observación.

F = Es la función de distribución acumulada (CDF) de la distribución normal bajo prueba.

Esta fórmula mide las discrepancias entre la distribución empírica y la teórica, dando mayor peso a las observaciones en los extremos de la distribución. Esta característica hace que la prueba de Anderson-Darling sea especialmente sensible a las desviaciones en las colas, lo que resulta muy útil en varios contextos aplicados.

En el sector financiero, por ejemplo, la capacidad de la prueba de Anderson-Darling para examinar detalladamente las colas de la distribución lo convierte en una herramienta valiosa. Los eventos extremos que se representan en estas colas tienen

un impacto crucial en la gestión de riesgos y en la valoración de activos. De manera similar, en la industria de seguros, donde es vital estimar eventos poco frecuentes, pero de gran impacto, la sensibilidad de la prueba a las colas proporciona información crítica para la modelización actuarial y la determinación de primas.

Una ventaja destacada de la prueba de Anderson-Darling es su flexibilidad en cuanto al tamaño de la muestra. A diferencia de otros tests que pueden perder precisión o potencia en muestras muy pequeñas o grandes, el Anderson-Darling mantiene su capacidad para detectar desviaciones en un amplio rango de tamaños muestrales. Esto lo convierte en una herramienta robusta y adaptable a diversas situaciones de investigación y análisis de datos (Indira y Llinás, 2014).

En el ámbito más amplio de la estadística inferencial, la prueba de Anderson-Darling complementa y enriquece las herramientas disponibles para la evaluación de normalidad y ajuste de distribuciones. Su uso en combinación con otros métodos, como el Shapiro-Wilk para muestras pequeñas o el Kolmogórov-Smirnov para una evaluación general, permite un análisis más completo y matizado de las características de las distribuciones de datos. La Figura 4 presenta el Gráfico de la prueba Anderson-Darling.

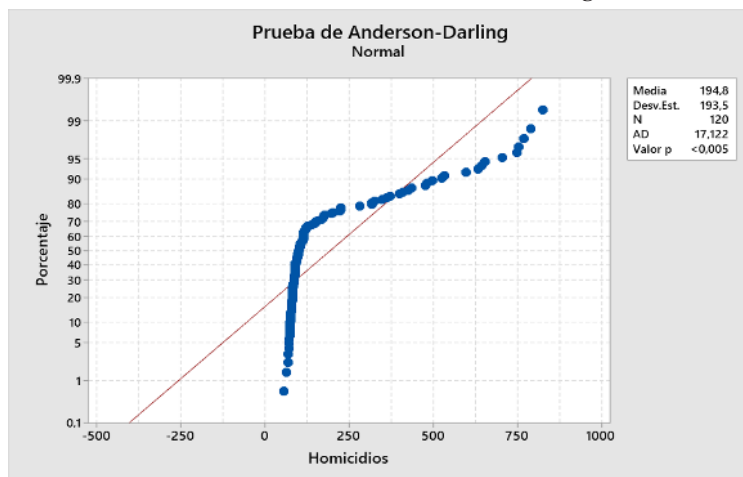


Figura 4. Gráfico de la prueba Anderson-Darling

Nota. Los datos muestran el gráfico de la prueba Anderson-Darling de la serie histórica de homicidios intencionales (2014-2023).

Fuente: Ministerio del Interior, (2024).

La gráfica de la prueba de Anderson-Darling para la normalidad de los datos de homicidios revela información crucial sobre su distribución. En primera

instancia, se observa una clara desviación de los puntos azules, que representan los datos observados, respecto a la línea roja diagonal que indicaría una

distribución normal perfecta. Esta desviación es particularmente pronunciada en los extremos del gráfico, sugiriendo una divergencia significativa de la normalidad.

La forma de "S" invertida que trazan los puntos es indicativa de una asimetría positiva en la distribución, con una cola larga extendiéndose hacia la derecha. Esta característica se ve reforzada por la alta concentración de puntos en la parte inferior izquierda del gráfico, lo que señala que una proporción considerable de los valores se encuentra por debajo de la media. Adicionalmente, la presencia de puntos aislados en la parte superior derecha del gráfico sugiere la existencia de valores atípicos altos, que contribuyen a la asimetría observada.

Los estadísticos proporcionados en la gráfica ofrecen un respaldo cuantitativo a estas observaciones visuales. Con una media de 194.8 y una desviación estándar de 193.5 para una muestra de 120 observaciones, se evidencia una alta variabilidad en los datos. El estadístico de Anderson-Darling (AD) de 17.122 es notablemente elevado, y el valor p asociado, inferior a 0.005, indica un rechazo contundente de la hipótesis nula de normalidad.

