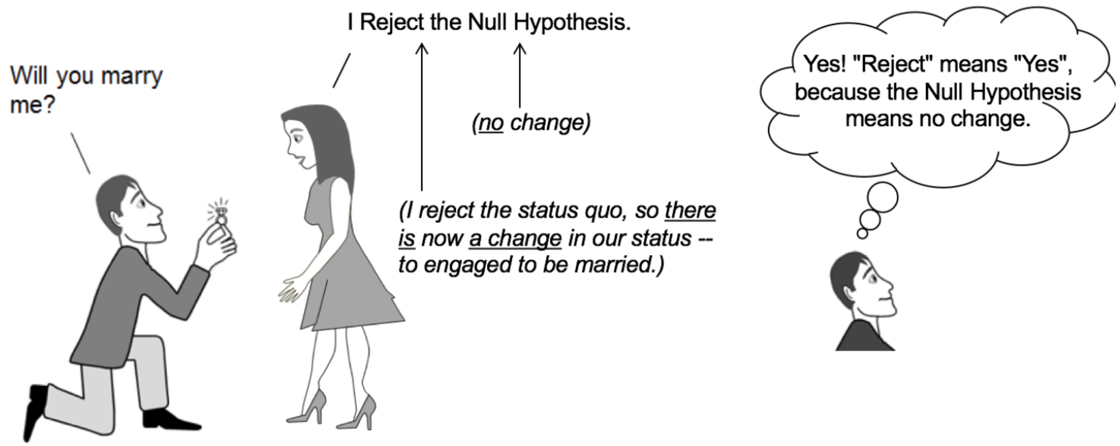
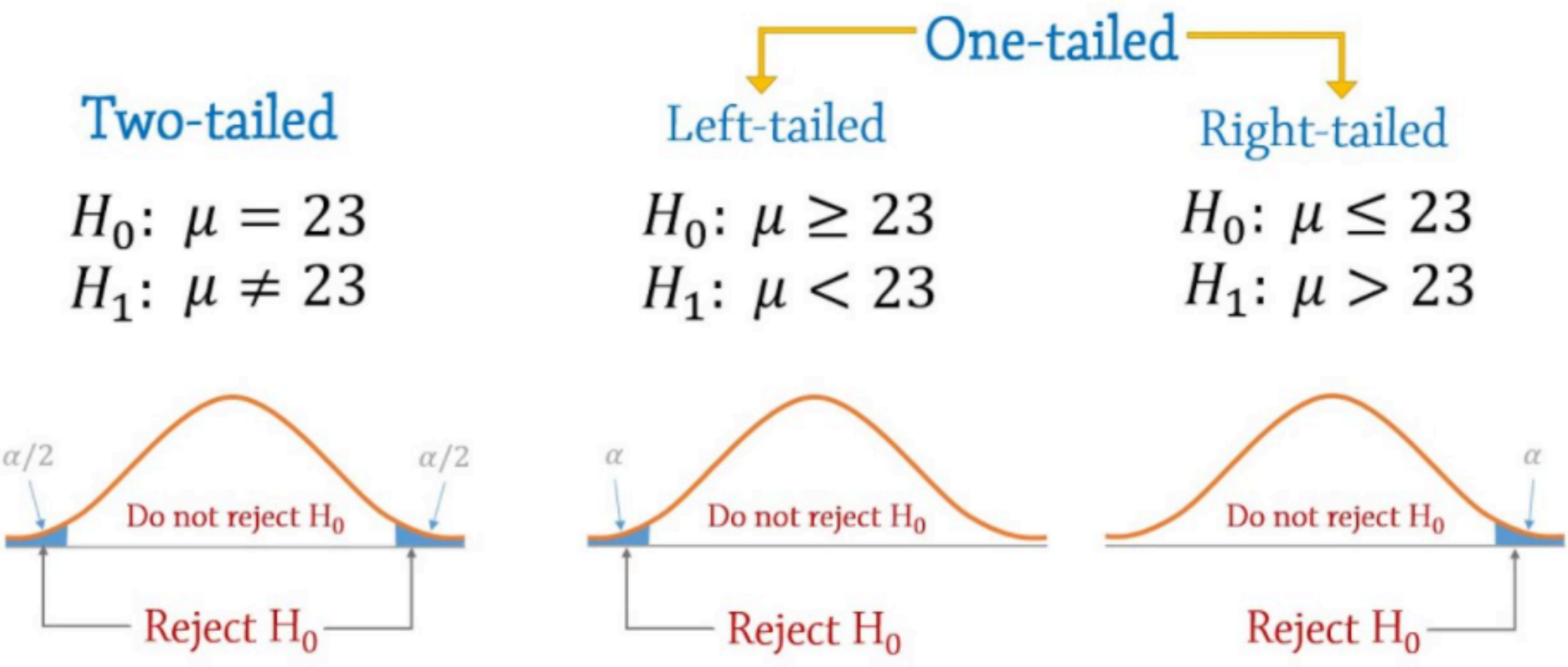
Hypothesis Testing



Hypothesis Testing



Hypothesis Testing

- A statistical hypothesis is an assumption made by the researcher about the data of the population collected for any experiment
- It is a formal process of validating the hypothesis
- It should consider the entire population, but in practice, this is not possible
- Instead, we use random samples from a population
- On the basis of the result from testing over the sample data, it either accepts or rejects the hypothesis
- It is not mandatory for this assumption to be true every time

https://data-flair.training/blogs/hypothesis-testing-in-r/

Hypothesis Testing

- **Null Hypothesis**
  - Tests the validity of a claim or assumption that is made about the population
  - It ascribes the claim to samples of the population
  - The null hypothesis testing is denoted by H0
- **Alternative Hypothesis**
  - An alternative hypothesis would be considered valid if the null hypothesis is fallacious
  - The evidence that is present is basically the sample data and the statistical computations that accompany it.
  - The alternative hypothesis testing is denoted by H1or Ha.

