Together for Tomorrow!
Enabling People
Education for Future Generations

Chapter 4.

Probability and Statistics

UNIT 4.

Statistical Hypothesis Testing

Samsung Innovation Campus

Unit 4. Statistical Hypothesis Testing

What this unit is about:

- You will learn about the principles of hypothesis testing.
- You will learn about the t-test for one sample.
- You will learn about the t-test for independent samples and paired samples.
- You will learn about the different kinds of Chi-squared test.
- You will learn about the F-test.

Expected outcome:

- Ability to apply the different kinds of hypothesis testing methods.

How to check your progress:

- Coding Exercises.
- Quiz.

Together for Tomorrow! Enabling People
Education for Future Generations

Chapter 4.

Probability and Statistics

Hypothesis Testing (1/8)

- Principles of the hypothesis testing in statistics:
  - Tests a hypothesis about the population parameters using the sample statistics.

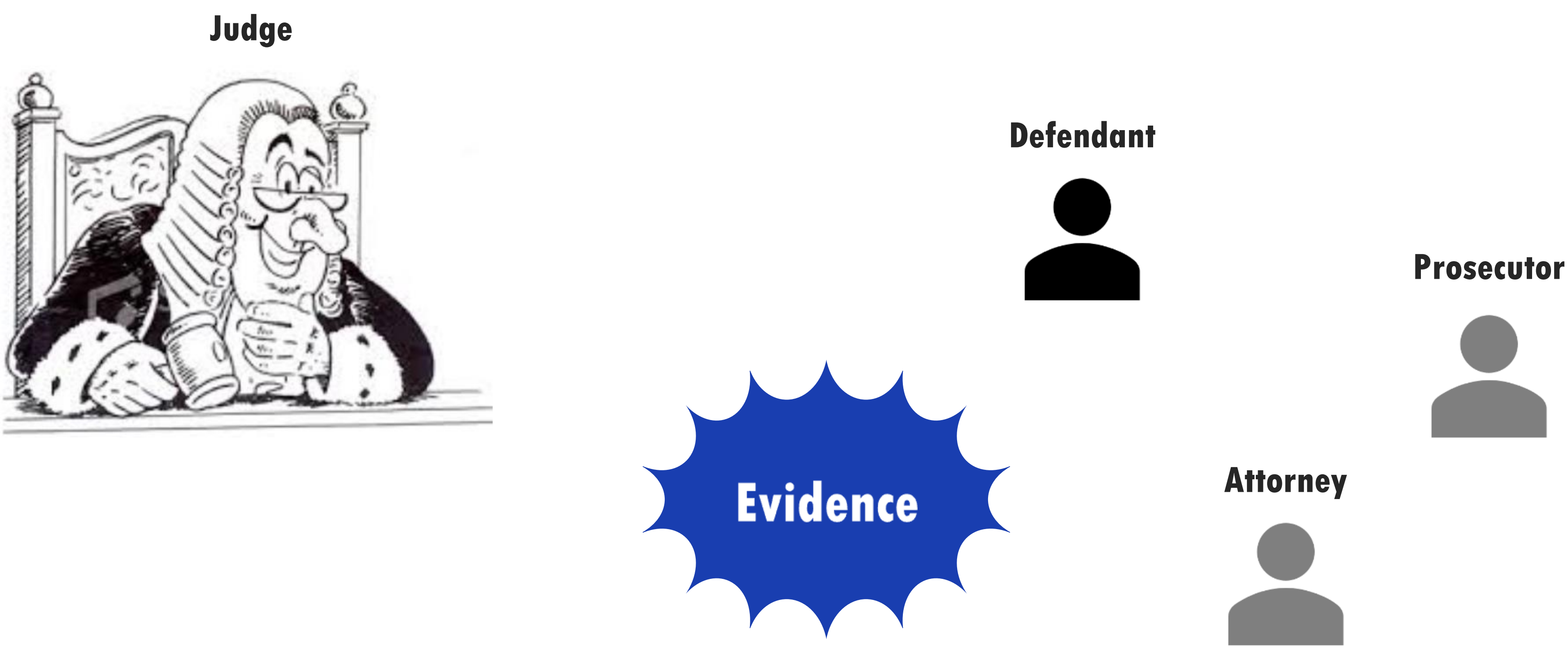
Ex) "The average sleep time of adults in USA is 7 hours." $\Leftarrow$ True or False?

Ex) "The average annual wage in Islamabad is Rs. 85,000." $\Leftarrow$ True or False?

Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Testing (2/8)

- Principles of the hypothesis testing in statistics:



Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Testing (3/8)

- Principles of the hypothesis testing in statistics:
  - The defendant is assumed innocent until proven otherwise.
  - The defendant doesn't need to prove his/her innocence.
  - The main objective of the trial is to prove the guiltiness of the defendant.
  - The judge delivers the verdict based on the presented evidence.

Hypothesis Testing (4/8)

Principles of the hypothesis testing in statistics:

- Null hypothesis $H0$: what is assumed until proven otherwise.
- Alternative hypothesis $H1$: what needs to be proven.
- Test statistic: a quantity used as the evidence; calculated from the sample.
- p-value: probability of observing the current test statistic or more extreme one assuming that the null hypothesis is true.

  a) If the p-value is small: weakens the assumption of the null hypothesis.

  b) If the p-value is large: strengthens the assumption of the null hypothesis.

Hypothesis Testing (5/8)

Principles of the hypothesis testing in statistics:

- Significance level α:

  a) It's the reference probability used when deciding whether the p-value is small enough or not.

  b) If p-value $\geq \alpha$, the H_0 is maintained.

  If p-value $< \alpha$, the H_0 is rejected in favor of the H_1.

  c) This can also be interpreted as the maximum probability of rejecting the H_0 even when it's true.

  d) $1-\alpha$ is the probability that the true H_0 is retained.