En conjunto, tanto la representación gráfica como los estadísticos calculados proporcionan evidencia sólida y congruente de que los datos de homicidios se alejan significativamente de una distribución normal. La asimetría positiva, la presencia de posibles valores atípicos y la alta variabilidad son características prominentes de esta distribución. Estos hallazgos tienen implicaciones importantes para cualquier análisis estadístico posterior, sugiriendo la necesidad de considerar métodos robustos o no paramétricos

que no dependan del supuesto de normalidad para el tratamiento adecuado de estos datos.

Jarque-Bera

Por su parte, la prueba de Jarque-Bera se basa en los momentos de tercer y cuarto orden de la distribución, es decir, asimetría y curtosis (Flores y Flores, 2021). En donde, su fórmula se detalla a continuación:

$$JB = \frac{n}{6} + \left[S^2 + \frac{(K-3)^2}{4} \right] \quad (4)$$

Donde:

n: Representa el tamaño de la muestra (Número total de observaciones).

S: Es el coeficiente de asimetría (Skewness) de la muestra.

K: Es el coeficiente de curtosis de la muestra. (Curtosis de una distribución es 3, por lo que k-3 es la desviación de la curtosis respecto a la distribución normal).

El valor de JB sigue una distribución chi-cuadrado con 2 grados de libertad. Un valor alto del estadístico Jarque-Bera sugiere datos no se ajustan bien a una distribución normal. La facilidad de cálculo e interpretación del Jarque-Bera lo hace adecuado para analizar residuos en modelos de regresión y en análisis financieros, aunque su menor potencia en muestras pequeñas y la suposición de independencia entre asimetría y curtosis son limitaciones para tener en cuenta. La Figura 5 presenta el Gráfico de distribución de probabilidad y test de Jarque-Bera.

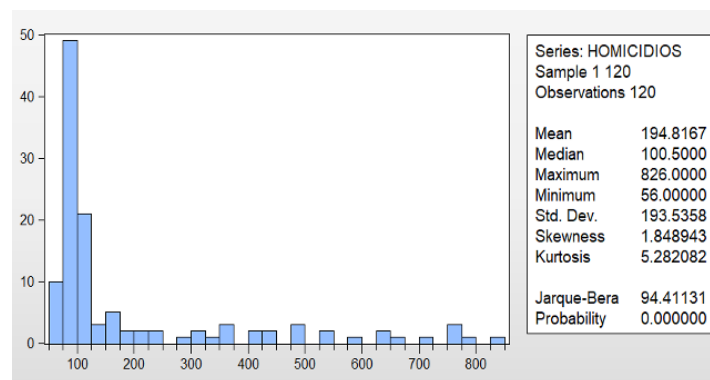


Figura 5. Gráfico de distribución de probabilidad y test de Jarque-Bera
Nota. Los datos muestran el gráfico de distribución de probabilidad de la serie histórica de homicidios intencionales y test de Jarque-Bera (2014 – 2023).

Fuente: Ministerio del Interior, (2024).

El estadístico Jarque-Bera, que muestra un valor de 94.41131 con una probabilidad asociada de 0.000000, indica un fuerte rechazo de la hipótesis nula de normalidad para la distribución de los datos de homicidios. Este hallazgo es coherente con otros estadísticos observados, especialmente la asimetría (skewness) de 1.848943 y la curtosis de 5.282082. La asimetría positiva sugiere que la distribución tiene una cola más larga hacia la derecha, lo cual se refleja en la notable diferencia entre la media (194.8167) y la mediana (100.5000).

La curtosis superior a 3 indica que la distribución es leptocúrtica, con colas más pesadas que una distribución normal. Estos valores de asimetría y curtosis explican el alto valor del Jarque-Bera, confirmando que los datos no siguen una distribución normal. Además, la considerable diferencia entre el valor mínimo (56) y el máximo (826), junto con una desviación estándar (193.5358) cercana a la media, refuerza la conclusión de que la distribución no es normal, con alta variabilidad y potencialmente valores atípicos en el extremo superior.

Prueba de Ryan-Joiner

La prueba de Ryan-Joiner es una herramienta estadística creada para evaluar si un conjunto de datos sigue una distribución normal. Su propósito principal es determinar si una muestra proviene de una población con distribución normal, comparando la distribución observada con una distribución normal teórica. El método calcula la correlación entre los datos ordenados y sus puntuaciones normales esperadas. Un valor de correlación cercano a 1 indica

un buen ajuste a la normalidad, mientras que valores más bajos sugieren desviaciones de esta distribución (Flores y Flores, 2021).