Exercise

- The IQ for the adult population is normally distributed with a mean of 100 with a standard deviation of 15.
- A researcher believes that this value has changed.
- The researcher decides to test the IQ of 75 adults chosen at random.
- The average IQ of the sample is 105
- Is there enough evidence to suggest that the average IQ has changed?
- Show your conclusion by doing a hypothesis testing

https://www.youtube.com/watch?v=lNoxKsuJ6Xc

Hypothesis Testing

1. State null (H_0) and alternative (H_1) hypothesis
2. Choose level of significance (α)
3. Find critical values
4. Find test statistic
5. Draw your conclusion

Hypothesis Testing

$$H_0: \mu = 100$$

$$H_1: \mu \neq 100$$

Hypothesis Testing

$$H_0: \mu = 100$$

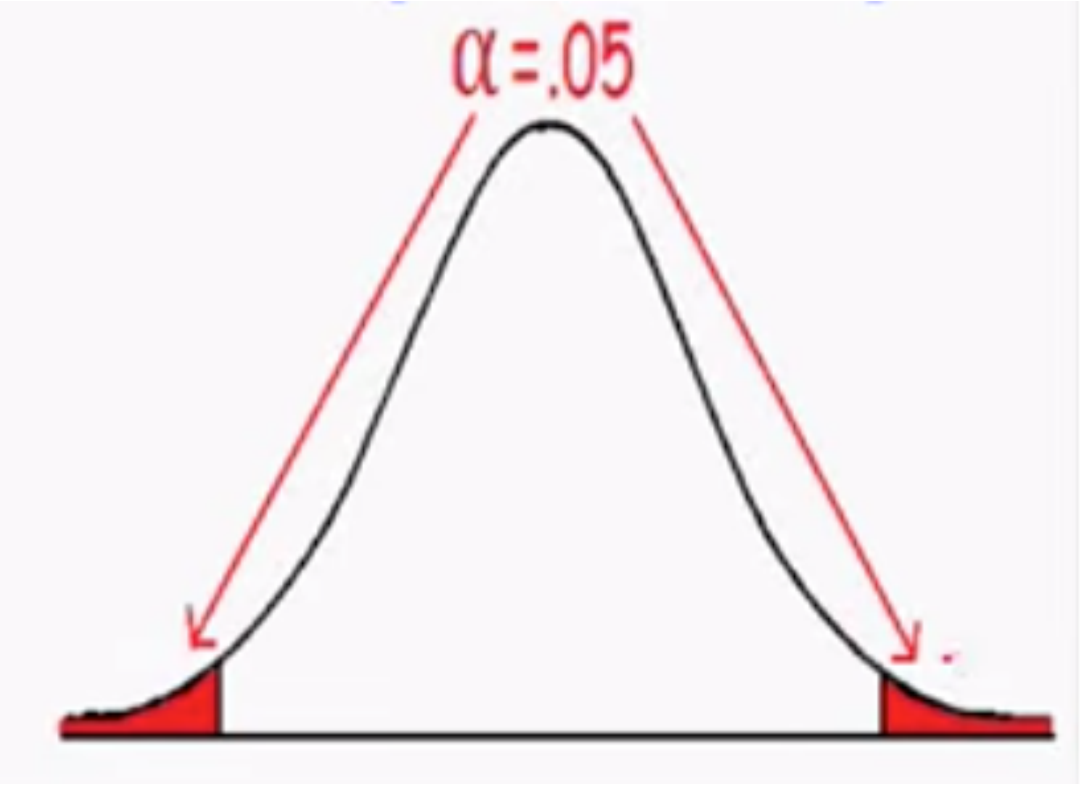
(2 tailed test)

$$H_1: \mu \neq 100$$

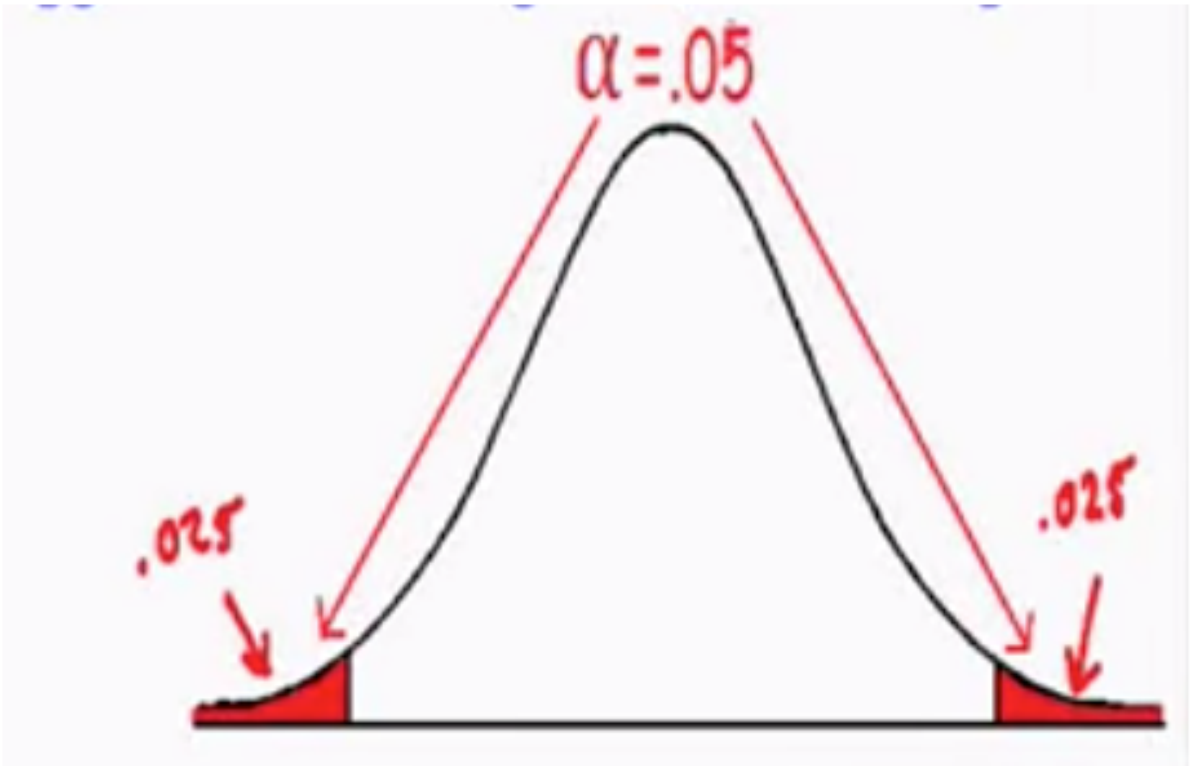
Hypothesis Testing

1. State null (H_0) and alternative (H_1) hypothesis
2. Choose level of significance (α)
3. Find critical values
4. Find test statistic
5. Draw your conclusion