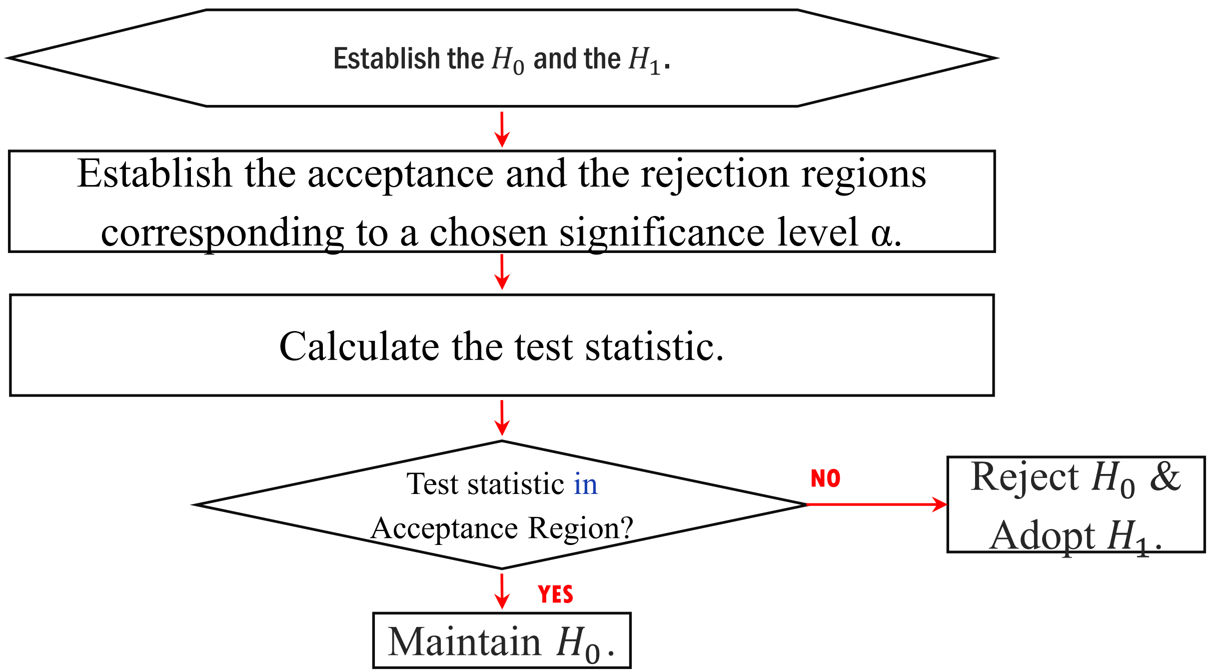
- Statistical power: probability of rejecting the false H_0 in favor of H_1.

Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Testing (6/8)

Principles of the hypothesis testing in statistics:

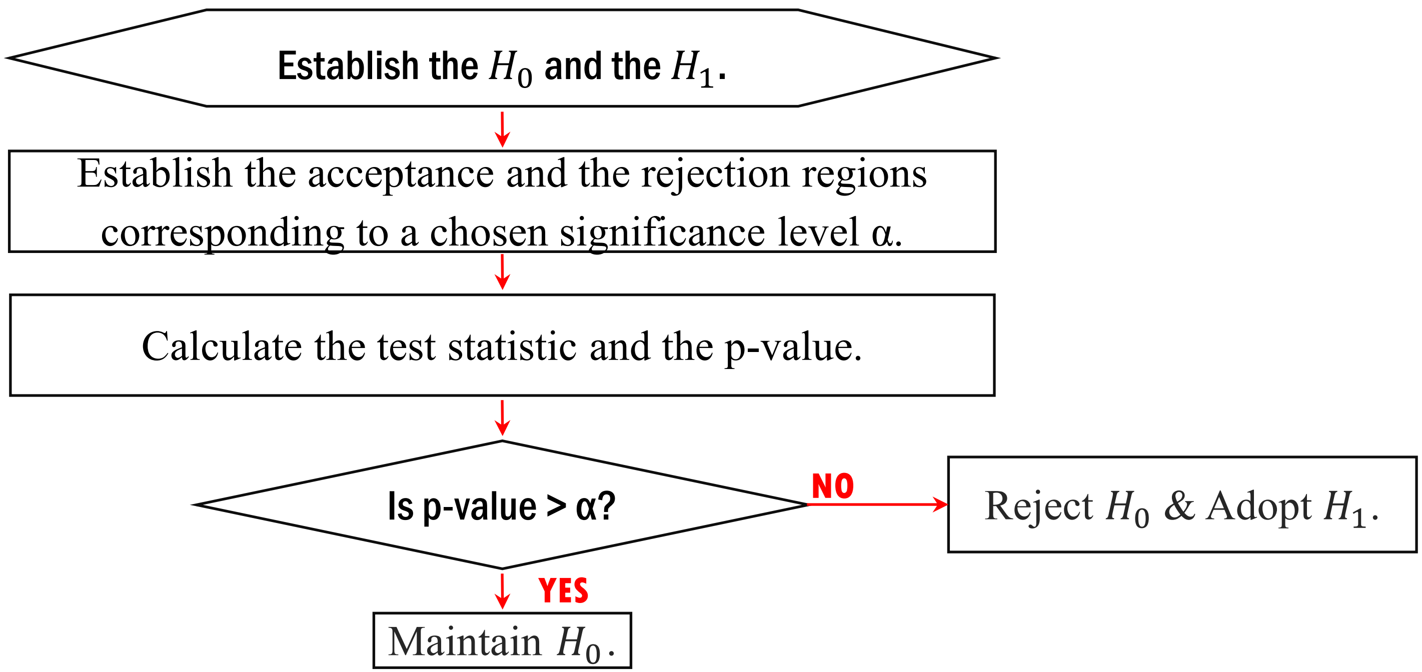
- Procedure using the test statistic:



Hypothesis Testing (7/8)

Principles of the hypothesis testing in statistics:

- Procedure using the p-value:



Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Testing (8/8)

Principles of the hypothesis testing in statistics:

ACTUALLY / TEST RESULT	H_0 IS TRUE	H_0 IS FALSE
H_0 RETAINED	**Correct Decision.** Probability $= 1 - \alpha$	**Type 2 error.** Probability$= \beta$
H_0 REJECTED & H_1 ADOPTED	**Type 1 error.** Probability $= \alpha$	**Correct Decision.** Probability $= 1 - \beta$

Chapter 4.