Esta prueba se basa en dos hipótesis: la hipótesis nula (H_0), que afirma que los datos siguen una distribución normal, y la hipótesis alternativa (H_1), que sostiene lo contrario. El valor p obtenido se usa para evaluar la hipótesis nula, donde un valor p menor que el nivel de significancia (usualmente 0.05) lleva a rechazar la hipótesis de normalidad. La prueba a menudo se complementa con representaciones gráficas, como gráficos de probabilidad normal, para visualizar el ajuste de los datos.

$$RJ = \frac{(\sum_{i=1}^n (Z_i - Z)^2) - \frac{1}{n} (\sum_{i=1}^n (Y_i - Y)^2)}{\frac{1}{n} (\sum_{i=1}^n (Y_i - Y)^2)} \quad (5)$$

Donde:

Z_i = Es la puntuación normal esperada para el i-ésimo dato ordenado.

Y_i = Es el i-ésimo dato ordenado de la muestra.

Z = Es la media de las puntuaciones normales esperadas.

Y = Es la media de los datos ordenados.

n = Es el tamaño de la muestra.

Entre sus ventajas, se destaca su robustez y eficacia para evaluar la normalidad, siendo particularmente útil para tamaños de muestra pequeños a moderados. La prueba de Ryan-Joiner es conceptualmente similar a otras pruebas de normalidad, como Shapiro-Wilk o Anderson-Darling. A continuación, se muestra la figura 6 de prueba de Ryan-Joiner.

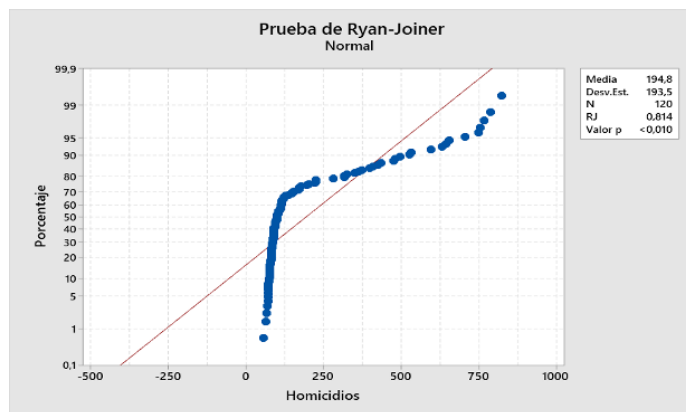


Figura 6. Prueba de Ryan-Joiner

Nota. Los datos muestran el gráfico de la prueba Ryan-Joiner de la serie histórica de homicidios intencionales (2014 - 2023).

Fuente: Ministerio del Interior, (2024).

El gráfico muestra una prueba de Ryan-Joiner para evaluar la normalidad de una distribución de datos sobre homicidios. En el eje X se representan los valores de homicidios, que van desde cerca de 0 hasta casi 1000. El eje Y muestra los porcentajes acumulados en una escala normal, desde 0.1% hasta 99.9%.

Los puntos azules representan los datos observados, estos forman una curva en forma de S que se desvía notablemente de la línea roja de referencia, especialmente en los extremos inferior y superior. Esta desviación sugiere que la distribución de los datos no es perfectamente normal.

La información estadística proporcionada en el

recuadro superior derecho indica una media de 194.8 homicidios, con una desviación estándar de 190.5. El tamaño de la muestra (N) es de 150 observaciones. El valor p de la prueba es 0.010, lo cual es menor que el nivel de significancia típico de 0.05.

Dado el valor p bajo y la desviación visible de los puntos respecto a la línea de referencia, se puede concluir que hay evidencia estadística para rechazar la hipótesis nula de normalidad. La distribución de homicidios en este conjunto de datos probablemente no sigue una distribución normal. A continuación, en la Tabla 4, se muestra un análisis comparativo de los métodos de prueba de normalidad presentados revela aspectos importantes sobre cada uno:

Tabla 4. Aspectos relevantes en las pruebas estadísticas

Prueba de normalidad	Aspectos importantes para considerar
Shapiro-Wilk	- Considerado uno de los más potentes, especialmente para muestras pequeñas a medianas ($n \leq 50$). - Muy sensible a diferentes formas de no normalidad. - En este caso, con un valor de 0.929 y $p=0.005$, rechazó la hipótesis de normalidad.
Kolmogórov-Smirnov	- Útil para muestras grandes y puede comparar con diversas distribuciones teóricas. - Menos sensible que Shapiro-Wilk para detectar no normalidad en muestras pequeñas. - Con un estadístico $D=0.313$ y $p<0.001$, también rechazó la normalidad.
Anderson-Darling	- Más sensible a las desviaciones en las colas de la distribución. - Mantiene su potencia en un amplio rango de tamaños muestrales. - El estadístico AD de 17.122 y $p<0.005$ indicó un fuerte rechazo de la normalidad.
Jarque-Bera	- Basado en asimetría y curtosis, es fácil de calcular e interpretar. - Útil en análisis financieros y de residuos en regresiones. - Con un valor de 94.41131 y $p=0.000000$, rechazó contundentemente la normalidad.
Ryan-Joiner	- Similar a Shapiro-Wilk, eficaz para muestras pequeñas a medianas. - Compara la correlación entre datos ordenados y puntuaciones normales esperadas. - Con $p=0.010$, también rechazó la hipótesis de normalidad.

Nota. Los datos muestran los aspectos más importantes a considerar sobre las diferentes pruebas de normalidad de la serie histórica de homicidios intencionales (2014 - 2023).

Fuente: Ministerio del Interior, (2024).

Las pruebas aplicadas coincidieron en rechazar la hipótesis de normalidad para los datos de homicidios. Entre ellos, los tests de Anderson-Darling y Jarque-Bera parecen mostrar los rechazos más contundentes, indicando una desviación significativa de la normalidad. Las pruebas de Shapiro-Wilk y Ryan-Joiner también rechazaron la hipótesis de normalidad, pero sus estadísticos fueron menos extremos en comparación con los anteriores.

La prueba de Kolmogórov-Smirnov proporcionó información adicional sobre la forma de la distribución, lo que complementa el análisis. En conjunto, estos métodos ofrecen una visión completa y coherente de la no normalidad de los datos. Cada prueba destaca diferentes aspectos de la distribución, lo que refuerza la conclusión de que los datos de homicidios no siguen una distribución normal. Esta concordancia en los resultados es crucial, ya que

tiene implicaciones significativas para los análisis estadísticos futuros.

La elección de la prueba adecuada depende de varios factores, incluyendo el tamaño de la muestra y la naturaleza específica de los datos. Mientras que la prueba de Shapiro-Wilk es más conveniente para muestras pequeñas, Kolmogórov-Smirnov y Anderson-Darling son preferibles para muestras grandes. Además, si las colas de la distribución son importantes, Anderson-Darling puede ser la mejor opción. En contextos como las finanzas, donde la asimetría y la curtosis son relevantes, la prueba de Jarque-Bera puede ser más pertinente.

Es crucial tener en cuenta que estas pruebas de normalidad aunque efectivas tienen limitaciones. En muestras muy grandes, pequeñas desviaciones de la normalidad pueden resultar en el rechazo de la hipótesis nula, mientras que, en muestras muy pequeñas, los tests pueden no ser lo suficientemente potentes para detectar desviaciones significativas. Por lo tanto, es recomendable complementar estas pruebas con métodos gráficos como gráficos Q-Q e histogramas y con el juicio experto basado en el conocimiento del área de estudio. A continuación, se muestra en la Tabla 5 análisis comparativo por cada prueba de normalidad empleada.

Tabla 5. Análisis comparativo por prueba de normalidad

Aspecto	Shapiro-Wilk	Kolmogórov-Smirnov	Anderson-Darling	Jarque-Bera	Ryan-Joiner
Tamaño de muestra óptimo	$n \leq 50$ (excelente)	$n > 100$ (recomendado)	$30 < n \leq 100$ (ideal)	$n > 30$ (eficiente)	$n \leq 50$ (muy bueno)
Potencia estadística	Muy alta	Moderada	Alta	Alta	Muy alta
Sensibilidad a asimetría	Buena	Baja	Muy buena	Excelente	Buena
Sensibilidad en colas	Buena	Baja	Excelente	Moderada	Buena
Facilidad de interpretación	Muy fácil	Fácil	Moderada	Muy fácil	Muy fácil
Uso en muestras grandes	No recomendado	Recomendado	Aceptable	Recomendado	No recomendado
Contexto financiero	Secundario	Poco usado	Complementario	Principal	Poco usado
Análisis de residuos	Principal	Poco usado	Complementario	Secundario	Confirmatorio
Investigación experimental	Principal	Poco usado	Complementario	Secundario	Confirmatorio
Fortalezas principales	Máxima potencia general	Comparación múltiples distribuciones	Sensibilidad en extremos	Basado en asimetría/ curtosis	Similar a Shapiro-Wilk

