Practical Work 03

Normality Tests

Graphical Methods and Statistical Testing

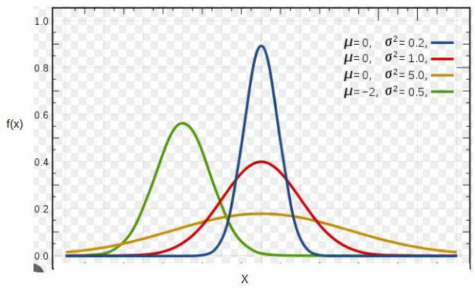
1. Objectives

1. Present graphical methods used to study the normality of a distribution.
2. Present technical procedures for conducting formal normality tests.

2. Introduction to Normal Distribution

The normal distribution (or Gaussian distribution) plays a vital role in statistics. It is a continuous probability distribution defined by two parameters: the mean (μ) and the variance (σ^2).

- **Symmetry:** The distribution is perfectly symmetrical around its mean.
- **Importance:** Many statistical tests (parametric tests) require the data to follow a normal distribution as a prerequisite.
- **Visualizing Variance:** As the standard deviation (σ) increases, the curve becomes flatter (lower peak). Conversely, a smaller standard deviation results in a narrower, taller peak.



3. Procedure for Normality Testing

To determine if a sample follows a normal distribution, we follow these standardized steps:

Step 1: Formulate Hypotheses

- H_0: The distribution is Gaussian (Normal).
- H_1: The distribution is not Gaussian.

Step 2: Choose the Statistical Test

- **Kolmogorov-Smirnov Test:** Used when the sample size is large ($n \geq 30$).
- **Shapiro-Wilk Test:** Used when the sample size is small ($n < 30$).

Step 3: Decision Rule We compare the calculated significance level (sig or p-value) with the alpha level (α):

- If $Sig < \alpha$: Reject H_0. The distribution is **not Normal**.
- If $Sig > \alpha$: Accept H_0. The distribution is **Normal**.

4. SPSS Procedure: The "Explore" Command

To perform normality tests and generate graphs in SPSS:

1. Go to **Analyze** > **Descriptive Statistics** > **Explore**.
2. Move the variable to the **Dependent List**.
3. Click **Plots**, check **Histogram**, and check **Normality plots with tests**.
4. Click **Statistics** and ensure **Descriptives** and **Percentiles** are selected.

5. Exercise 1: Bacterial Count Analysis

A study recorded the number of bacteria found in 30 different biological media.
Data: 31, 32, 32, 33, 33, 33, 34, 34, 34, 34, 34, 35, 35, 35, 35, 35, 35, 36, 36, 36, 36, 36, 37, 37, 37, 37, 38, 38, 39, 39.

5.1 Solutions and Interpretation

1. **Objective:** To test if the bacterial count follows a normal distribution.
2. **Variable:** Number of bacteria (Quantitative continuous).
3. **Hypotheses:** H_0: Distribution is Gaussian; H_1: Distribution is not Gaussian.