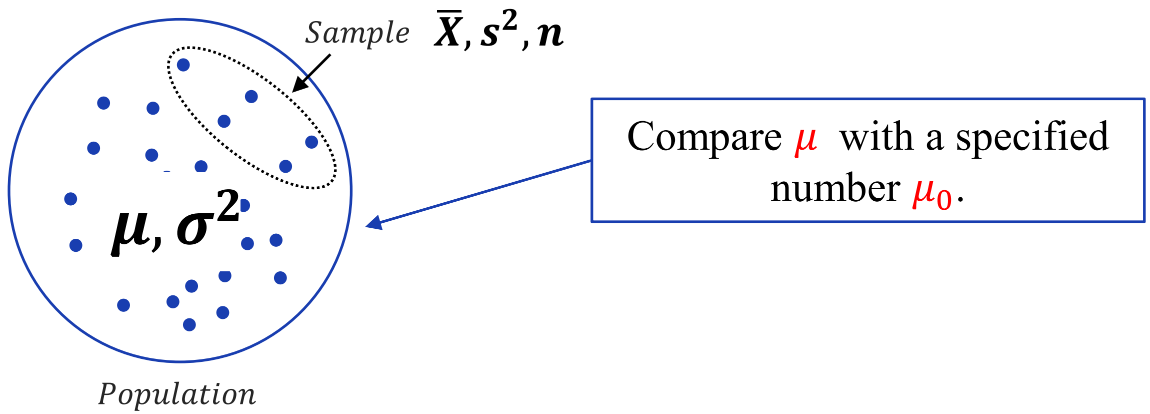
Probability and Statistics

Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Test of the Means (1/8)

One sample t-test:

- There is one population and one sample.



- Student-t distribution is used to interpret the test statistic calculated as following:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Hypothesis Test of the Means (2/8)

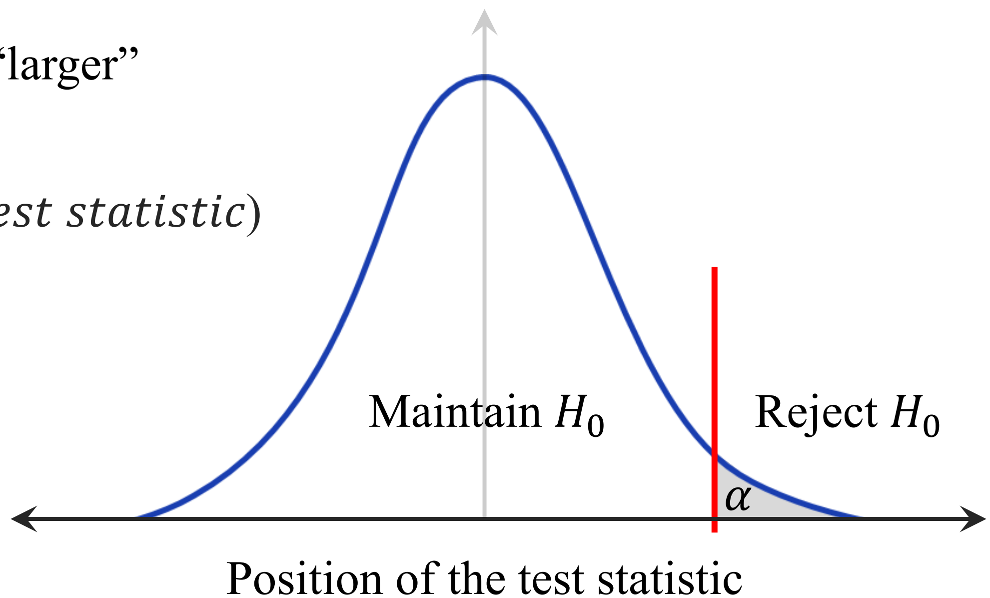
One sample t-test:

- Right tail test:

  $H_0 \; : \; \mu \leq \mu_0$

  $H_1 \; : \; \mu > \mu_0$ "larger"