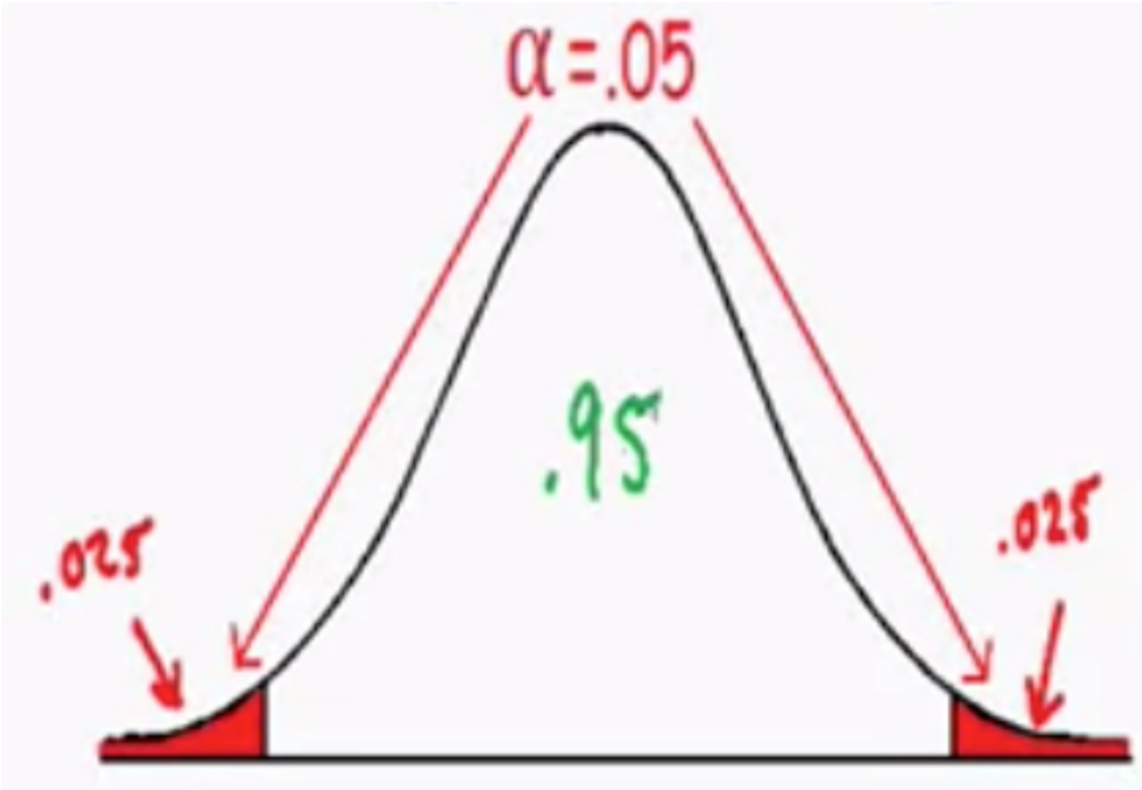
Hypothesis Testing



Hypothesis Testing



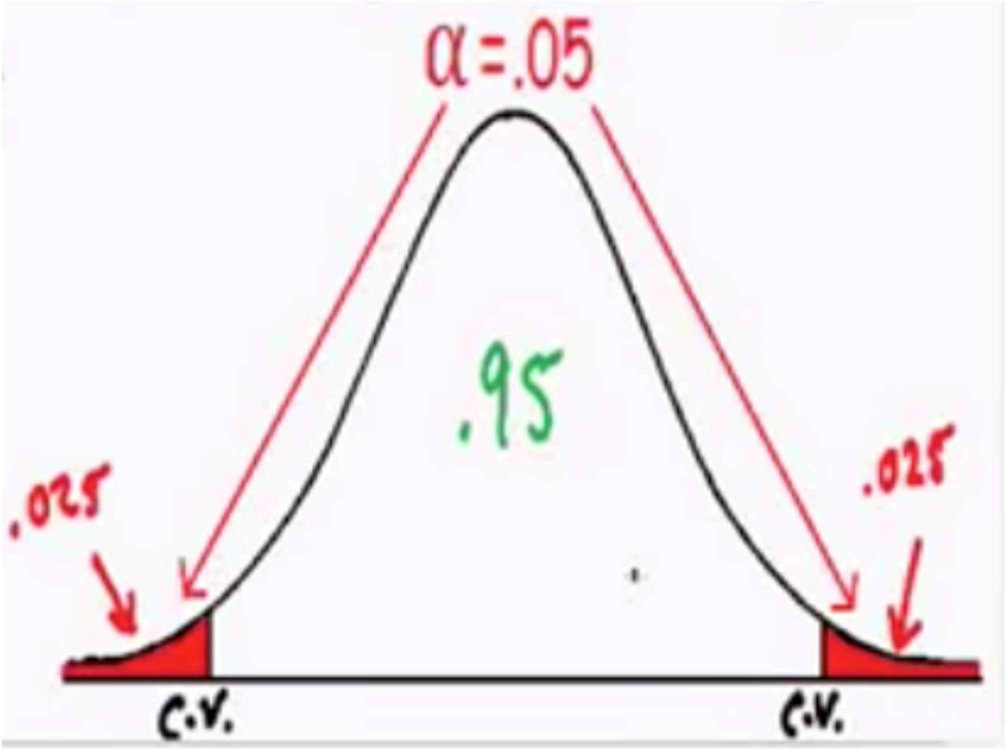
Hypothesis Testing



Hypothesis Testing

1. State null (H_0) and alternative (H_1) hypothesis
2. Choose level of significance (α)
3. Find critical values
4. Find test statistic
5. Draw your conclusion