5.1.1 Steps To Follow

Figure 1: Step 01

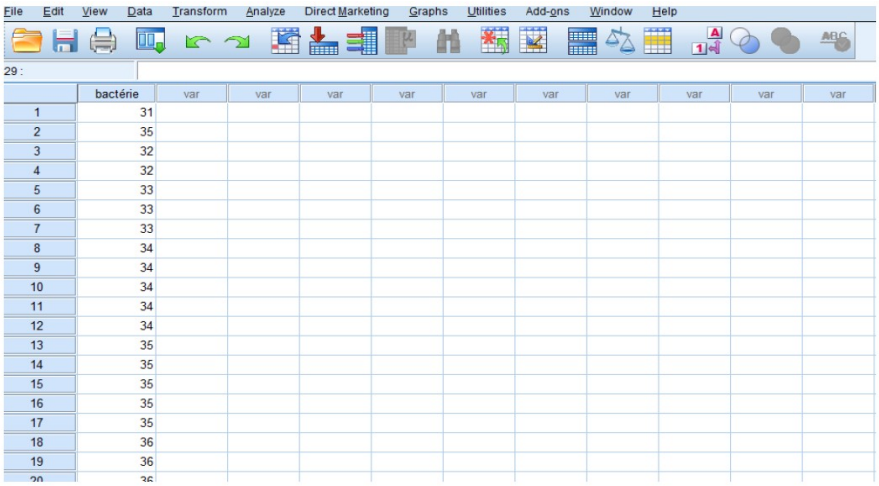


Figure 2: Step 02

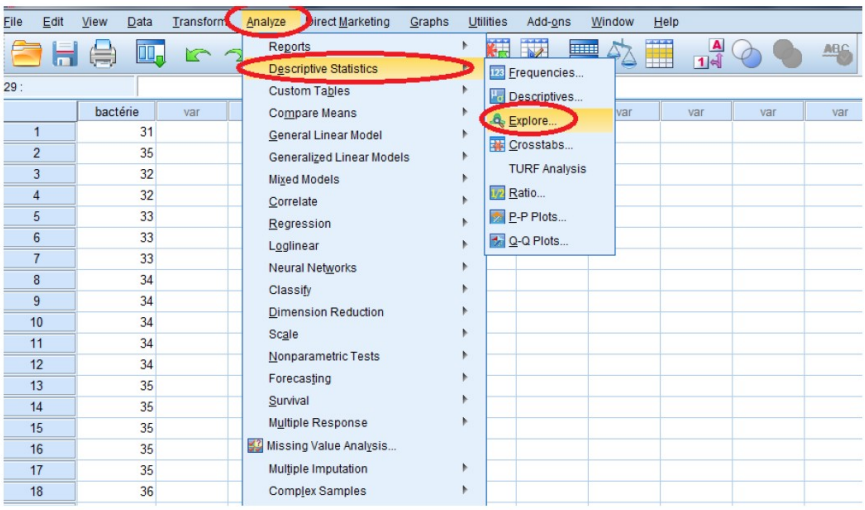


Figure 3: Step 04

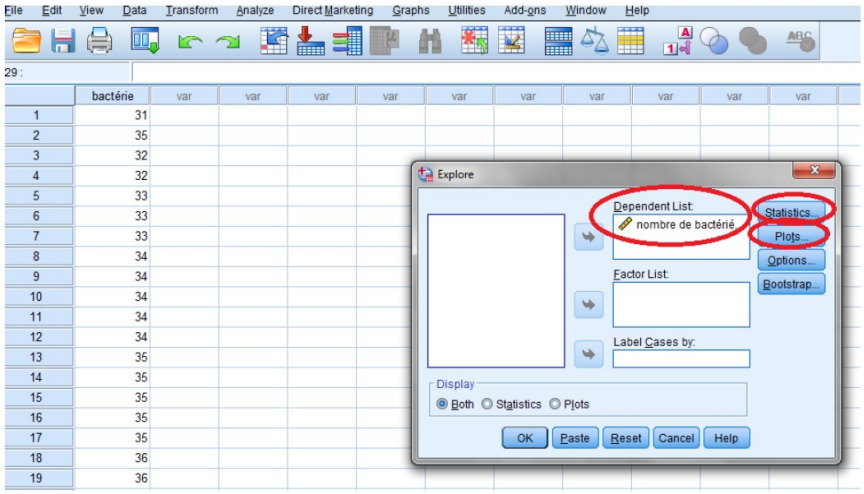


Figure 4: Step 05

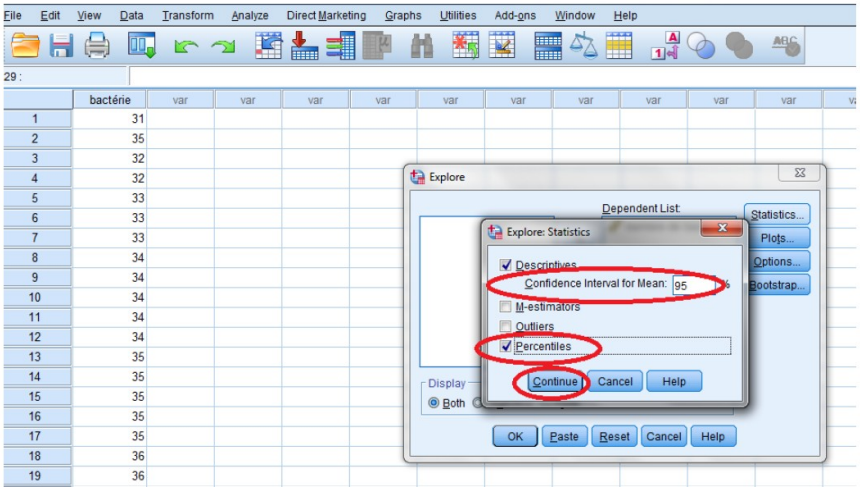


Figure 5: Step 06

5.1.2 Results Analysis

- **Symmetry:** The Mean (35.20) is nearly equal to the Median (35.00). The Skewness is -0.003 (very close to 0), indicating a symmetrical distribution.

