Understanding and Implementing the Jarque-Bera Test in Excel

Authored by

Mohammed looti

November 8, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *Understanding and Implementing the Jarque-Bera Test in Excel*. PSYCHOLOGICAL STATISTICS. Retrieved from https://statistics.arabpsychology.com/?p=13250

The Critical Role of the Jarque-Bera Test in Data Analysis

The Jarque-Bera test (JB test) stands as a highly respected and essential goodness-of-fit test within modern statistics and econometrics. Its primary purpose is to determine whether a given sample dataset follows a theoretical normal distribution. This determination is crucial because many advanced statistical procedures, particularly parametric tests and regression models, rely fundamentally on the assumption of data normality. If this core assumption is violated, the resulting statistical inferences can be flawed, leading to unreliable model predictions and incorrect conclusions.

The JB test specifically evaluates the data based on two key characteristics: its level of skewness and its level of kurtosis. For a dataset to be perfectly **normally distributed**, the distribution must be perfectly symmetrical, meaning its skewness (a measure of asymmetry) must be zero. Furthermore, its excess kurtosis (a measure of tail heaviness relative to the normal curve) must also be zero. The JB test elegantly combines deviations from these two theoretical values into a single, standardized test statistic.

Interpreting the test statistic is straightforward: a value close to zero indicates that the sample data closely mirrors the ideal properties of a normal distribution. Conversely, a large test statistic signals a significant departure from normality, often due to heavy tails or pronounced asymmetry. Understanding the underlying hypothesis framework is key: the null hypothesis (H_0) asserts that the data is normally distributed, while the alternative hypothesis (H_a) asserts that it is not. Because the test statistic is non-negative, larger positive values increase the evidence against H_0, making the Jarque-Bera test an indispensable preliminary tool for comprehensive data diagnostics.

Deconstructing the Jarque-Bera Test Statistic Formula

The calculation of the **Jarque-Bera (JB) test statistic** is rooted in the sample moments of the distribution. Specifically, it utilizes the third standardized moment (related to skewness) and the fourth standardized moment (related to kurtosis). The resulting formula integrates these two measures, weighted by the sample size, to generate the final statistic. This statistic is then evaluated against a known statistical distribution to calculate the probability of observing the current data characteristics if normality were truly present.

The exact mathematical definition of the test statistic, typically denoted as ***JB***, is expressed as follows:

JB =(n/6) * (S2 + (C2/4))

To successfully calculate this statistic in Excel, analysts must first accurately derive three critical components directly from their raw dataset:

n: This represents the total **number of observations** or data points contained within the sample dataset.
S: This is the **sample skewness**, calculated from the third standardized moment. It measures both the extent and the direction of asymmetry in the data's probability distribution.
C: This is the **sample excess kurtosis**. Derived from the fourth standardized moment, it quantifies the "peakedness" and "tailedness" of the distribution relative to the standard normal distribution.

Crucially, under the assumption that the data is genuinely normally distributed (i.e., the null hypothesis is true), the JB statistic follows, asymptotically, a Chi-squared distribution with exactly two degrees of freedom, written as $JB sim text{X}^2(2)$. These two degrees of freedom correspond precisely to the two constraints being tested simultaneously: skewness equals zero and excess kurtosis equals zero. This established distributional relationship is what permits the calculation of critical values and the essential p-value needed for formal hypothesis testing.

Setting Up Excel for Jarque-Bera Calculation

Unlike specialized statistical software packages that offer a single function for the Jarque-Bera test, Microsoft Excel requires the user to meticulously combine several built-in functions to arrive at the final statistic. This necessitates careful organization and understanding of Excel's statistical function definitions. The prerequisite steps involve structuring the data correctly and identifying the specific functions required for the components n, S, and C.

Before any calculation begins, ensure your dataset is neatly organized, ideally in a single column or row. For the purposes of this guide, we will assume the data is arranged vertically in one column. The three fundamental Excel functions needed to calculate the component parts are: **COUNT** (to determine n), **SKEW** (to calculate the sample skewness S), and **KURT** (to calculate the sample excess kurtosis C). It is vital to recognize that the Excel **KURT** function specifically calculates the **excess kurtosis**, which conveniently aligns perfectly with the requirement of the Jarque-Bera formula (as the expected value for a normal distribution is zero).

We strongly recommend creating clearly labeled intermediate cells (e.g., "Sample Size," "Skewness Value," "Kurtosis Value," "JB Statistic") to ensure complete transparency throughout the process and minimize the risk of cell reference errors. Once the individual components are calculated using these functions, the final JB statistic can be determined by inputting the full formula, $JB = (n/6) times (S^2 + (C^2/4))$, into the dedicated output cell. This rigorous, step-by-step methodology guarantees accuracy when performing this complex statistical procedure within a spreadsheet environment.

Step-by-Step Implementation: Data Preparation and Sample Size

The initial stage of performing the Jarque-Bera test involves the accurate input and preparation of the raw data within the Excel spreadsheet. This section outlines the precise steps required to establish the foundation for the subsequent calculations of skewness and kurtosis.

Step 1: Input the Data.

Begin by entering all of your raw data observations into a single, contiguous column in Excel. For the purposes of this demonstration, we will assume the data occupies the range of cells A1 through A30. This initial, systematic organization is critical as all subsequent formulas will reference this specific range.

	A	B	C	D	E	F
1	**Data**					
2	4					
3	5					
4	5					
5	6					
6	9					
7	12					
8	13					
9	14					
10	14					
11	19					
12	22					
13	24					
14	25					
15						
16						
17						
18						

Once the data is correctly positioned, the next essential calculation is determining the sample size, **n**. This quantity serves as the weighting factor in the Jarque-Bera equation. To calculate n, utilize the Excel **COUNT** function, which accurately tallies the number of numerical entries within the defined range. In our running example, where data spans A1:A30, the formula executed in a separate, labeled cell (e.g., B2) would be `=COUNT(A1:A30)`. We recommend dedicating separate, clearly labeled cells for the sample size, skewness, and kurtosis before moving to the next phase, which significantly simplifies the final calculation of the JB statistic.