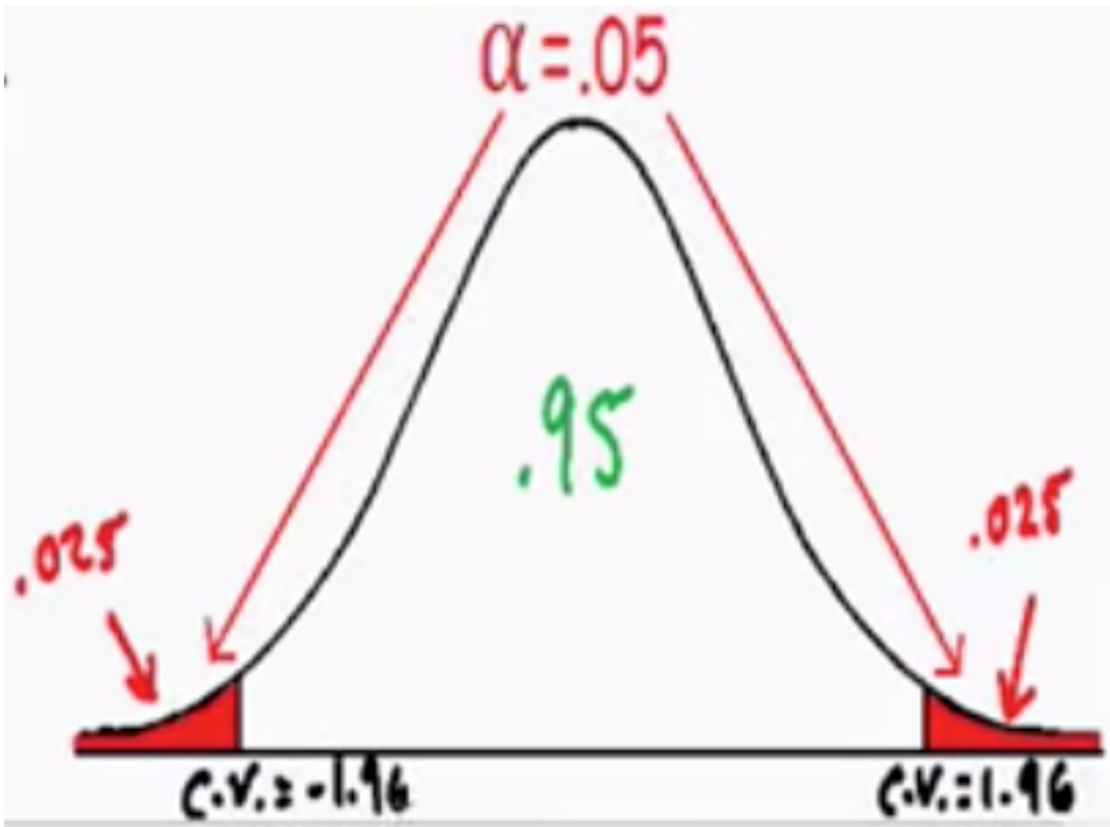
Critical Values



Z-Scores Table

Confidence Level	Area between 0 and z-score	Area in one tail (alpha/2)	z-score
50%	0.2500	0.2500	0.674
80%	0.4000	0.1000	1.282
90%	0.4500	0.0500	1.645
95%	0.4750	0.0250	1.960
98%	0.4900	0.0100	2.326
99%	0.4950	0.0050	2.576

Rejection Region for Null Hypothesis (In Red)



Hypothesis Testing

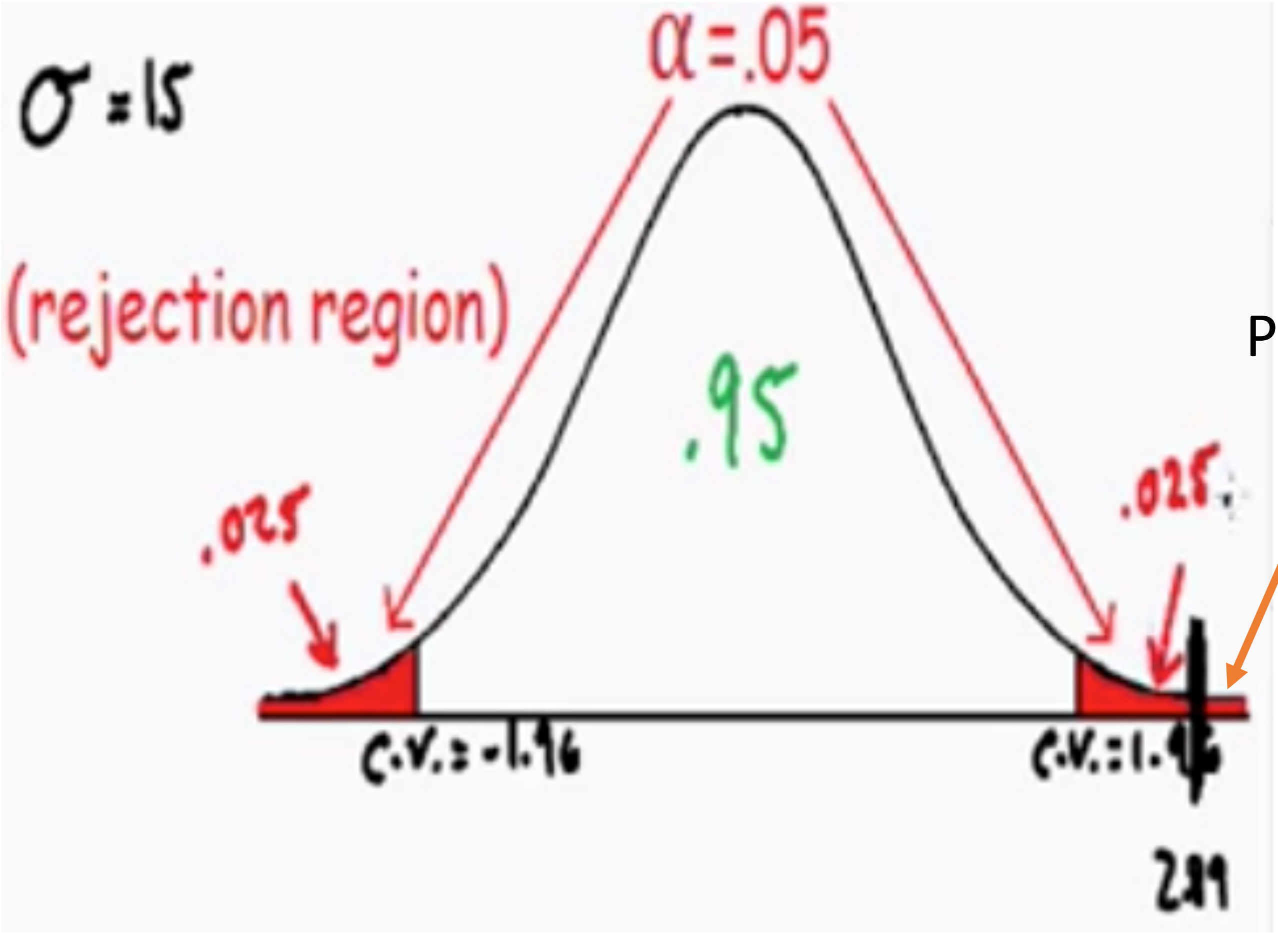
1. State null (H_0) and alternative (H_1) hypothesis
2. Choose level of significance (α)
3. Find critical values
4. Find test statistic
5. Draw your conclusion