Aspecto	Shapiro-Wilk	Kolmogórov-Smirnov	Anderson-Darling	Jarque-Bera	Ryan-Joiner
Limitaciones principales	Excesivamente sensible en n grandes	Baja potencia en n pequeñas	Cálculo más complejo	Limitado a 2 momentos	Excesivamente sensible en n grandes
Cuando usar como principal	$n \leq 30$ o análisis de residuos	$n > 100$ o múltiples distribuciones	Sospecha problemas en colas	Datos financieros o $n > 30$	Confirmación de Shapiro-Wilk
Cuando usar como confirmatoria	$n > 30$ como segunda opción	Validación en muestras grandes	Cualquier tamaño muestral	Muestras pequeñas/medianas	$n \leq 30$ como segunda opción
Si solo esta prueba rechaza	Desviación sutil detectada	Revisar cálculos	Problema en colas	Asimetría/curtosis específica	Desviación sutil detectada
Si solo esta prueba NO rechaza	Poco probable	Posible baja potencia	Problema no en colas	Asimetría/curtosis normales	Poco probable
Recomendación en tu caso	Evidencia sólida ($p=0.005$)	Confirmación fuerte ($p<0.001$)	Validación robusta ($p<0.005$)	Rechazo contundente ($p=0.000000$)	Consenso final ($p=0.010$)
Estrategia $n \leq 30$	Prueba principal	No recomendado	Confirmatoria	Confirmatoria	Confirmatoria
Estrategia $30 < n \leq 100$	Confirmatoria	Opcional	Prueba principal	Prueba principal	Opcional
Estrategia $n > 100$	No recomendado	Prueba principal	Opcional	Prueba principal	No recomendado
Análisis gráfico recomendado	Q-Q plot + histograma	Q-Q plot + histograma	Q-Q plot + análisis colas	Histograma + coef. asimetría	Q-Q plot + histograma
Decisión final consenso	Todas rechazan = No normalidad robusta	Proceder con métodos no paramétricos	Considerar transformaciones	Evidencia concluyente	Análisis completo

Nota. Los datos muestran un análisis comparativo por cada prueba de normalidad empleada.

Fuente: Dagnino (2014)

Por último, las pruebas de normalidad son herramientas valiosas para los analistas de datos, pero deben usarse con una comprensión clara de sus fortalezas y limitaciones. La combinación de diferentes pruebas, junto con métodos visuales y consideraciones teóricas, proporciona una evaluación más robusta de la normalidad. Esta evaluación cuidadosa es esencial para asegurar la validez de los análisis estadísticos y la solidez de las conclusiones científicas derivadas.

Los resultados obtenidos en el análisis de normalidad de la serie histórica de homicidios intencionales revelan aspectos cruciales sobre la naturaleza de estos datos y sus implicaciones para futuros análisis estadísticos (Molina & Molina, 2025). En diversos estudios se ha subrayado la relevancia de las pruebas de normalidad, como Shapiro-Wilk, Kolmogórov-Smirnov, Anderson-Darling, Jarque-Bera y Ryan-Joiner, para determinar si una

población de la cual se obtuvo una muestra sigue una distribución normal (Flores & Flores, 2021; Gandica, 2021).

Estas pruebas son fundamentales para garantizar la validez de los supuestos estadísticos, ya que la suposición de normalidad es crucial para la aplicación de técnicas estadísticas paramétricas. Ignorar este supuesto puede resultar en conclusiones erróneas y una interpretación incorrecta de los resultados (Gandica, 2021).

Los resultados obtenidos en este estudio confirman la importancia de estas pruebas. Por ejemplo, Shapiro-Wilk (Flores, Muñoz, & Sánchez, 2019) mostró un estadístico de 0.92 con un valor p de 0.005, lo que indica un rechazo claro de la normalidad. Este resultado es consistente con la literatura, ya que la prueba de Shapiro-Wilk es considerada una de las más robustas para evaluar la normalidad (Molina y otros, 2023).

Asimismo, la prueba de Kolmogórov-Smirnov reveló un estadístico D de 0.31 y un valor p inferior a 0.001, lo que confirmó la no normalidad y proporcionó información sobre la posible asimetría positiva de los datos (Platas, 2021). La prueba de Anderson-Darling, centrada en las colas de la distribución, también apoyó esta conclusión, con un estadístico AD de 17.12 y un valor p menor a 0.005, sugiriendo desviaciones significativas en los extremos de la distribución. Este hallazgo es particularmente relevante dado que los eventos extremos en el contexto de los homicidios pueden tener implicaciones críticas para la comprensión y gestión del fenómeno criminal (Molina & Molina, 2025).