- p-value = $P(X > test\ statistic)$



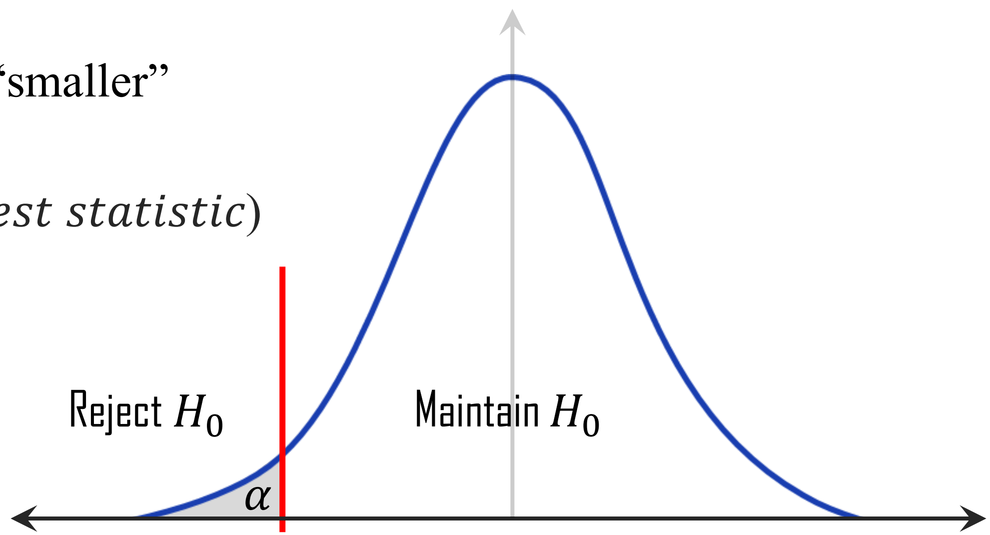
Hypothesis Test of the Means (3/8)

- One sample t-test:
  - Left tail test:

    $H_0 \ : \ \mu \geq \mu_0$

    $H_1 \ : \ \mu < \mu_0$ "smaller"

  - p-value = $P(X < test\ statistic)$



Position of the test statistic

Together for Tomorrow! Enabling People
Education for Future Generations

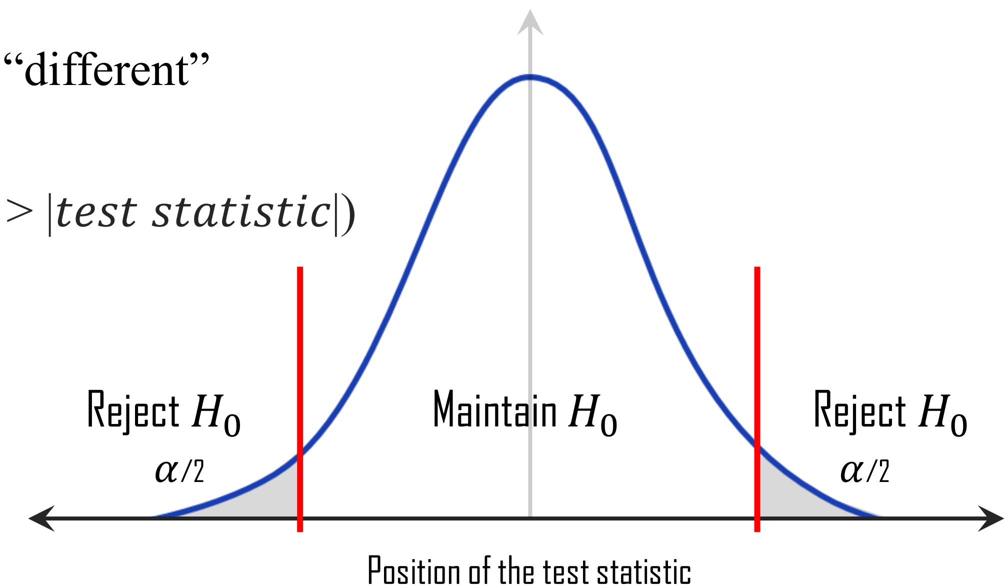
Hypothesis Test of the Means (4/8)

- One sample t-test:
  - Two tail test:

$$H_0 \; : \; \mu = \mu_0$$

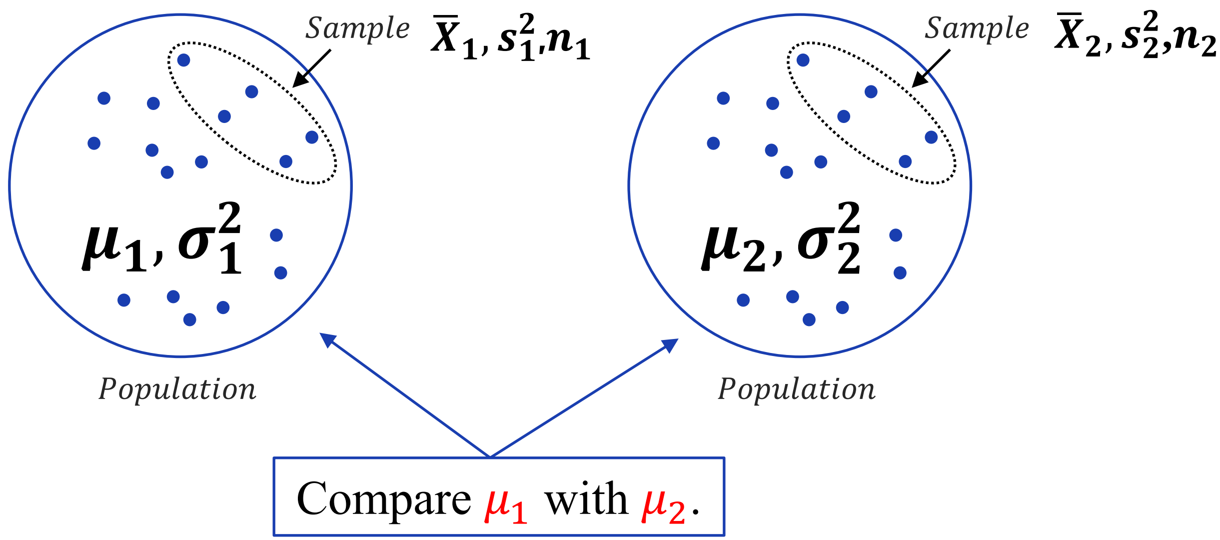
$$H_1 \; : \; \mu \neq \mu_0 \qquad \text{“different”}$$

  - p-value $= 2 \times P(X > |test\ statistic|)$



Hypothesis Test of the Means (5/8)

- Independent two sample t-test:
  - There are two populations and two samples.



Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Test of the Means (6/8)

Independent two sample t-test:

- Right tail test:

$$H_0 : \mu_1 - \mu_2 \leq 0 \quad \Leftrightarrow \quad \mu_1 \leq \mu_2$$

$$H_1 : \mu_1 - \mu_2 > 0 \quad \Leftrightarrow \quad \mu_1 > \mu_2$$

- Left tail test:

$$H_0 : \mu_1 - \mu_2 \geq 0 \quad \Leftrightarrow \quad \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 - \mu_2 < 0 \quad \Leftrightarrow \quad \mu_1 < \mu_2$$

- Two tail test:

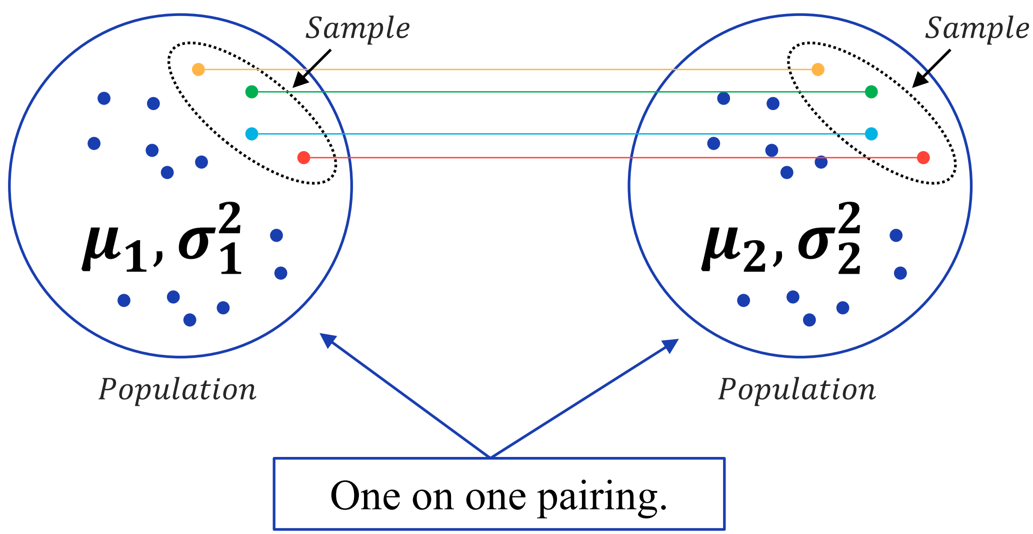
$$H_0 : \mu_1 - \mu_2 = 0 \quad \Leftrightarrow \quad \mu_1 = \mu_2$$

$$H_1 : \mu_1 - \mu_2 \neq 0 \quad \Leftrightarrow \quad \mu_1 \neq \mu_2$$

You should consider two cases:
1) Equal variances.
2) Unequal variances.

Hypothesis Test of the Means (7/8)

- Paired two sample t-test:
  - There are two populations and two samples. There is “one on one” pairing.



Ex) Change in the blood pressure of the same test subjects before and after taking a new drug.

Hypothesis Test of the Means (8/8)

Analysis of Variance (ANOVA):

- So far with the t-test, we had one or two groups (samples).
- ANOVA can detect differences in the means of two or more groups.

Null hypothesis $\boldsymbol{H_0}: \mu 1 = \mu 2 = \mu 3 = \ldots$

Alternative hypothesis $\boldsymbol{H_1}$: There is at least one case where $\mu i \neq \mu j$

- Assumptions: 1) The distribution of the data is Normal.
  2) The group variances are the same.
  3) The groups are independent from each other.
- F distribution is used to calculate the p-value.

Coding Exercise #0307

Follow practice steps on 'ex_0307.ipynb' file.

Hypothesis Test of the Frequencies (1/4)

Chi-squared test for one way table:

- One way table or "frequency table" summarizes a categorical variable.
- Compares the observed frequencies with a given model (expected frequencies).

  Null hypothesis $\boldsymbol{H_0}$: The observed frequency table and the expected model agree.

  Alternative hypothesis $\boldsymbol{H_1}$: The observed frequency table and the expected model are different.
- Also called the "Goodness of fit test".

Hypothesis Test of the Frequencies (2/4)

- Chi-squared test for one way table:
  - The test statistic is:

$$test\ statistic = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

  - E_i are the expected frequencies and O_i are the observed frequencies.
  - k is the number of categories or types.
  - The test statistic follows the Chi-square distribution of degree of freedom = $k - 1$.

Hypothesis Test of the Frequencies (3/4)

Chi-squared test for two way table:

- A contingency table summarizes two categorical variables.

  Ex) Confusion matrix (machine learning).
- Uses the frequencies to test the existence of relationship between two categorical variables.

  Null hypothesis H_0 : The categorical variables are independent.

  Alternative hypothesis H_1 : The categorical variables are not independent.
- Also called “Independence test”.

Hypothesis Test of the Frequencies (4/4)

- Chi-squared test for two way table:
  - The test statistic is:

$$test\ statistic = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}}$$

  - E_i are the expected frequencies and O_i are the observed frequencies.
  - r is the number of rows and c is the number of columns in the two way table.
  - The test statistic follows the Chi-square distribution of degree of freedom = $(r-1) \times (c-1)$.

Together for Tomorrow!
Enabling People
Education for Future Generations

Coding Exercise #0308

Follow practice steps on 'ex_0308.ipynb' file.