Find Test Statistic

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{105 - 100}{15/\sqrt{75}} = 2.89$$

test statistic

Find Test Statistic



P-value < α

Hypothesis Testing

1. State null (H_0) and alternative (H_1) hypothesis
2. Choose level of significance (α)
3. Find critical values
4. Find test statistic
5. Draw your conclusion

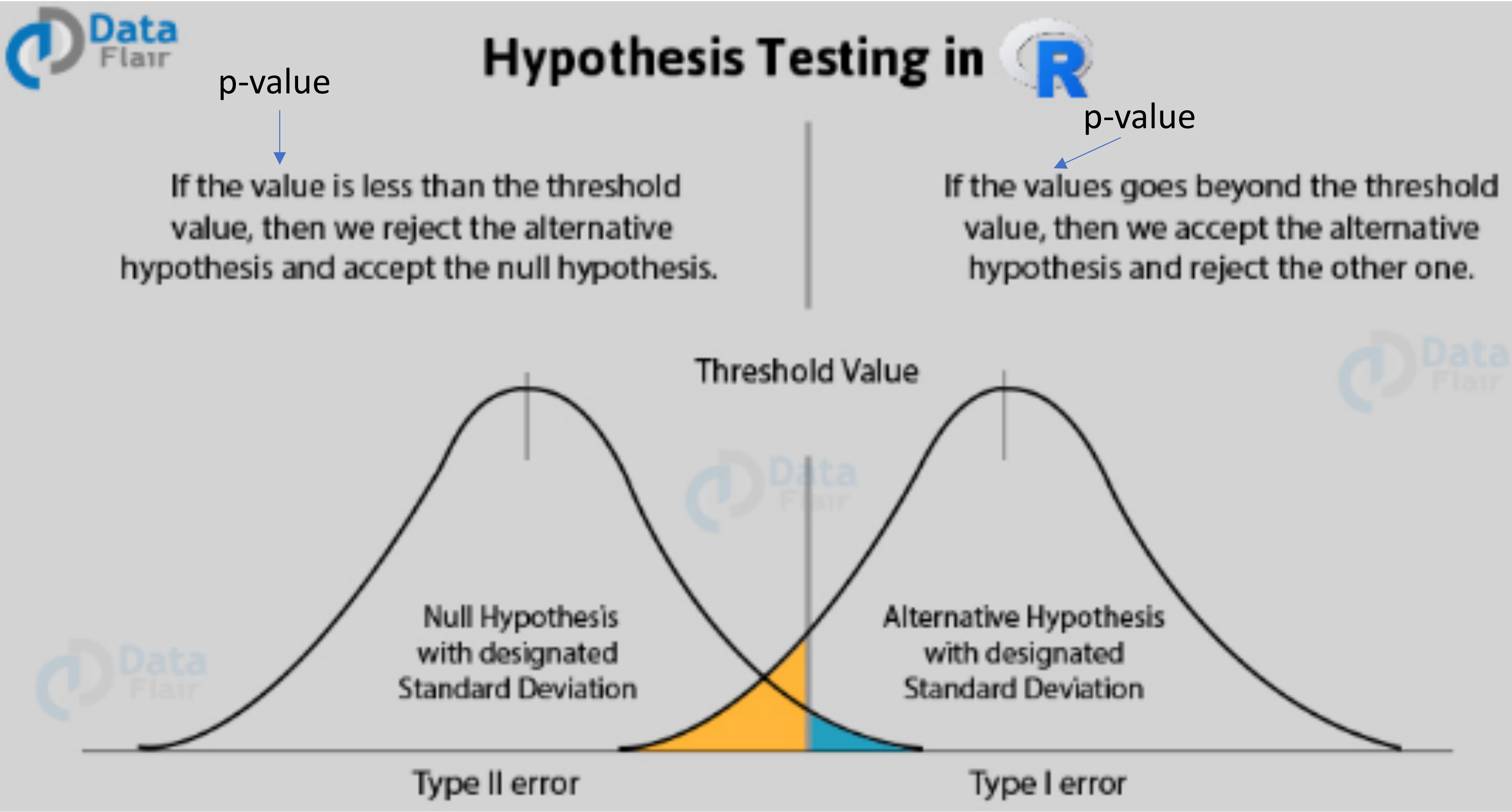
The p-value

- We use a **p-value** to weigh the strength of the evidence in the sample data.
- The p-value ranges between 0 and 1.
- It is interpreted in the following way:
  - A small p-value (typically ≤ 0.05) indicates strong evidence against the null hypothesis, so you reject it.
  - A large p-value (> 0.05) indicates weak evidence against the null hypothesis, so you fail to reject it.