En cuanto a la prueba de Jarque-Bera, los resultados mostraron un estadístico de 94.41 con una probabilidad asociada cercana a cero, lo que refuerza la conclusión de no normalidad, al tiempo que proporciona información sobre la asimetría (1.84) y la curtosis (5.28) de la distribución. Estos resultados sugieren la presencia de períodos con tasas de homicidios excepcionalmente altas, y una distribución leptocúrtica, lo que indica una mayor frecuencia de valores extremos tanto altos como bajos (Lind y otros, 2012).

La prueba de Ryan-Joiner también rechazó la hipótesis de normalidad, con un valor p de 0.010. Aunque el valor fue menos extremo que el de otros tests, la representación gráfica asociada mostró una clara desviación de los puntos respecto a la línea de referencia normal, especialmente en los extremos, lo que aportó una visualización intuitiva de la no normalidad de los datos (Soto & Herrera, 2024).

La convergencia de estos resultados tiene implicaciones significativas para el análisis y la interpretación de los datos de homicidios. La no normalidad observada sugiere que los métodos estadísticos paramétricos que asumen normalidad podrían no ser adecuados para estos datos (Webster, 2000). Por lo tanto, se recomienda considerar enfoques no paramétricos o métodos robustos que no dependan de supuestos de normalidad (Pedrosa y otros, 2015). Además, la asimetría positiva y la presencia de colas pesadas indican la necesidad de prestar especial atención a los valores extremos y de emplear modelos que puedan capturar adecuadamente estas características.

Es importante destacar que la no normalidad en los datos de homicidios no es sorprendente y

puede reflejar la complejidad inherente al fenómeno criminal (Romero M., 2016). Factores como eventos sociopolíticos, cambios en las políticas de seguridad o dinámicas específicas del crimen organizado pueden contribuir a patrones no lineales y a la presencia de valores atípicos en las series temporales de homicidios.

IV. CONCLUSIONES

La normalidad sigue siendo un componente fundamental en el análisis estadístico, siendo crucial para la aplicación de muchas técnicas y métodos paramétricos. Ignorar el supuesto de normalidad puede llevar a conclusiones incorrectas y a una interpretación errónea de los resultados, lo que resalta la importancia de evaluar cuidadosamente esta condición antes de aplicar métodos que la asumen. La distribución normal, con su curva simétrica en forma de campana, es esencial tanto para la teoría estadística como para una variedad de modelos y pruebas inferenciales en diversas áreas.

Los tests de normalidad continúan siendo herramientas clave en el análisis estadístico, pero su aplicación e interpretación requieren un enfoque detallado. La elección adecuada de la prueba, la consideración del tamaño de la muestra y la comprensión de las implicaciones de la no normalidad son aspectos críticos que deben evaluarse en cualquier análisis estadístico riguroso. La suposición de normalidad sigue siendo crucial, pero es esencial que los investigadores reconozcan sus implicaciones, validen su aplicabilidad mediante pruebas adecuadas y consideren alternativas cuando sea necesario.

Para evaluar la normalidad, se han desarrollado métodos que combinan pruebas formales y enfoques gráficos. Estos tests miden la discrepancia entre la distribución empírica de los datos y la distribución normal teórica. La combinación de métodos gráficos, como los gráficos Q-Q e histogramas, con pruebas estadísticas formales permite una evaluación más completa de la normalidad. Este enfoque multifacético brinda a los investigadores una comprensión más profunda de la estructura de los datos y facilita la selección de los métodos de análisis adecuados.

A pesar de su utilidad, las pruebas de normalidad tienen limitaciones. En muestras grandes, incluso pequeñas desviaciones de la normalidad pueden

resultar en el rechazo de la hipótesis nula, mientras que, en muestras pequeñas, estas pruebas pueden no detectar desviaciones significativas. Esto resalta la necesidad de considerar el contexto y el tamaño de la muestra al interpretar los resultados, y no depender únicamente de los valores p .

Cuando los datos no siguen una distribución normal, se pueden aplicar transformaciones o métodos no paramétricos. Aunque estas opciones pueden ser útiles, las transformaciones deben aplicarse con cautela para evitar complicaciones en la interpretación de los resultados. Esta tensión entre cumplir con los supuestos estadísticos y mantener la interpretabilidad de los resultados refleja los desafíos en la práctica investigativa.

