

Introduction to Hypothesis Testing

Prepared by Dr. Hesham Mousli

MHA 613
Statistics in Healthcare

Lecture 5

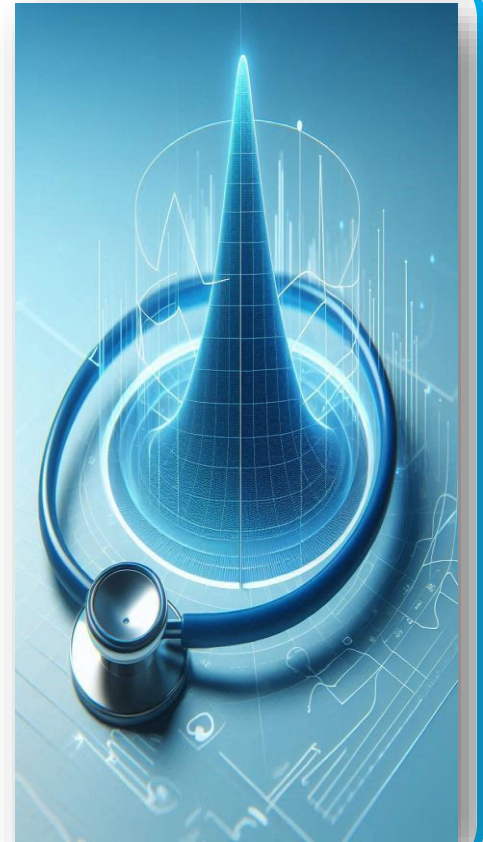


INTENDED OBJECTIVES

1. Define hypothesis testing and its role in Healthcare-related decision-making & research.
2. Understand the concept of normal distribution and its importance.
3. Learn how hypothesis testing and p-values are interconnected.
4. Differentiate between Type I and Type II errors.
5. Learn graphical & numerical methods to test normality.

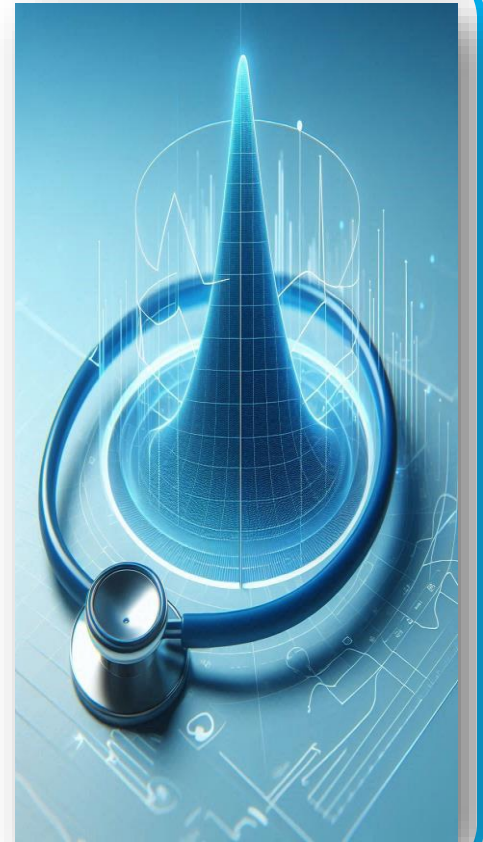
What is the Hypothesis Testing?

- In healthcare-related research, hypothesis testing is a fundamental statistical method used to make decisions based on data.



What is the Hypothesis Testing?

- It allows researchers to evaluate whether the observed data supports a specific claim (**hypothesis**) about a **population parameter**.



What is the Hypothesis Testing?

- **The process begins by formulating two competing hypotheses:**
 - ❑ **The null hypothesis (H_0):** assumes no effect or no difference.
 - ❑ **The alternative hypothesis (H_1):** suggests the presence of an effect or difference.

Why Hypothesis Testing Matters?

- Through **hypothesis testing**, we **assess evidence from a sample** to determine **whether** it is **strong enough** to **reject the null hypothesis** in favor of the **alternative**.

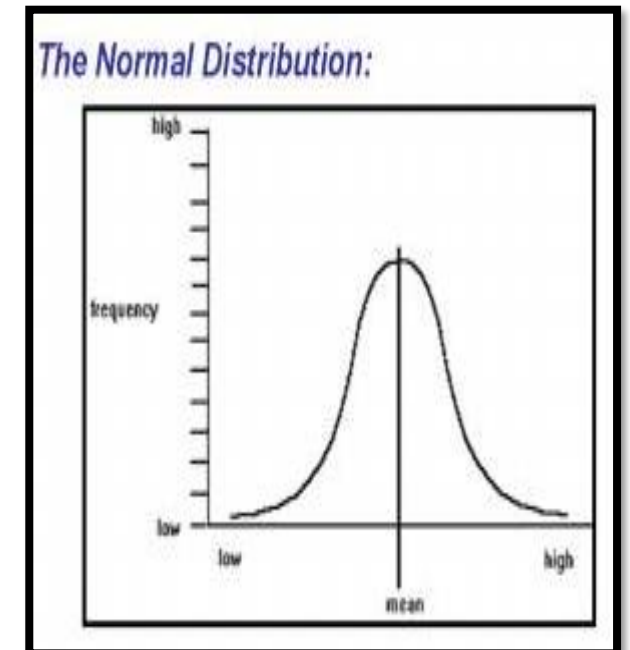
Why Hypothesis Testing Matters?