https://data-flair.training/blogs/hypothesis-testing-in-r/

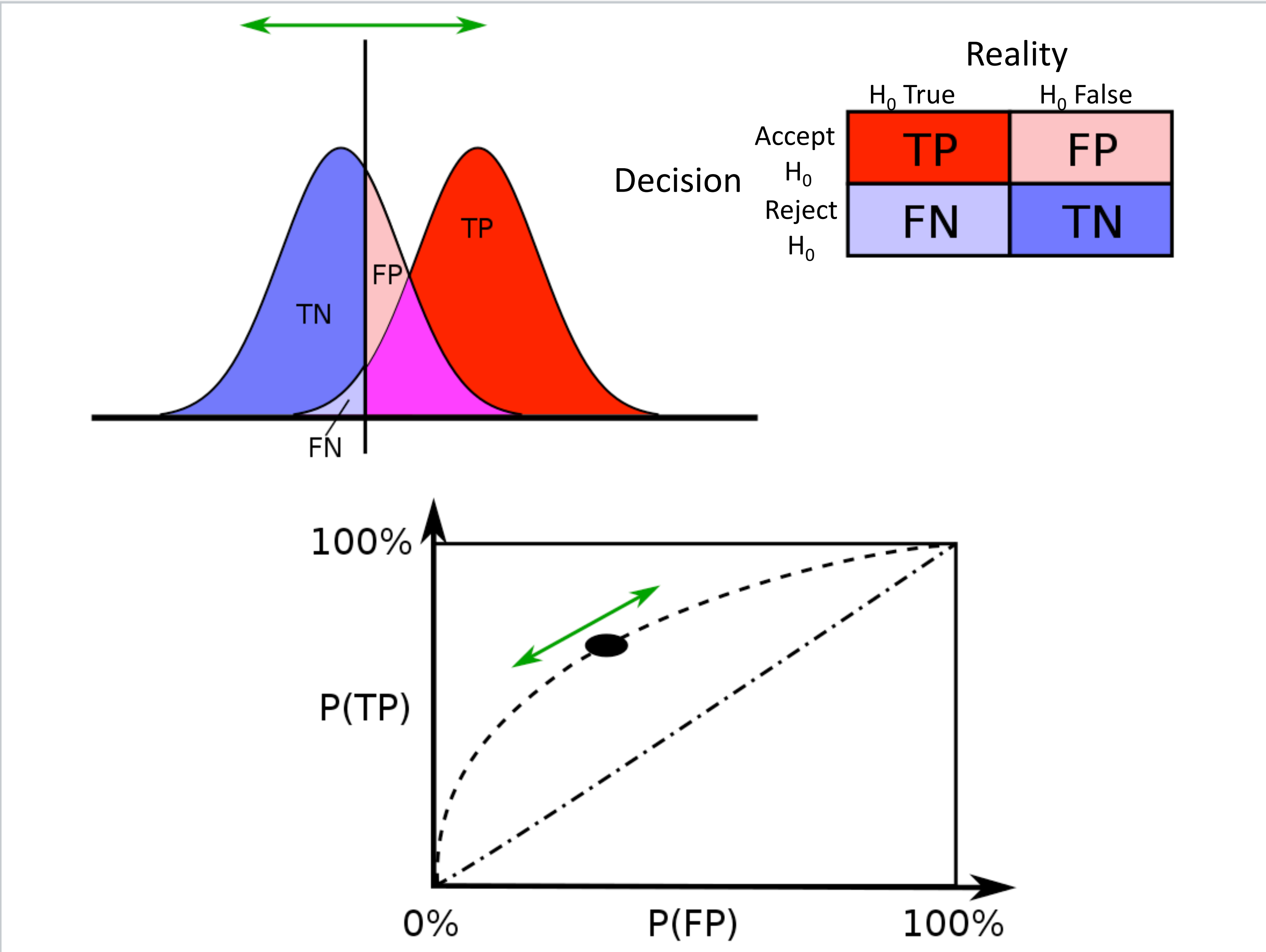
Draw a Conclusion

- Reject H_o
- Accept H_1

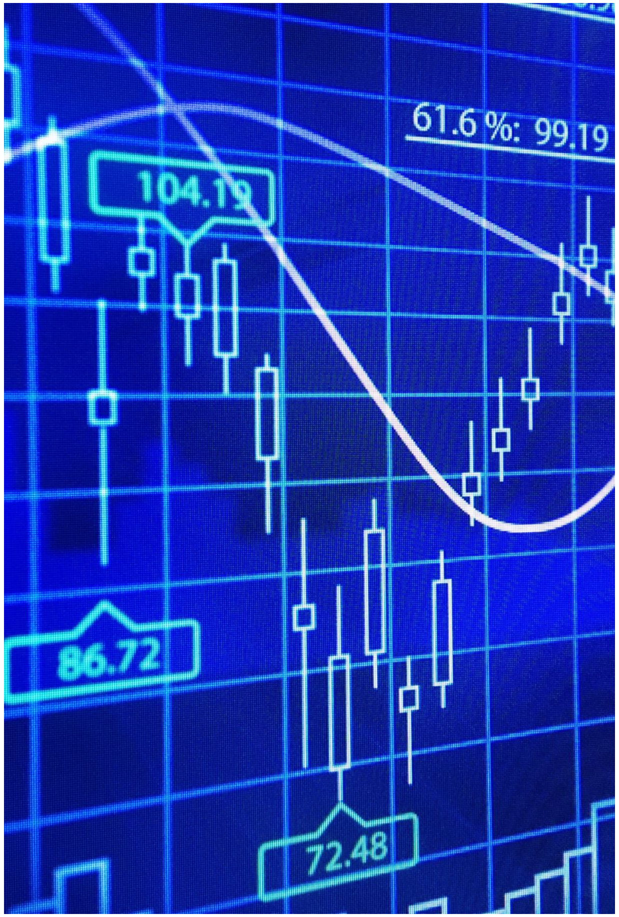

Data Flair
Hypothesis Testing in R
p-value
If the value is less than the threshold value, then we reject the alternative hypothesis and accept the null hypothesis.
p-value
If the values goes beyond the threshold value, then we accept the alternative hypothesis and reject the other one.
Threshold Value
Null Hypothesis with designated Standard Deviation
Alternative Hypothesis with designated Standard Deviation
Type II error
Type I error

Decision Errors

- The two types of error that can occur from hypothesis testing:
- **Type I Error**
  - Occurs when the researcher rejects a null hypothesis when it is true
  - It is called a False Positive (FP)
  - **Significance lev**el is the probability of a Type I error
  - The significance level is represented by the symbol α (Blue area)
- **Type II Error**
  - Accepting a false null hypothesis H0 is referred to as the **Type II** error
  - The term power of the test is used to express the probability of Type II error while testing hypothesis
  - The power of the test is represented by the symbol β (Yellow area).