Together for Tomorrow! Enabling People
Education for Future Generations

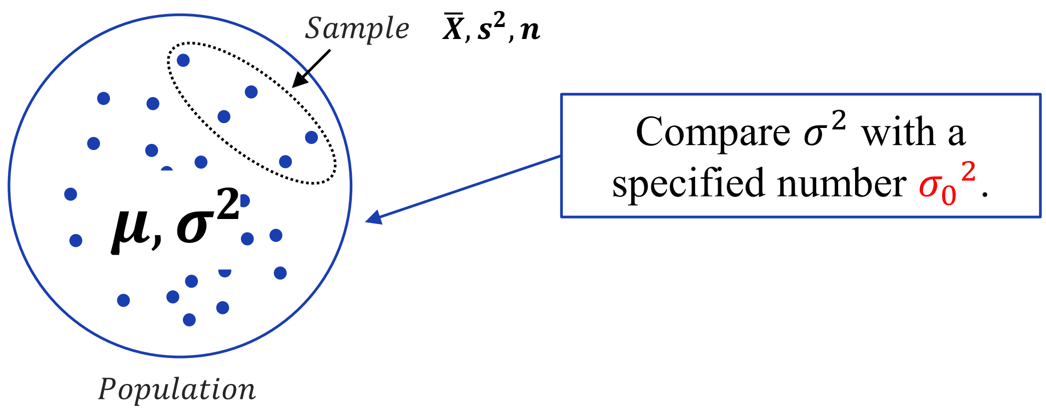
Coding Exercise #0309

Follow practice steps on 'ex_0309.ipynb' file.

Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Test of the Variances (1/4)

- One sample t-test:
  - There is one population and one sample.



Hypothesis Test of the Variances (2/4)

Chi-squared test of variance:

- There are left tail test, right tail test and two tail test. ⇐ Just like in t-test.
- The test statistic is calculated as: (n= sample size)

$$test\ statistic = \frac{(n-1)S^2}{\sigma_0^2}$$

- The test statistic follows the Chi-square distribution of degree of freedom = $n - 1$.

Together for Tomorrow! Enabling People
Education for Future Generations

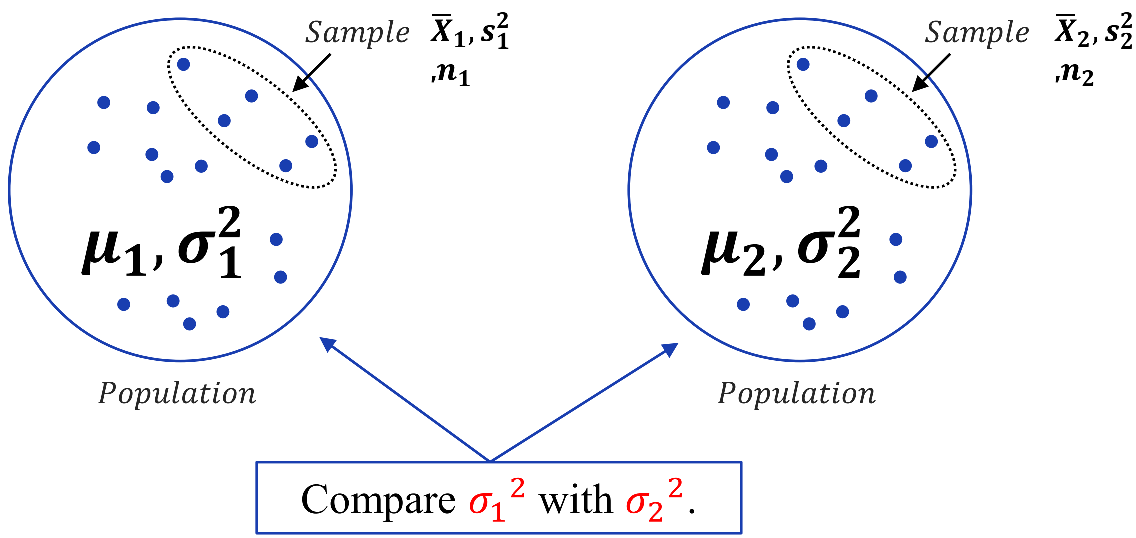
Coding Exercise #0310

Follow practice steps on 'ex_0310.ipynb' file.

Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Test of the Variances (3/4)

- F-test of variance ratio:
  - There are two populations and two samples.



Hypothesis Test of the Variances (4/4)

F-test of variance ratio:

- There are left tail test, right tail test and two tail test.
- The test statistic is calculated as a ratio of the sample variances:

$$test\ statistic = \frac{S_1^2}{S_2^2}$$

- The test statistic follows the F distribution $F(n_1 - 1, n_2 - 1)$. Here n_1 and n_2 are the sample sizes.

Hypothesis Test Summary

- Here, this summarizes the hypothesis tests we have covered so far:

HYPOTHESIS TEST	PROBABILITY DENSITY DISTRIBUTION
One sample t-test, Independent two sample t-test, Paired sample t-test.	Student-t
ANOVA	F
Chi-squared test of one way table, Chi-squared test of two way table.	Chi-square
Chi-squared test of variance.	Chi-square
F-test of variance ratio.	F