- This process helps researchers and clinicians make informed decisions in clinical and healthcare settings.



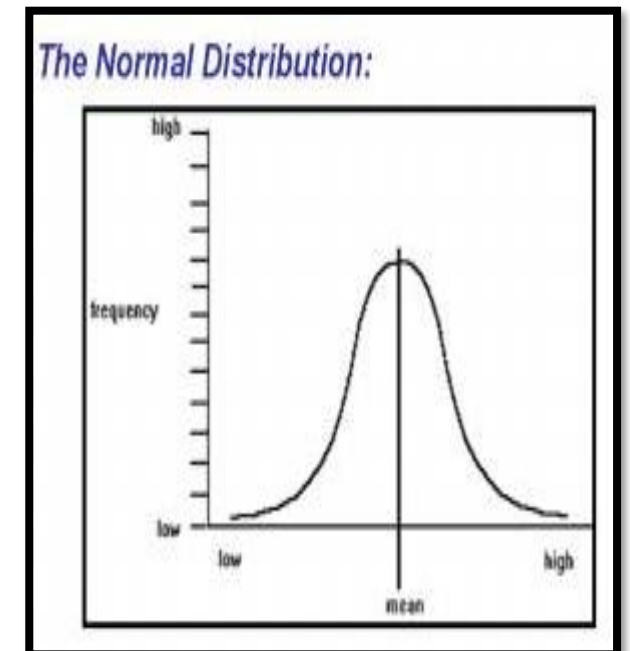
Why Hypothesis Testing Matters?

- A **Normal Distribution** is a symmetrical probability distribution with most values clustered around the mean and fewer values tapering off equally on both sides.



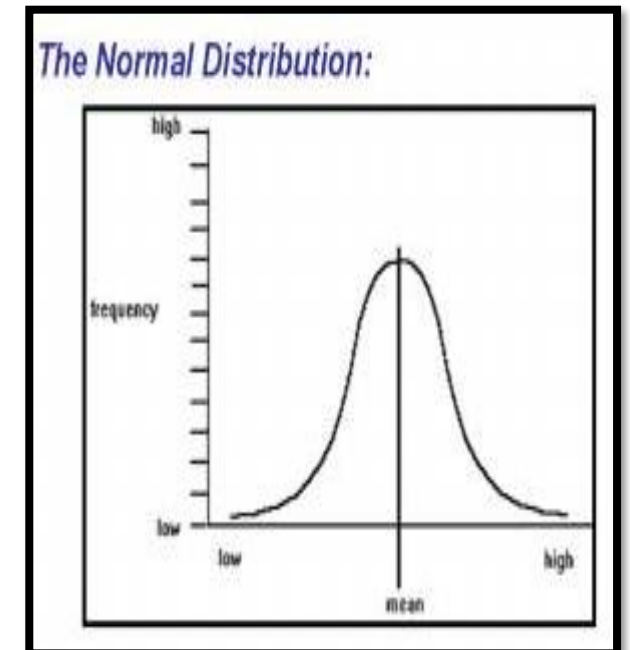
Why Hypothesis Testing Matters?

- It has the shape of a **bell**.
- Can **entirely** described by its **mean** and **standard deviation**.



Why Hypothesis Testing Matters?

- It's found **naturally** in **social and human sciences**.
- **Example:** Population systolic blood pressure.