The results obtained from positive sample (left curve) overlap with the results obtained from negative samples (right curve). By moving the result cutoff value (vertical bar), the rate of false positives (FP) can be decreased, at the cost of raising the number of false negatives (FN), or vice-versa.



Test Statistics to find p-values

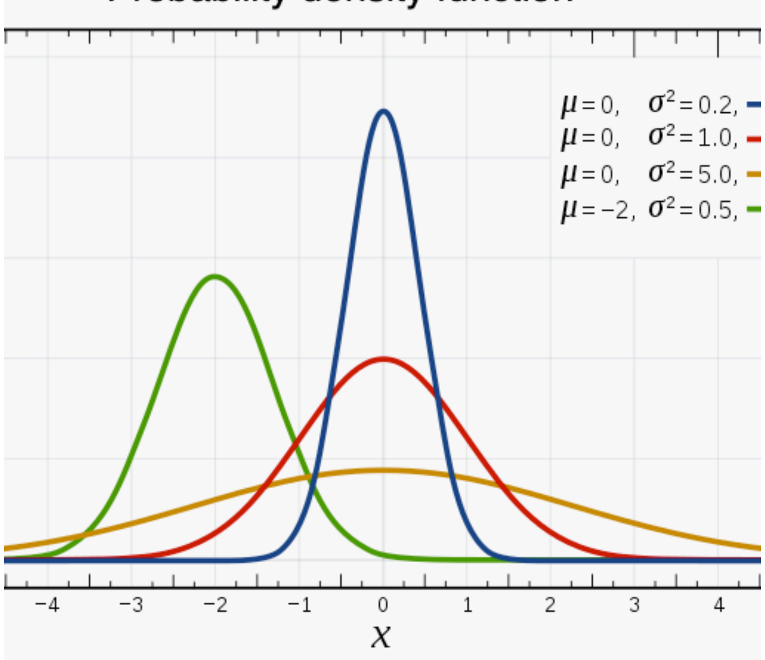
- Z scores Test
- t-Student scores Test
- F scores Test - ANOVA
- Xi^2 scores Test

Z-Scores

- A Z-score estimates how far from the mean a data point is
- it’s a measure of how many standard deviations below or above the population mean a raw score is
- A Z-score can be placed on a normal distribution curve
- Z-scores range from -3 standard deviations (which would fall to the far left of the normal distribution curve) up to +3 standard deviations (which would fall to the far right of the normal distribution curve).
- In order to use a Z-score, you need to know the mean μ and also the population standard deviation σ.

https://www.statisticshowto.datasciencecentral.com/probability-and-statistics/z-score/

- https://en.wikipedia.org/wiki/Normal_distribution



he red curve is the *standard normal distribution*

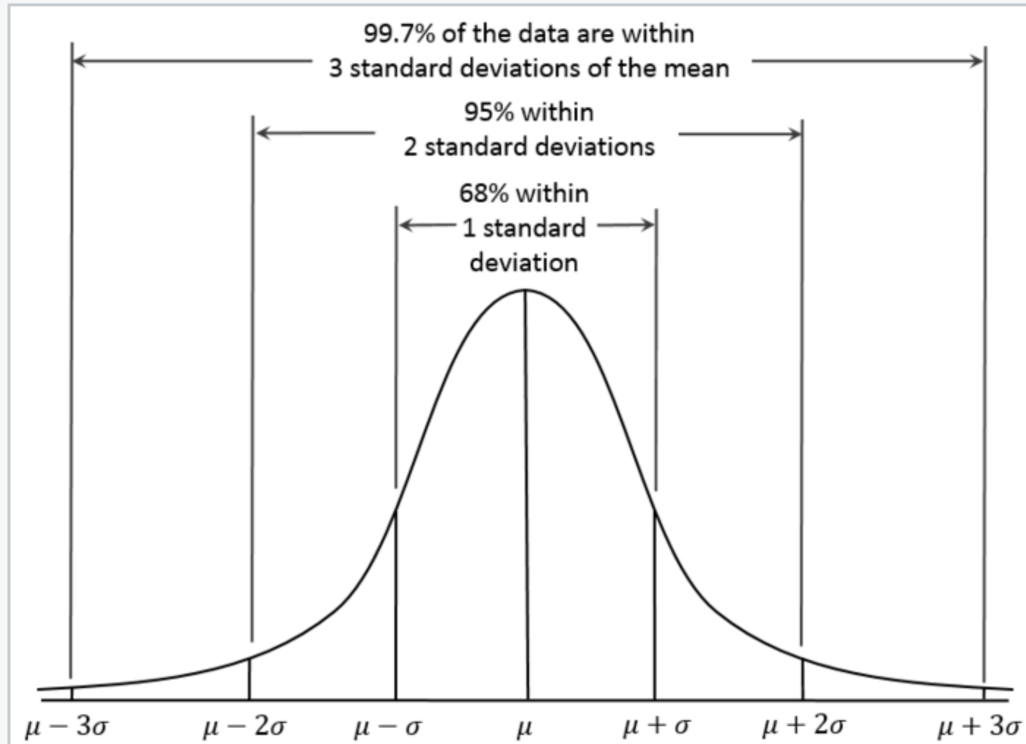
Z-scores -> Normal Distribution

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

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

Normal Distribution

- https://en.wikipedia.org/wiki/Normal_distribution



For the normal distribution, the values less than one standard deviation away from the mean account for 68.27% of the set; while two standard deviations from the mean account for 95.45%; and three standard deviations account for 99.73%.