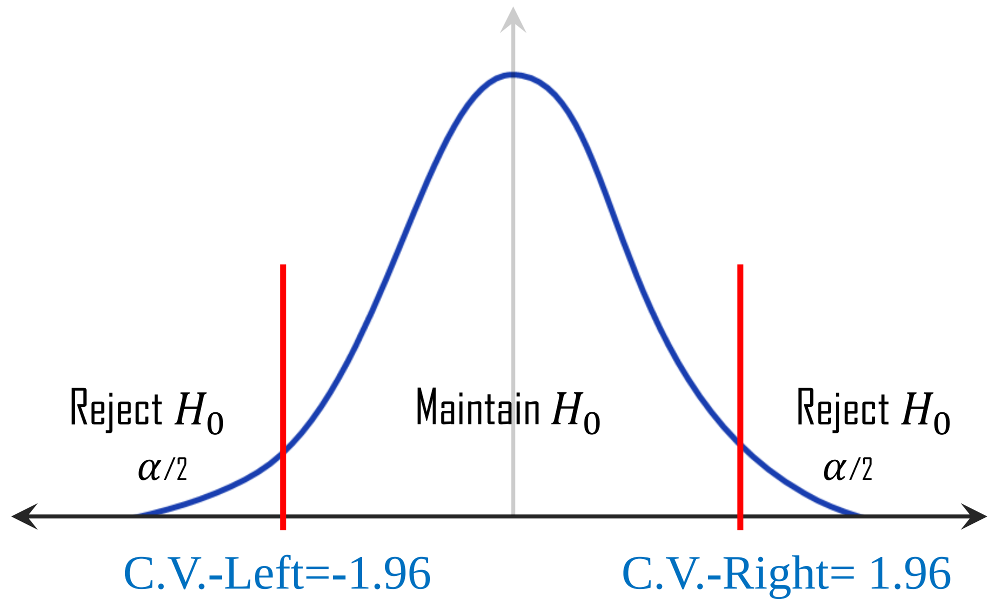
Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Test Example

The average height of Pakistani men is 167 cm with a standard deviation of 15. A researcher believes that this value has changed. The researcher decides to test the height of 75 random male adults. The average height of the sample is 172 cm. Is there enough evidence to suggest that the average height of Pakistani men has changed.

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion



Hypothesis Test Example

- The average height of Pakistani men is 167 cm with a standard deviation of 15. A researcher believes that this value has changed. The researcher decides to test the height of 75 random male adults. The average height of the sample is 172 cm. Is there enough evidence to suggest that the average height of Pakistani men has changed.

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion

5 Steps to be kept in mind:

1. H0: $\mu=167$
   H1: $\mu \neq 167$
   (Two-Tailed Test)

2. $\alpha = 5\%$

Hypothesis Test Example

- The average height of Pakistani men is 167 cm with a standard deviation of 15. A researcher believes that this value has changed. The researcher decides to test the height of 75 random male adults. The average height of the sample is 172 cm. Is there enough evidence to suggest that the average height of Pakistani men has changed.

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion

5 Steps to be kept in mind:

3. C.V. of the rejection region:
   C.V.-Left= -1.96
   C.V.-Right= +1.96

$$4.\ \text{test statistic} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

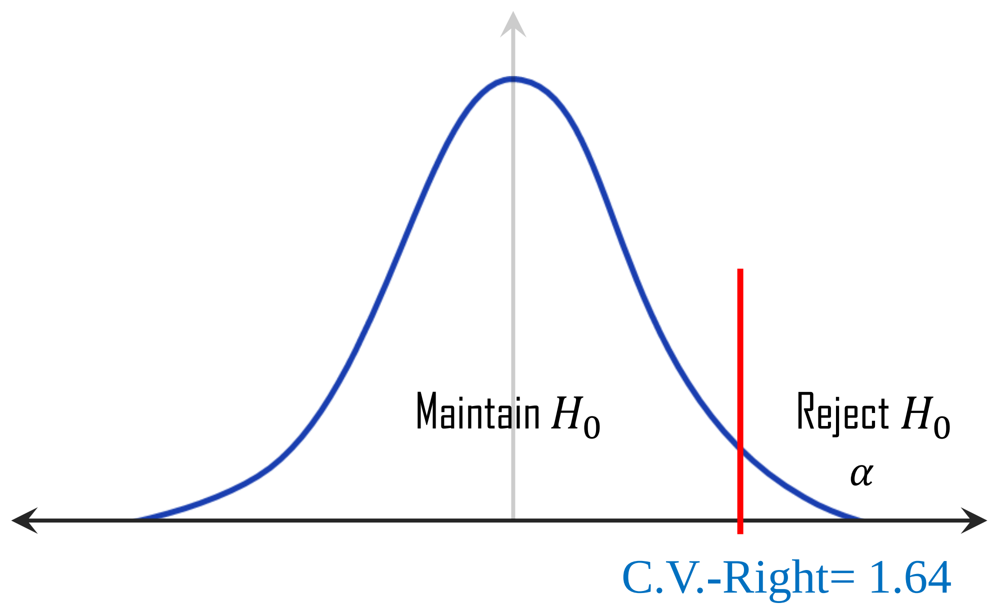
Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Test Example

The average height of Pakistani men is 167 cm. A researcher believes that this value is higher. The researcher decides to test the height of 75 random male adults. The average height of the sample is 172 cm with a standard deviation of 12. Is there enough evidence to suggest that the average height of Pakistani men has changed.

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion



Hypothesis Test Example

The average height of Pakistani men is 167 cm with a standard deviation of 15. A researcher believes that this value has changed. The researcher decides to test the height of 75 random male adults. The average height of the sample is 172 cm. Is there enough evidence to suggest that the average height of Pakistani men has changed.

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion

5 Steps to be kept in mind:

1. H0: $\mu=167$
   H1: $\mu< 167$
   (One-Tailed Test)

2. $\alpha = 5\%$

Hypothesis Test Example

- The average height of Pakistani men is 167 cm with a standard deviation of 15. A researcher believes that this value is higher. The researcher decides to test the height of 75 random male adults. The average height of the sample is 172 cm. Is there enough evidence to suggest that the average height of Pakistani men has changed.