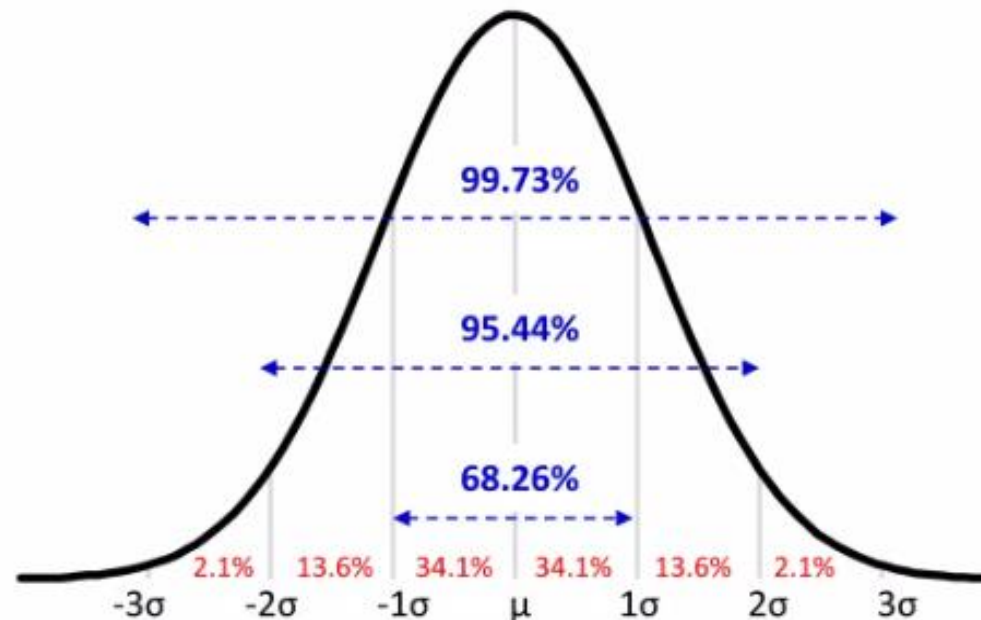


Empirical Rule:

□ For any normally distributed data:

- **68%** of the data fall within **1** standard deviation of the mean.
- **95%** of the data fall within **2** standard deviations of the mean.
- **99.7%** of the data fall within **3** standard deviations of the mean.

95% Confidence Interval for
the population (95%CI)



Hypothesis Testing

➤ **Hypothesis testing refers to:**

1. **Making an assumption, called a *hypothesis*, about a *population parameter*.**
2. **Collecting the sample data.**
3. **Calculating a sample statistic, P-value, and *Confidence Interval (CI)*.**

Hypothesis Testing

➤ **Hypothesis testing refers to:**

- 4. Using sampling to evaluate the hypothesis** (how likely is it that our hypothesized parameter or conclusion regarding the population is correct).

HYPOTHESIS TESTING

Null hypothesis, H_0

- State the hypothesized value of the parameter before sampling.
- The assumption we wish to test (or the assumption we are trying to reject)
- E.g population mean $\mu = 20$
- There is no difference between coke and diet coke

Alternative hypothesis, H_A

All possible alternatives other than the null hypothesis.

E.g $\mu \neq 20$

$\mu > 20$

$\mu < 20$

There is a difference between coke and diet coke

Hypothesis Testing

➤ Resulting P-value of Hypothesis Testing:

- P-value is defined as the probability of obtaining a result equal to or more extreme than what was actually observed.
- The smaller the P-value, the larger the significance of my alternative hypothesis.

Hypothesis Testing

➤ Resulting P-value of Hypothesis Testing:

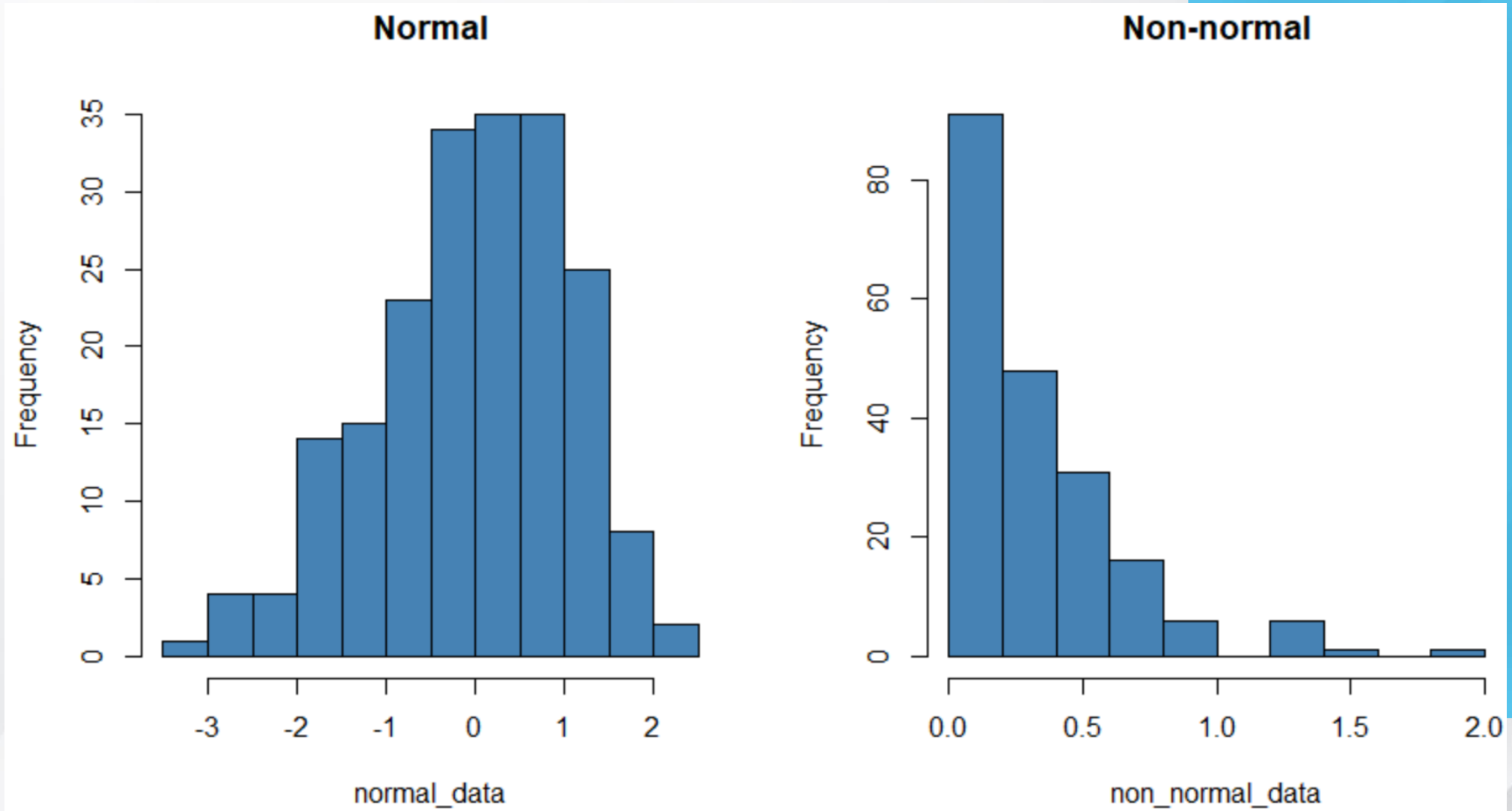
- For any hypothesis testing:
 - A. Arbitrary cut-off 0.05 (5% chance of false positive conclusion) is commonly set.
 - B. If $P\text{-value} < 0.05$: Reject the Null Hypothesis, and accept the Alternative hypothesis.
 - C. If $P\text{-value} > 0.05$: Accept the Null Hypothesis, and reject the Alternative hypothesis.