Skewness and Kurtosis

- Real life data rarely, if ever, follow a perfect normal distribution.

- The skewness and kurtosis coefficients measure how different a given distribution is from a normal distribution

Skewness

- Measures the symmetry of a distribution
- The normal distribution is symmetric and has a skewness of zero
- If the distribution of a data set has a skewness less than zero, or negative skewness, then the left tail of the distribution is longer than the right tail
- Positive skewness implies that the right tail of the distribution is longer than the left.

Kurtosis

- Measures the thickness of the tail ends of a distribution in relation to the tails of the normal distribution
- Distributions with large kurtosis exhibit tail data exceeding the tails of the normal distribution (e.g., five or more standard deviations from the mean)
- Distributions with low kurtosis exhibit tail data that is generally less extreme than the tails of the normal distribution
- The normal distribution has a kurtosis of three, which indicates the distribution has neither fat nor thin tails.
- If an observed distribution has a kurtosis greater than three, the distribution is said to have heavy tails when compared to the normal distribution
- If the distribution has a kurtosis of less than three, it is said to have thin tails when compared to the normal distribution.

Normality Tests with Python

Download Python script normal-distribution.py

https://machinelearningmastery.com/statistical-hypothesis-tests-in-python-cheat-sheet/

Z-scores Test

- To do a z-core test we need to calculate z-scores
- To calculate a z-score we need the population **mean** and the population **standard deviation**.
- A one-sample z-test allows for us to see if a particular piece/group of data is actually from a larger population of data.
- Situations which warrant a z-test:
  - Sample size greater than 30
  - Independent data points
  - Normally distributed data
  - Randomly selected data
  - Equal sample sizes
- **Null Hypothesis:** Sample mean is same as the population mean
- **Alternate:** Sample mean is not the same as the population mean

Z-scores Test with Python

- Load Python script z-test.py

https://towardsdatascience.com/hypothesis-testing-in-machine-learning-using-python-a0dc89e169ce

T- Test (Student's)

- The Student's T-test is a method for comparing one or two samples
- Similar to z-scores
- Determines whether a sample is different from the population distribution
- Determines whether two samples are different or have the same mean
- Used for small-size samples or if you don't know the population standard deviation
- This is a parametric test, and the data should be normally distributed
  - https://www.statisticshowto.datasciencecentral.com/t-statistic/

T-Test

- T-tests differ from z-tests because the mean and standard deviations of the population are not known.
- Situations to use a T-Test:
  - Small samples (Less that 30)
  - Data is independent
  - Approximately normally distributed data
  - Similar amount of variance within the groups being compared (homogenous)

T-Test

- There are three T-Test types:
- **One Sample T-Test:** If there is one group being compared against a standard value
- **Two Sample T-Test (Independent):** If the groups come from different populations
- **Paired T-Test:** If the groups come from a single population
- **T-Scores:**
  - **$t = (x1 — x2) / (\sigma / \sqrt{n1} + \sigma / \sqrt{n2})$,** where
  - x1 = mean of sample 1
  - x2 = mean of sample 2
  - n1 = size of sample 1
  - n2 = size of sample 2

T-Test with Python

Load Python script t-test.py

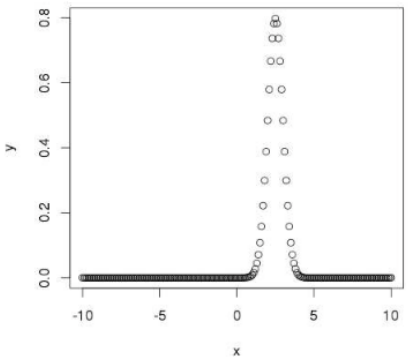
https://towardsdatascience.com/hypothesis-testing-in-machine-learning-using-python-a0dc89e169ce

17 Normality Tests with Python

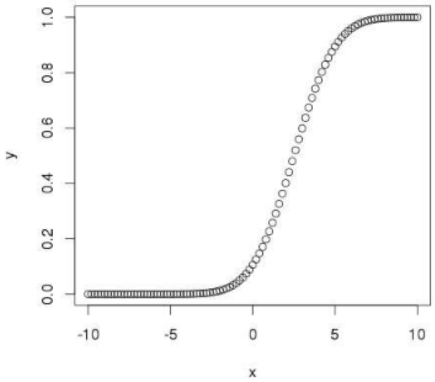
Download Python script
normality-tests.py

https://machinelearningmastery.com/statistical-hypothesis-tests-in-python-cheat-sheet/

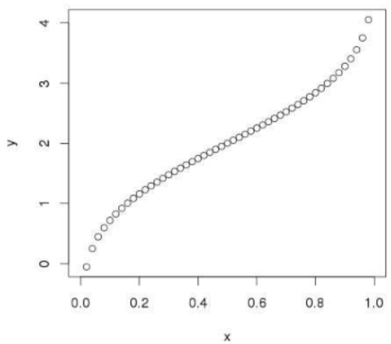
Z-scores Test with R



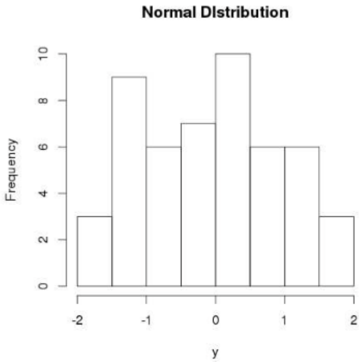

```
x <- seq(-10, 10, by = .1)
y <- dnorm(x, mean = 2.5, sd = 0.5)
plot(x,y)
```




```
x <- seq(-10,10,by = .2)
y <- pnorm(x, mean = 2.5, sd = 2)
plot(x,y)
```




```
x <- seq(0, 1, by = 0.02)
y <- qnorm(x, mean = 2, sd = 1)
plot(x,y)
```




```
y <- rnorm(50)
hist(y, main = "Normal DIstribution")
```

https://stat.ethz.ch/R-manual/R-devel/library/stats/html/Normal.html