V. REFERENCIAS

- Álvarez, E., y Barreda, L. (2020). La estadística descriptiva en la formación investigativa del Instructor de Arte. *Revista Conrado*, 16(73), 100-107. <http://scielo.sld.cu/pdf/rc/v16n73/1990-8644-rc-16-73-100.pdf>
- Arguello, V., Cargua, S., Arguello, A., y Argurello, A. (2023). Distribución de Poisson en el análisis de defunciones en menores de un año. *Polo del conocimiento*, 8(5), 292-303. <https://doi.org/10.23857/pc.v8i5.5564>
- Bernal, C. (2010). *Metodología de la investigación*. Pearson. <http://librodigital.sangregorio.edu.ec/librosusgp/Boo61.pdf>
- Cruz, M., Álvarez, S. E., y Pérez, F. (2014). Sobre la Enseñanza de la Distribución de Poisson en Carreras de Ingeniería. *Boletim de Educação Matemática*, 28(50), 1117-1134. <http://www.redalyc.org/articulo.oa?id=291232906007>
- Dagnino, J. (2014). La distribución normal. *Rev Chil Anest*, 43, 116-121. <https://revistachilenadeanestesia.cl/PII/revchilanestv43n02.o8.pdf>
- Delgado, R. (2022). El teorema central del límite: una visión ilustrativa. *Revista Varianza*, 20, 35-42. <https://doi.org/https://doi.org/10.53287/uufr-9587pp15b>
- Flores, C., y Flores, K. (2021). Pruebas para comprobar la normalidad de datos en procesos productivos: Anderson Darling, Ryan-Joiner, Shapiro-Wilk Y Kolmogórov-Smirnov. *Societas. Revista de Ciencias Sociales y Humanísticas*, 23(2), 83-97. <http://portal.amelica.org/ameli/jatsRepo/341/3412237018/3412237018.pdf>
- Flores, P., Muñoz, L., y Sánchez, T. (2019). Estudio de potencia de pruebas de normalidad usando distribuciones desconocidas con distintos niveles de no normalidad. *Perfiles*, 1(21), 4-11. <https://doi.org/https://doi.org/10.47187/perf.v1i21.42>
- Gandica, E. (2021). Potencia y Robustez en Pruebas de Normalidad con Simulación Montecarlo. *Revista Científica*, 5(18), 108-119. <https://doi.org/https://doi.org/10.29394/Scientific.issn.2542-2987.2020.5.18.5.108-119>
- Hernández, R. (2010). *Metodología de la investigación*. México: Macgraw Hill. <https://www.smuje-rescoahuila.gob.mx/wp-content/uploads/2020/05/Sampieri.Met.Inv.pdf>
- Indira, L., y Llinás, H. (2014). Distribuciones Poisson y Gamma: Una Discreta y Continua Relación. *Prospect*, 12(1), 99-107. <https://doi.org/10.15665/rp.v12i1.156>
- Lemus, N., y Huincahue, J. (2019). Distribución Normal: Análisis histórico-epistemológico e implicaciones didácticas. *Revista Académica UC Maule*, 56, 1-24. <http://portal.amelica.org/ameli/journal/600/6002792003/>
- Lemus, N., y Huincahue, J. (2019). Distribución normal: análisis histórico-epistemológico e implicaciones didácticas. *UC Maule*(56), 29-57. <https://doi.org/http://doi.org/10.29035/ucmaule.56.29>
- Lind, Marchal, y Wathen. (2012). *Estadística Aplicada a los Negocios y la Economía*. Mcgraw Hill. https://www.academia.edu/17986600/Mcgraw_Hill_Estadística_Aplicada_a_los_Negocios_y_la_Economía
- Lugo, J., y Pino, L. (2022). Razonamiento inferencial de docentes de matemáticas de enseñanza media sobre