Type I vs. Type II error

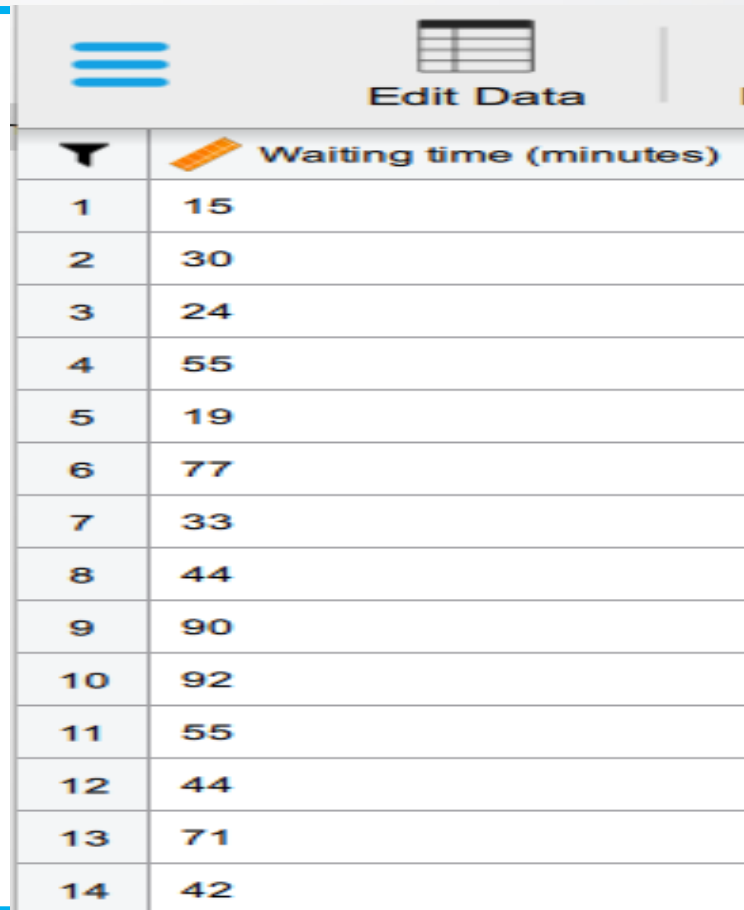
- A hospital management researcher is conducting a study to assess the effectiveness of a new staff training program designed to reduce medication errors in the hospital.

P-value	Manager concluded	The truth	Type of error
< 0.05	The new training program reduces medication errors compared to the standard training.	The new training program does not reduce medication errors compared to the standard training.	Type I error
> 0.05	The new training program does not reduce medication errors compared to the standard training.	The new training program reduces medication errors compared to the standard training.	Type II error
< 0.05	The new training program reduces medication errors compared to the standard training.	The new training program reduces medication errors compared to the standard training.	
> 0.05	The new training program does not reduce medication errors compared to the standard training.	The new training program does not reduce medication errors compared to the standard training.	