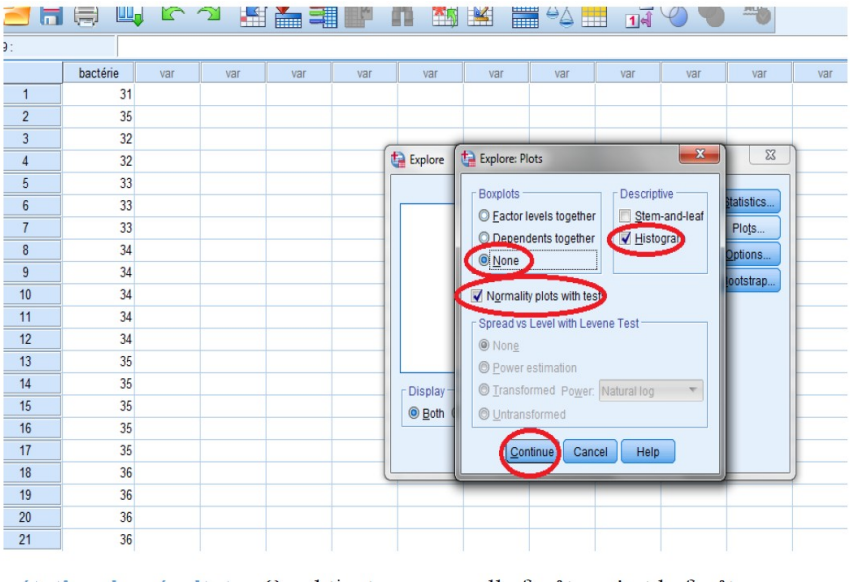


Figure 6: Step 07

Descriptives

			Statistic	Std. Error
nombre de bactéries	Mean		35.2000	.37263
	95% Confidence Interval for Mean	Lower Bound	34.4379	
		Upper Bound	35.9621	
	5% Trimmed Mean		35.2037	
	Median		35.0000	
	Variance		4.166	
	Std. Deviation		2.04096	
	Minimum		31.00	
	Maximum		39.00	
	Range		8.00	
	Interquartile Range		3.00	
	Skewness		-.003	.427
	Kurtosis		-.439	.833

Figure 7: Descriptive Statistics Table

- **Normality Test:** Since $n = 30$, we look at the **Kolmogorov-Smirnov** results. The *Sig* value is **0.200**.
- **Conclusion:** Since $0.200 > 0.05$, we accept H_0. The distribution is Gaussian at a 95% confidence level.

5.1.3 Graphical Methods

- **Histogram:** The bars follow the bell-shaped curve closely.
- **Q-Q Plot:** The data points lie directly on or very close to the diagonal line, confirming normality.

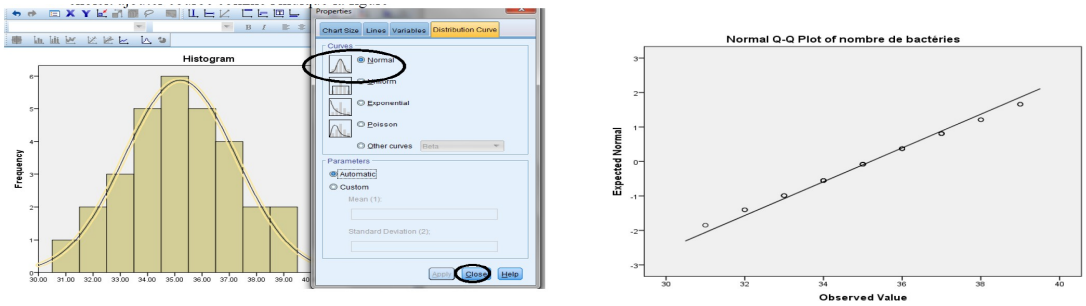


Figure 8: Histogram with Normal Curve and Q-Q Plot