Calculating the Distributional Shape Measures (S and C)

With the sample size n successfully calculated, attention shifts to quantifying the distribution's shape using the measures of asymmetry and tailedness. These calculations rely entirely on Excel's specialized statistical functions, which efficiently compute the complex summations required for the standardized moments.

To derive the **sample skewness (S)**, analysts must employ the Excel function **SKEW**. This function should reference the entire data range (A1:A30). The output value S mathematically quantifies the degree and direction of the dataset's asymmetry. A positive value for S indicates a distribution with a longer tail extending to the right (positive skew), whereas a negative S signals a longer tail extending to the left (negative skew). Recall that for a truly normal distribution, the expected value of S is zero.

Following skewness, the **sample excess kurtosis (C)** must be calculated using the **KURT** function. This measure assesses how sharply peaked the distribution is and how thick its tails are in comparison to the benchmark normal distribution. If C is positive, the distribution is characterized as leptokurtic (possessing heavier, fatter tails than normal), and if C is negative, it is platykurtic (possessing lighter tails). The ideal value for C in a normal distribution is 0. These calculated values, S and C, are the components that must be squared (S^2 and C^2) and integrated into the final JB formula.

Step 2: Assemble the Jarque-Bera Test Statistic.

The final computation of the test statistic involves combining n, S, and C into the comprehensive formula: $JB = (n/6) times (S^2 + (C^2/4))$. It is best practice to reference the cell locations where n, S, and C were individually calculated, rather than hardcoding the values. The visual aid below demonstrates how the intermediate components are derived using the appropriate Excel functions (Column F) before being synthesized into the final JB statistic formula in the designated cell.

	A	B	C	D	E	F	G
1	**Data**						
2	4		observations	n	13	=COUNT(A2:A14)	
3	5		sample skewness	S	0.3367	=SKEW(A2:A14)	
4	5		sample kurtosis	C	-1.2171	=KURT(A2:A14)	
5	6						
6	9		JB test statistic	JB	1.0481	=(E2/6) * (E3^2+(E4^2)/4)	
7	12						
8	13						
9	14						
10	14						
11	19						
12	22						
13	24						
14	25						
15							
16							
17							
18							

The resulting JB statistic is a single, quantitative summary of the data's overall deviation from the theoretical normal curve. A higher magnitude of this value provides stronger statistical evidence suggesting that the data is non-normal, leading directly to the final, necessary stage of the test: calculating and interpreting the associated p-value.

Interpreting Normality via the P-Value and Decision Rule

While the calculated JB statistic provides a measure of deviation, it cannot alone lead to a formal conclusion. The true power of the test lies in its associated **p-value**. The p-value represents the probability of observing a test statistic as extreme as, or more extreme than, the one calculated, assuming that the null hypothesis--that the data is normally distributed--is actually true.

Step 3: Calculate the p-value of the test.

As established, the JB statistic asymptotically follows a Chi-squared distribution with two degrees of freedom ($X^2(2)$). Therefore, to find the p-value in Excel, we must utilize the **CHISQ.DIST.RT** function (Chi-squared distribution right tail). We employ the right-tail function because large, positive JB values indicate significant non-normality, placing these critical results in the right tail of the distribution.

The syntax for the p-value calculation in Excel is standardized as: `=CHISQ.DIST.RT(JB_Statistic_Cell, 2)`. The first argument must be the cell containing the calculated JB value, and the second argument is fixed at 2, representing the degrees of freedom for the standard Jarque-Bera test.

	A	B	C	D	E	F	G
1	**Data**						
2	4		observations	n	13	=COUNT(A2:A14)	
3	5		sample skewness	S	0.3367	=SKEW(A2:A14)	
4	5		sample kurtosis	C	-1.2171	=KURT(A2:A14)	
5	6						
6	9		JB test statistic	JB	1.0481	=(E2/6) * (E3^2+(E4^2)/4)	
7	12						
8	13		p-value	p	0.5921	=CHISQ.DIST.RT(E6, 2)	
9	14						
10	14						
11	19						
12	22						
13	24						
14	25						
15							
16							

In the example shown above, the resultant p-value is **0.5921**. This figure must then be compared against a predefined significance level, denoted as alpha ($alpha$), which is conventionally set at 0.05 (5%). The decision rule for the hypothesis test is simple and critical:

If the p-value is **less than $alpha$ (0.05)**, we **reject the null hypothesis**, concluding the data is not normally distributed.
If the p-value is **greater than or equal to $alpha$ (0.05)**, we **fail to reject the null hypothesis**.

Since our calculated p-value (0.5921) is substantially larger than the standard 0.05 significance level, we **fail to reject the null hypothesis of normality**. This outcome signifies that the statistical evidence provided by the Jarque-Bera test is insufficient to conclude that the dataset deviates significantly from a normal distribution. Consequently, an analyst can confidently proceed with parametric statistical modeling, assuming other necessary requirements (such as homogeneity of variance or linearity) are also met.

Summary of Key Jarque-Bera Concepts

The Jarque-Bera test provides a powerful quantitative assessment of data normality by focusing on two fundamental shape parameters: skewness and kurtosis. Mastering the calculation in Excel requires a clear understanding of the formula components and the correct application of the SKEW and KURT functions, coupled with the final interpretation using the Chi-squared distribution. This systematic process ensures that critical data assumptions are verified before complex modeling begins, thus safeguarding the validity of all subsequent statistical findings.