Testing Normality for Continuous Data (Graphically)



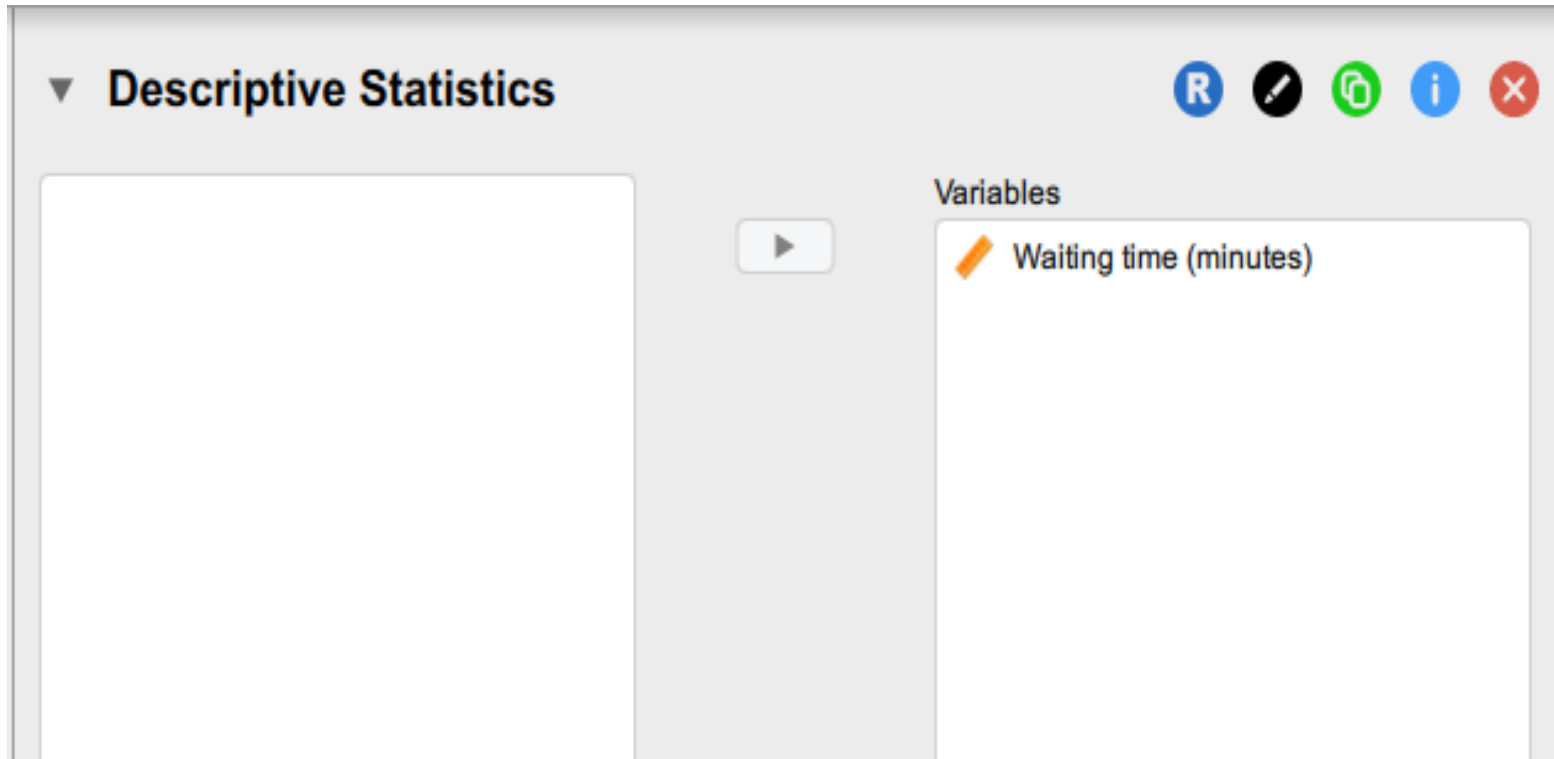
Testing Normality for Continuous Data (Graphically)



