INFERENTIAL STATISTICS

INTRODUCTION

prepared by:

Gyro A. Madrona
Electronics Engineer

TOPIC OUTLINE

Inferential Statistics

Distribution

Normal Distribution

Histogram

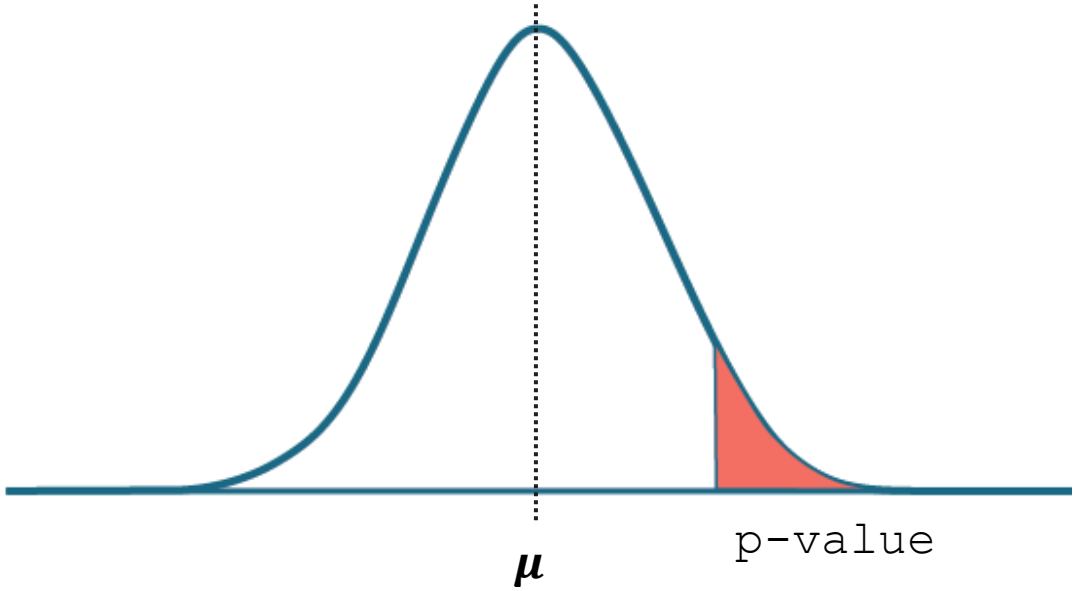
Normality Test

DISTRIBUTION

INFERENTIAL STATISTICS

Inferential statistics is a branch of statistics that analyzes and interprets data to make **conclusions** beyond the observed dataset. It focuses on drawing meaningful inferences about a population based on a sample using techniques such as **hypothesis testing**, **confidence intervals**, and **regression analysis**.

Normal Distribution:



DISTRIBUTION

Distribution or the probability distribution describes the **probabilities** or **frequencies** of different outcomes in an experiment or observed data.

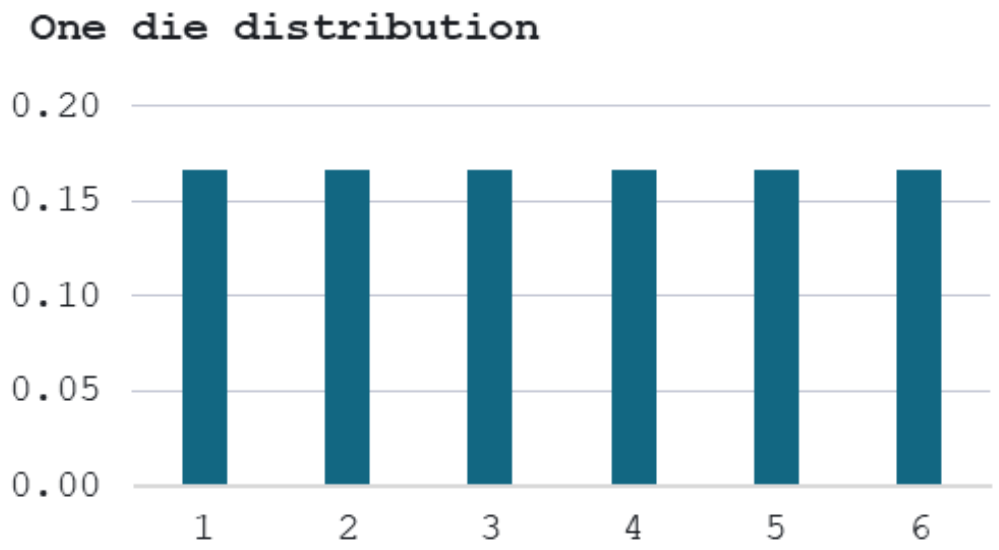
Rolling one die:

Outcome	Probability
1	1/6 or 0.17
2	1/6 or 0.17
3	1/6 or 0.17
4	1/6 or 0.17
5	1/6 or 0.17
6	1/6 or 0.17
7 (all else)	0

Sum of probabilities = 1 or 100%

UNIFORM DISTRIBUTION

Discrete data:



Rolling one die:

Outcome	Probability
1	1/6 or 0.17
2	1/6 or 0.17
3	1/6 or 0.17
4	1/6 or 0.17
5	1/6 or 0.17
6	1/6 or 0.17
7 (all else)	0

Sum of probabilities = 1 or 100%

DISTRIBUTION

36 possible outcomes:

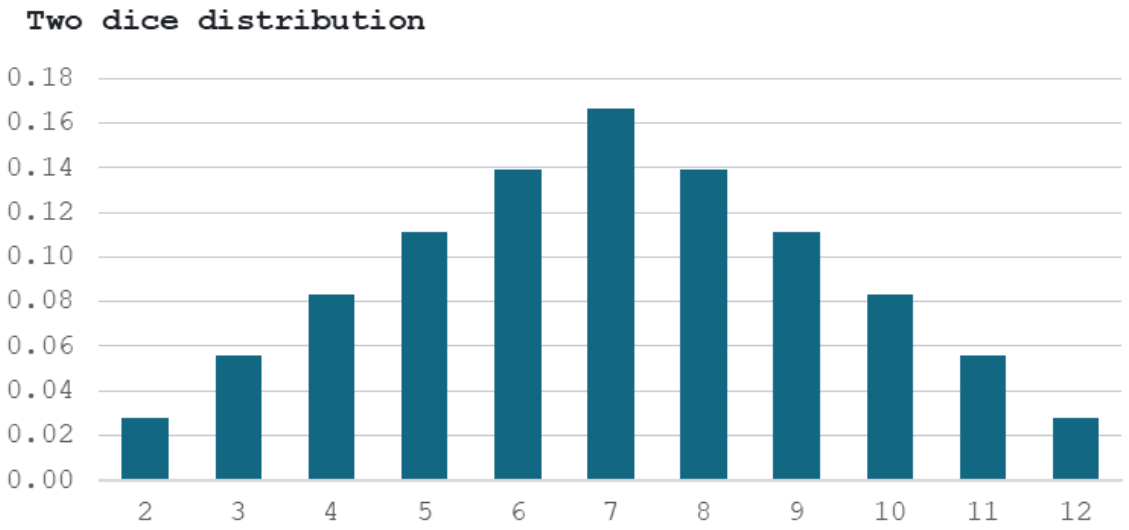
(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

Rolling two dice:

Sum	Probability
2	0.03
3	0.06
4	0.08
5	0.11
6	0.14
7	0.17
8	0.14
9	0.11
10	0.08
11	0.06
12	0.03
All else	0

DISTRIBUTION

Discrete data:



Rolling two dice:

Sum	Probability
2	0.03
3	0.06
4	0.08
5	0.11
6	0.14
7	0.17
8	0.14
9	0.11
10	0.08
11	0.06
12	0.03
All else	0

NORMAL DISTRIBUTION

NORMAL DISTRIBUTION

A **normal distribution** is a probability distribution where the values of a random variable are distributed symmetrically. Also known as **Gaussian** distribution or bell curve because of its shape.