- el estadístico t-Student. *Uniciencia*, 36(1), 1-29. <https://doi.org/https://doi.org/10.15359/ru.36-1.25>
- Mares, A., y Domínguez, J. (2022). Evaluación estadística de índices de desempeño para el proceso de división de rollos de EVA. *Ingeniería, investigación y tecnología*, XXIII(2), 1-15. <https://doi.org/https://doi.org/10.22201/fi.25940732e.2022.23.2.013>
- Marques, J., Andrade, T., y Wagner dos, S. (2023). Relación entre la titulación del profesorado y la calidad de las instituciones de la Red Federal de Educación Profesional, Científica y Tecnológica brasileña. *Revista de estudios y experiencias en educación*, 22(48), 207-221. <https://doi.org/https://doi.org/10.21703/0718-5162.v22.n48.2023.012>
- Martínez , P., y Carleos, C. (2011). sobre el estadístico de Cramér – Von Mises. *Revista de Matematica: Teoría y Aplicaciones*, 19(1), 80-101. <https://www.researchgate.net/publication/262701897>
- Ministerio del Interior. (16 de Agosto de 2024). *Datos abiertos*. <https://datosabiertos.gob.ec/dataset/?organization=ministerio-del-interior>
- Molina, M. (2022). Análisis de normalidad. Una imagen vale más que mil palabras. *Rev elect anestesiar*, 14(4). <https://revistaanestesiar.org/index.php/rear/article/view/1093/1402>
- Molina, P., Molina, D., Pico, F., y Guevara, P. (2023). La valoración de empresas y las decisiones de inversión en el sector comercial farmacéutico. *Res Non Verba*, 13(1), 15-35. <https://doi.org/https://doi.org/10.21855/resnonverba.v13i1.689>
- Molina, P., Molina, D., Pico, F., y Velasco, K. (2023). Modelo de fragilidad empresarial Springate en empresas comerciales de productos veterinarios en Ecuador. (37), 7-14. <https://doi.org/https://doi.org/10.31243/bcoyu.37.2023.2081>
- Molina, P., Morán, E., Molina, D., y Caiza, E. (2023). Ineficiencia del mercado de valores de Ecuador a través del modelo de valoración de activos de capital (CAPM). *Revista de Investigación Sigma*, 10(2), 82-105. <https://doi.org/https://doi.org/10.24133/ris.v10i02.3127>
- Molina, P., y Molina, D. (2025). Crímenes en expansión: homicidios intencionales y los desafíos para los derechos humanos y la seguridad en Ecuador. *Tsafiqui - Revista Científica En Ciencias Sociales*, 15(1), 103-120. <https://doi.org/https://doi.org/10.29019/tsafiqui.v15i1.1463>
- Moraguez, A., Pilar, E., y Morales, L. (2017). La prueba de hipótesis Kolmogorov-Smirnov para dos muestras grandes con una cola. *Luz*, 16(3), 77-89. <https://www.redalyc.org/pdf/5891/589166504009.pdf>
- Muñoz , P., Muñoz , L., y Sánchez, T. (2021). Estudio de potencia de pruebas de normalidad usando distribuciones desconocidas con distintos niveles de no normalidad. *Perfiles*, 1(21), 4-11. <https://doi.org/https://doi.org/10.47187/perf.v1i21.42>
- Padilla, F. (2022). El valor p: cómo analizarlo para separarse del positivismo extremo e inductivismo ingenuo? *Investigación en educación médica*, 11(44), 105-114. <https://www.redalyc.org/journal/3497/349775300011/349775300011.pdf>
- Pedrosa, I., Baster, J., Robles, A., Basteiro, J., y García, E. (2015). Pruebas de bondad de ajuste en distribuciones simétricas, ¿qué estadístico utilizar?*. *Universitas Psychologica*, 14(1), 245-254. <https://doi.org/http://dx.doi.org/10.11144/Javeriana.upsy13-5.pbad>
- Pek, Wong, y Wong. (2017). Confidence Intervals for the Mean of Non-Normal Distribution: Transform or Not to Transform. *Open Journal of Statistics*(7), 405-421. https://www.scirp.org/pdf/OJS_2017060814493596.pdf
- Platas, V. (2021). *Contrastes de normalidad*. https://minerva.usc.es/xmlui/bitstream/handle/10347/28978/Platas_Garc%C3%ADa_V%C3%ADctor.pdf?sequence=1&isAllowed=y
- Romero, H., Real, J., Ordoñez, J., Gavino, G., y Saldarriaga, G. (2020). *Metodologías de la investigación*. https://acvenisproh.com/libros/index.php/Libros_categoria_Academico/article/view/22/29
- Romero, M. (2016). Pruebas de bondad de ajuste a una distribución normal. *Revista Enfer-*

- mería del Trabajo*, 6(3), 105-114. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-50572018000300092
- Schaum. (1991). *Estadística de Schaum* (2da ed.). México: McGraw Hill. https://materiales.untrefvirtual.edu.ar/documentos_extras/0714_estadisticas/estadistica_schaum_1.pdf
- Soto, J., y Herrera, A. (2024). Conceptos básicos de la Distribución Normal. *Rev. Red Inv. Sal. Trab.*, 7(12), 73-82. [https://rist.zaragoza.unam.mx > article > download](https://rist.zaragoza.unam.mx/article/download)
- Webster, A. (2000). *Estadísticas aplicadas a los negocios y a la economía* (3ra ed.). México: McGraw Hill. https://dennismontes2.wordpress.com/wp-content/uploads/2014/11/estadistica_negocios.pdf