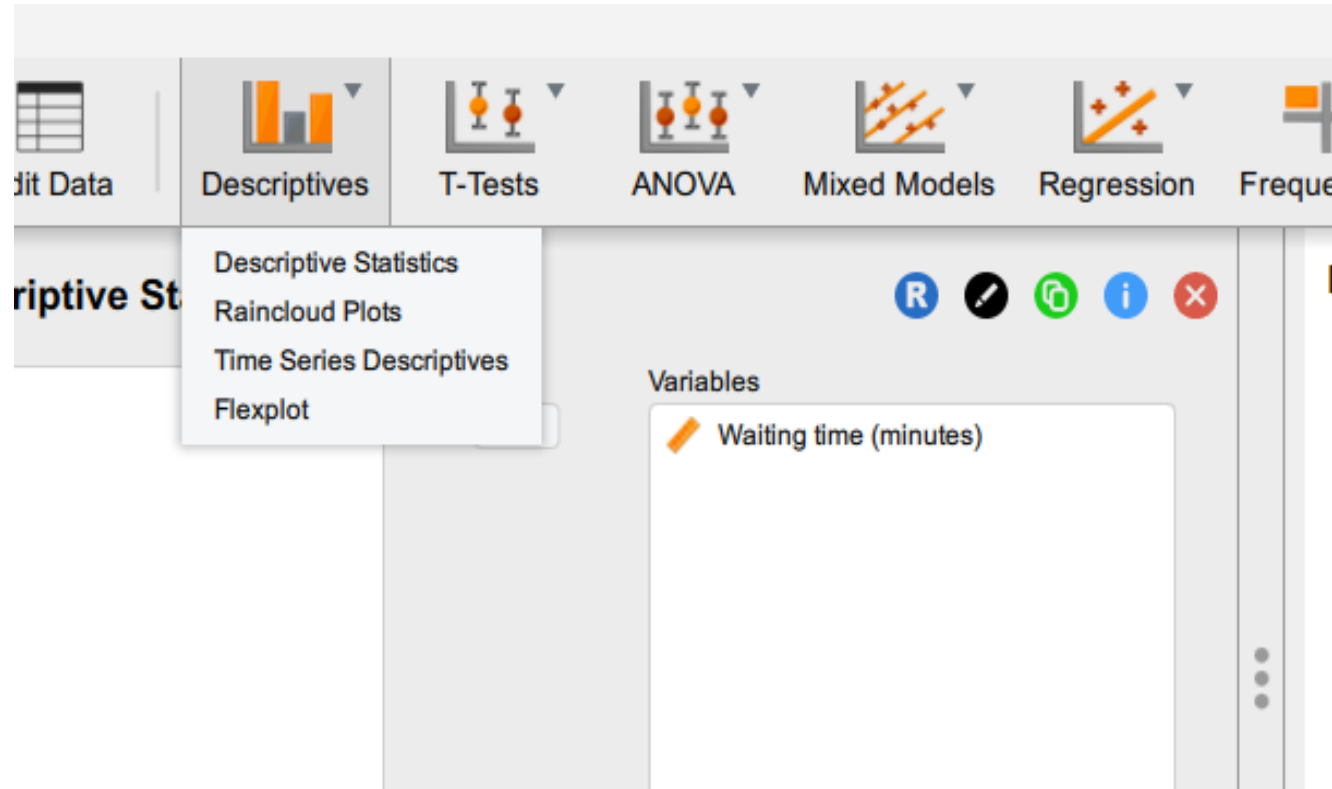
The screenshot shows a data entry interface with a header bar containing a menu icon, a table icon, and the text 'Edit Data'. Below the header is a table with 14 rows. The first column contains numbers 1 through 14, and the second column contains waiting times in minutes. The table is titled 'Waiting time (minutes)' with a pencil icon.

	Waiting time (minutes)
1	15
2	30
3	24
4	55
5	19
6	77
7	33
8	44
9	90
10	92
11	55
12	44
13	71
14	42

Testing Normality for Continuous Data (Graphically)



Testing Normality for Continuous Data (Graphically)



Testing Normality for Continuous Data (Graphically)

The screenshot displays the Minitab software interface. At the top, a menu bar includes 'Edit Data', 'Descriptives', 'T-Tests', 'ANOVA', 'Mixed Models', and 'Regression'. The 'Descriptives' menu is open, showing a list of variables with 'Waiting time (minutes)' selected and highlighted by a red box. Below the variable list, the 'Basic plots' section is expanded, and the 'Distribution plots' checkbox is checked, also highlighted by a red box. Other options in the 'Basic plots' section include 'Correlation plots', 'Interval plots', 'Q-Q plots', 'Pie charts', and 'Dot plots'. The 'Bin width type' is set to 'Sturges' and the 'Number of bins' is set to '30'.

Testing Normality for Continuous Data (Graphically)