Johann Carl Friedrich Gauss

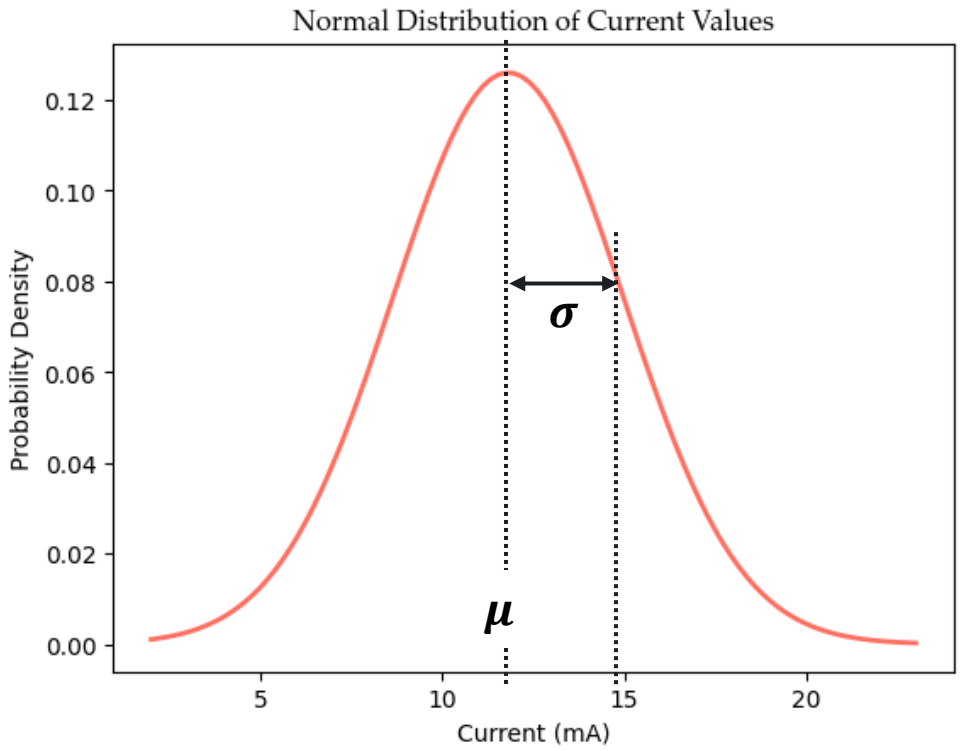
NORMAL DISTRIBUTION

A **normal distribution** is a probability distribution where the values of a random variable are distributed symmetrically. Also known as **Gaussian** distribution or **bell curve** because of its shape.

Denoted by:

$$N \sim (\mu, \sigma^2)$$

Bell Curve:



mean = median = mode

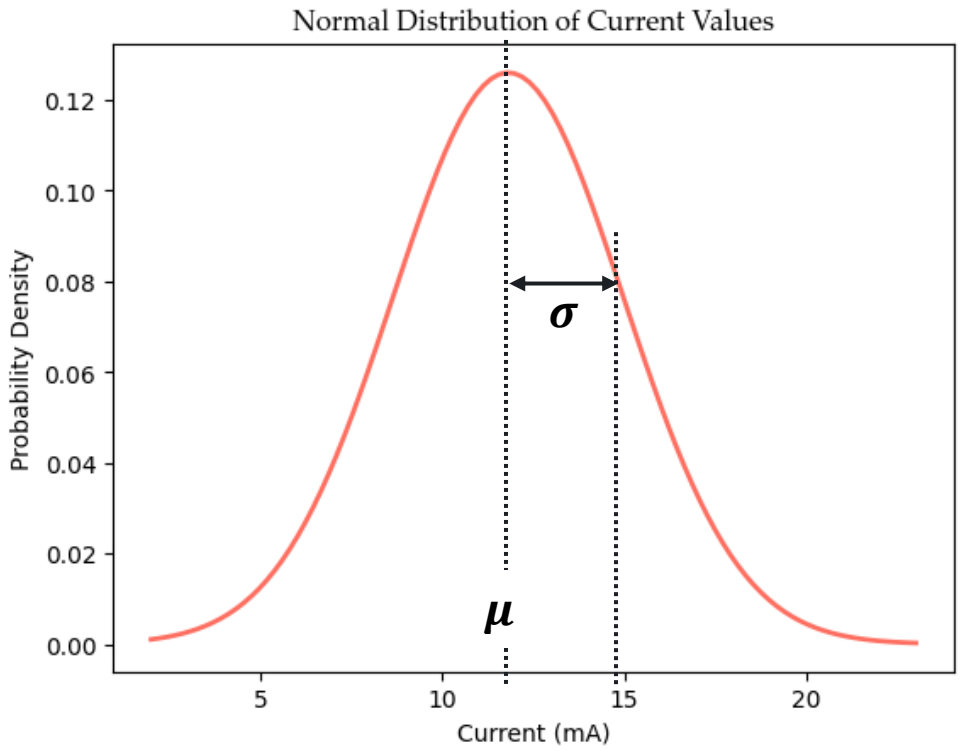
NORMAL DISTRIBUTION

A **normal distribution** is a probability distribution where the values of a random variable are distributed symmetrically. Also known as **Gaussian** distribution or **bell curve** because of its shape.

Formula:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

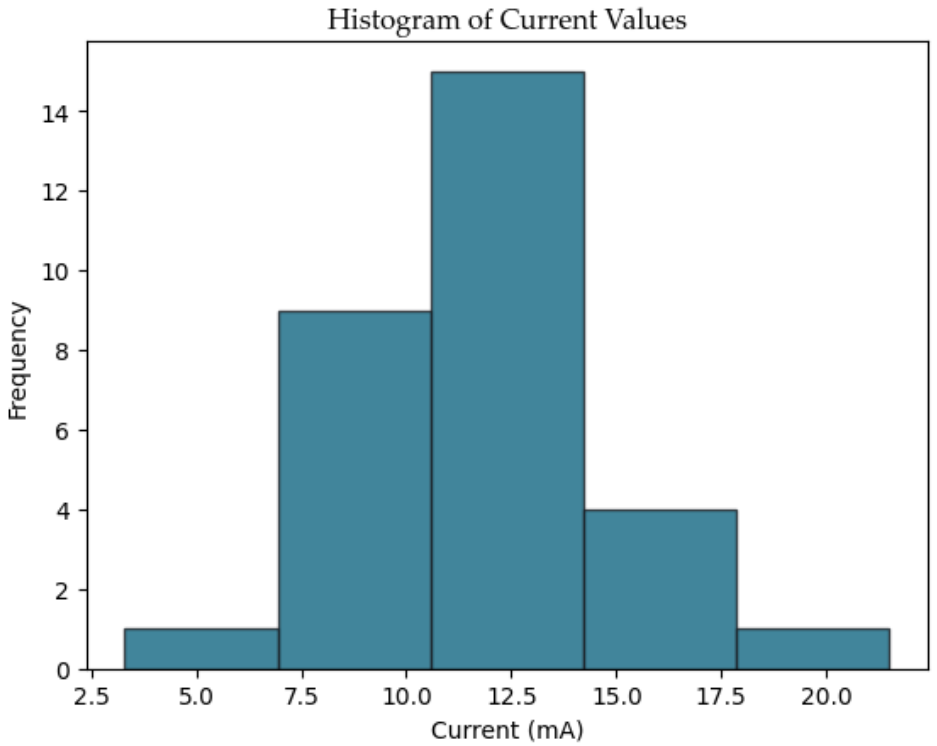
Bell Curve:



mean = median = mode

HISTOGRAM

Histogram:



Histograms are used to visualize the **shape**, **spread**, and **central tendency** of data, making them a useful tool for assessing whether a dataset follows a normal distribution or deviates from it.

NORMALITY TEST

Test	Recommendation
Shapiro-Wilk	Small sample sizes (< 50)
Anderson-Darling	Moderate sample sizes ($50 - 5000$)
Kolmogorov-Smirnov	Large sample sizes (> 5000)

Interpreting the p-value in a Normality Test

High p-value ($p > 0.05$):

The data may be normally distributed.

Low p-value ($p \leq 0.05$):

The data is likely **not** normally distributed.

EXERCISE

The dataset consists of 30 samples of current measurements (in mA). Generate a **normal distribution plot** and assess the normality of the data using the **Shapiro-Wilk** test in a Jupyter Notebook.

Dataset:

L14-current-data.csv

Current Response

Sample	Current
1	12.0
2	15.0
3	8.3
4	9.7
5	12.0
6	13.9
7	14.1
8	9.2
9	12.4
10	13.7
11	10.6
12	21.5
13	12.0

LABORATORY