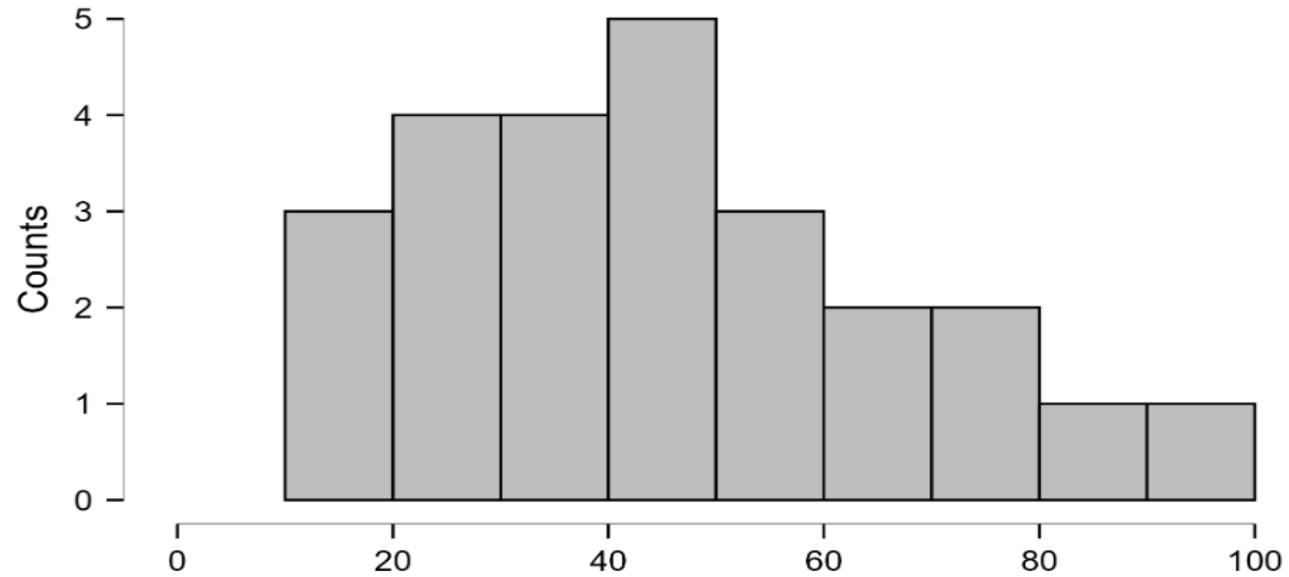
Descriptive Statistics

Descriptive Statistics

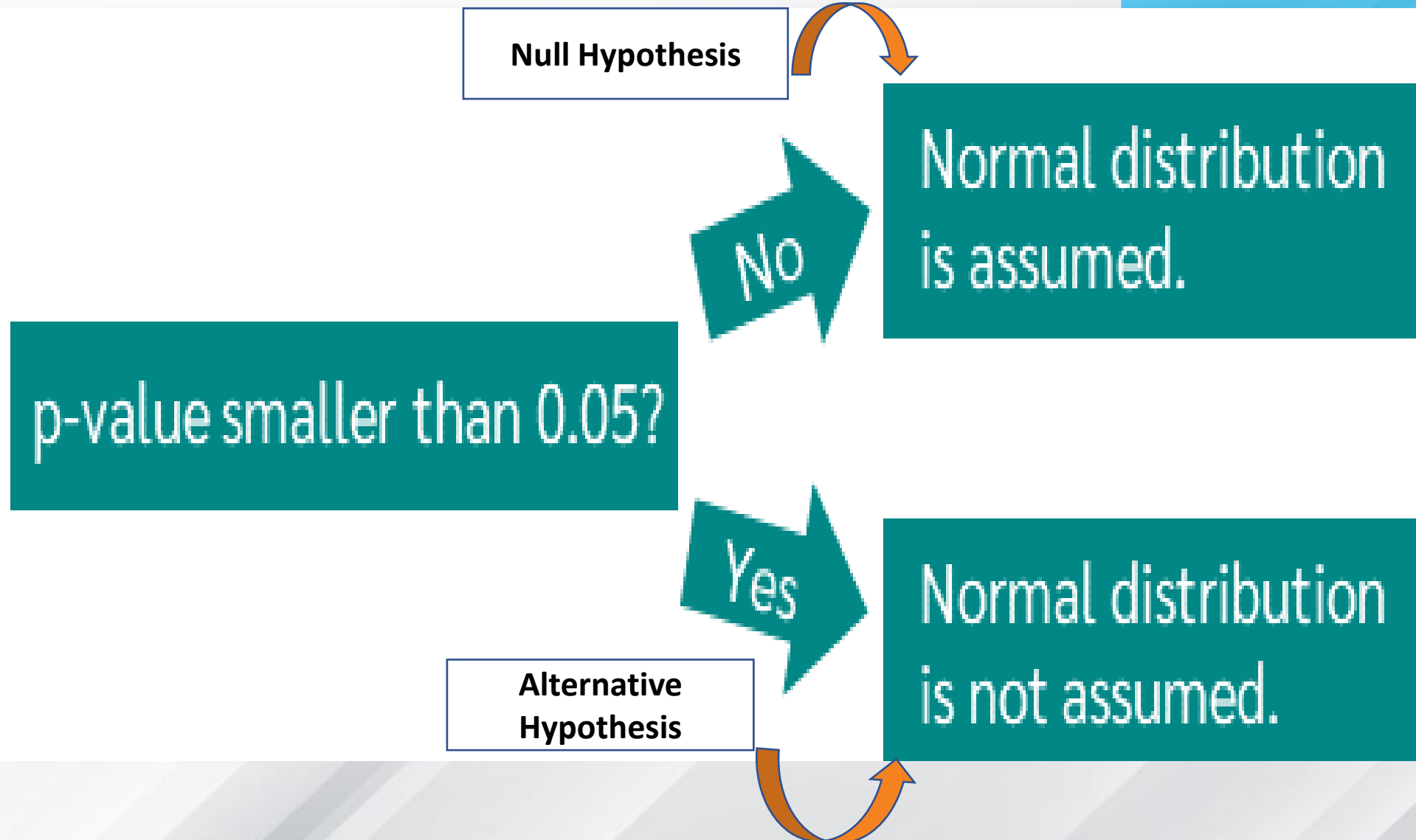
Waiting time (minutes)	
Valid	25
Missing	0
Mean	44.760
Std. Deviation	21.745
Minimum	15.000
Maximum	92.000

Distribution Plots

Waiting time (minutes)



Testing Normality for Continuous Data (Numerically)



Testing Normality for Continuous Data (Numerically)

Testing Normality for Continuous Data (Numerically)

- The Kolmogorov-Smirnov (K_S) / Shapiro-Wilk test is used to test the **null hypothesis** that a set of data comes from a **Normal distribution**.

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Reaction time trial 1	.073	233	.0047	.984	233	.0099
Reaction time trial 2	.090	233	.0001	.840	233	.0000
Reaction time trial 3	.385	235	.0000	.697	235	.0000
Reaction time trial 4	.045	226	.2000*	.998	226	.0759
Reaction time trial 5	.120	235	.0000	.991	235	.0000

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

P-value > 0.05, so, normally distributed data.

P-value < 0.05, so, non-normally distributed data.

Testing Normality for Continuous Data (Numerically)

☐ Transpose descriptives table

▼ Statistics

Sample size

- ☒ Valid
- ☒ Missing

Central tendency

- ☐ Mode
- ☐ Median
- ☒ Mean

Dispersion

- ☒ Std. deviation
- ☐ MAD
- ☐ IQR
- ☐ Range
- ☒ Maximum
- ☐ Coefficient of variation
- ☐ MAD robust
- ☐ Variance
- ☒ Minimum

Quantiles

- ☐ Quartiles
- ☐ Cut points for: equal groups
- ☐ Percentiles:

Distribution

- ☐ Skewness
- ☐ Kurtosis
- ☒ Shapiro-Wilk test
- ☐ Sum

Testing Normality for Continuous Data (Numerically)

Descriptive Statistics

Descriptive Statistics

Waiting time (minutes)	
Valid	25
Missing	0
Mean	44.760
Std. Deviation	21.745
Shapiro-Wilk	0.941
P-value of Shapiro-Wilk	0.155
Minimum	15.000
Maximum	92.000

Case Scenario (1)

➤ **Electronic Health Record (EHR) Implementation and Operational Efficiency:**

- A healthcare system introduces a new EHR system to reduce administrative errors and improve workflow efficiency. Management wants to evaluate whether the implementation of the EHR system reduces the average time physicians spend on documentation.

Case Scenario (1)

➤ **Electronic Health Record (EHR) Implementation and Operational Efficiency:**

- **Null Hypothesis (H_0):** The implementation of the EHR system does not reduce the time physicians spend on documentation.
- **Alternative Hypothesis (H_1):** The implementation of the EHR system reduces the time physicians spend on documentation.

Case Scenario (2)

➤ Healthcare Costs and Readmission Rates:

- A hospital introduces a new discharge protocol aimed at reducing 30-day readmission rates for patients with chronic heart failure. The hospital wants to test whether this intervention reduces overall healthcare costs.

Case Scenario (2)

➤ Healthcare Costs and Readmission Rates:

- **Null Hypothesis (H_0):** The new discharge protocol does not reduce overall healthcare costs associated with chronic heart failure readmissions.
- **Alternative Hypothesis (H_1):** The new discharge protocol reduces overall healthcare costs associated with chronic heart failure readmissions.

Case Scenario (3)

➤ Training Programs and Clinical Error Rates:

- Management implements a new training program for surgical staff to reduce the occurrence of clinical errors during operations. They need to test if the training program is effective.

Case Scenario (3)

➤ **Training Programs and Clinical Error Rates:**

- **Null Hypothesis (H_0):** The new training program does not reduce the clinical error rate in surgeries.
- **Alternative Hypothesis (H_1):** The new training program reduces the clinical error rate in surgeries.

Case Scenario (4)

➤ **Telemedicine and Follow-Up Compliance:**

- A hospital implements a telemedicine platform to improve follow-up compliance among chronic disease patients. Management wants to investigate whether this improves patient compliance with follow-up visits.

Case Scenario (4)

➤ **Telemedicine and Follow-Up Compliance:**

- **Null Hypothesis (H_0):** The telemedicine platform does not improve follow-up compliance among chronic disease patients.
- **Alternative Hypothesis (H_1):** The telemedicine platform improves follow-up compliance among chronic disease patients.

TO WRAP UP

- **Hypothesis Testing** is a fundamental tool in healthcare-related research for making decisions.
- **Many statistical tests** rely on the assumption of normality, making it crucial to assess whether your data meets this assumption.
- **P-value Interpretation** is a key metric in determining whether your results are statistically significant.

TO WRAP UP

- **Type I vs. Type II Errors:** Both errors have implications for research and decision-making. Minimizing these errors is critical for reliable conclusions.
- **Testing Normality can be done using** graphical and numerical methods to assess the normality of your data before conducting hypothesis testing.
- **You can apply these concepts** to real-life healthcare scenarios to make informed, evidence-based decisions.



Thank You