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion

5 Steps to be kept in mind:

3. C.V. of the rejection region:
   C.V.-Right= +1.64

4. test statistic $= \dfrac{\bar{X} - \mu}{\bar{s}/\sqrt{n}}$

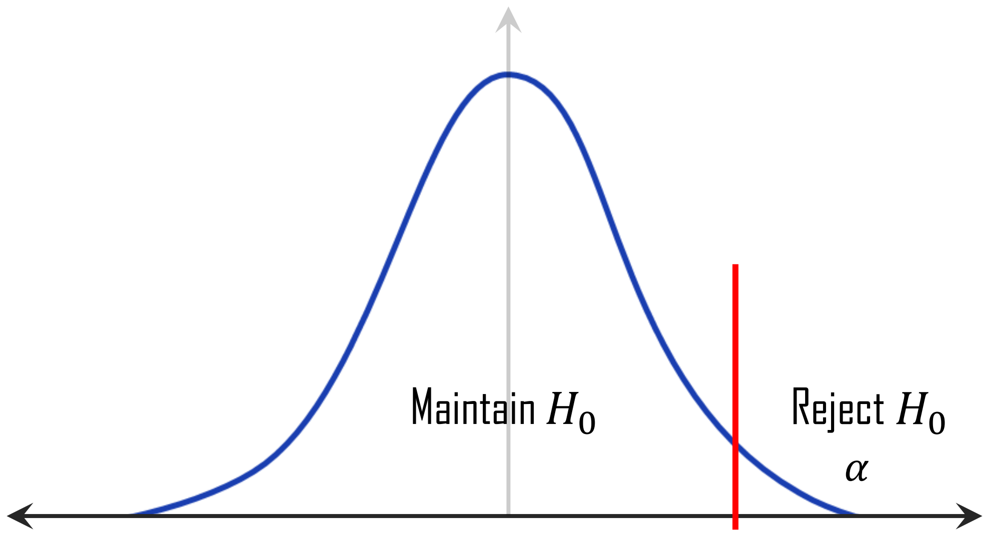
Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Test Example

The average height of Pakistani men is 167 cm with a standard deviation of 15. A researcher believes that this value is higher. The researcher decides to test the height of 75 random male adults. The average height of the sample is 172 cm. Is there enough evidence to suggest that the average height of Pakistani men has changed.

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion



C.V.-Right= 1.64

Hypothesis Test Example

The average height of Pakistani men is 167 cm with a standard deviation of 15. A researcher believes that this value has changed. The researcher decides to test the height of 75 random male adults. The average height of the sample is 172 cm. Is there enough evidence to suggest that the average height of Pakistani men has changed.

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion

5 Steps to be kept in mind:

1. H0: μ=167
   H1: μ< 167
   (One-Tailed Test)

2. $\alpha = 5\%$

Hypothesis Test Example

The average height of Pakistani men is 167 cm. A researcher believes that this value has increased. The researcher decides to test the height of 75 random male adults. The average height of the sample is 172 cm with a standard deviation of 15. Is there enough evidence to suggest that the average height of Pakistani men has changed.

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion

5 Steps to be kept in mind:

3. C.V. of the rejection region:
   C.V.-Right= +1.64

$$4.\ \text{test statistic} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Test Example

The Centers for Disease Control (CDC) reported on trends in weight, height and body mass index from the 1960's through 2002.[1] The general trend was that Americans were much heavier and slightly taller in 2002 as compared to 1960; both men and women gained approximately 24 pounds, on average, between 1960 and 2002. In 2002, the mean weight for men was reported at 191 pounds. Suppose that an investigator hypothesizes that weights are even higher in 2006 (i.e., that the trend continued over the subsequent 4 years).

In order to test the hypotheses, we select a random sample of American males in 2006 and measure their weights. Suppose we have resources available to recruit n=100 men into our sample. We weigh each participant and compute summary statistics on the sample data. Suppose in the sample we determine the following:

$n=100$ $\bar{X}= 197.1$ $s=25.6$

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion

Together for Tomorrow! Enabling People
Education for Future Generations

Hypothesis Test Example

The Centers for Disease Control (CDC) reported on trends in weight, height and body mass index from the 1960's through 2002.[1] The general trend was that Americans were much heavier and slightly taller in 2002 as compared to 1960; both men and women gained approximately 24 pounds, on average, between 1960 and 2002. In 2002, the mean weight for men was reported at 191 pounds. Suppose that an investigator hypothesizes that weights are even higher in 2006 (i.e., that the trend continued over the subsequent 4 years).

In order to test the hypotheses, we select a random sample of American males in 2006 and measure their weights. Suppose we have resources available to recruit n=100 men into our sample. We weigh each participant and compute summary statistics on the sample data. Suppose in the sample we determine the following:

n=100 $\bar{X}$= 192.1 s=25.6

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion

Together for Tomorrow!
Enabling People
Education for Future Generations

Hypothesis Test Example

Pakistanis believe that the chance of the Indians winning the coin toss is greater than 50%. In a random sample of 200 coin tosses, the Indians won 118 times. Is there enough evidence to suggest the Indians are cheaters?

If pn>5 and (1-p)n>5, then Z-Test

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion

Hypothesis Test Example

Pakistanis believe that the chance of the Indians winning the coin toss is greater than 50%. In a random sample of 200 coin tosses, the Indians won 118 times. Is there enough evidence to suggest the Indians are cheaters?

If pn>5 and (1-p)n>5, then Z-Test

$$\text{test statistic} = \frac{\bar{P} - P}{\sqrt{\frac{P(1-P)}{n}}}$$

5 Steps to be kept in mind:

1. Establish H0 and H1
2. Choose Level of Significance (α)
3. Find Critical Values
4. Find Test Statistic (z or t)
5. Draw your conclusion

Answer the following questions by providing Python code:

In a factory there are two packaging machines. Output samples are drawn from each machine.

n1=15n1=15, x¯1=5.0592kgx¯1=5.0592kg, s21=0.1130kg2s12=0.1130kg2
n2=12n2=12, x¯2=4.9808kgx¯2=4.9808kg, s22=0.0152kg2s22=0.0152kg2
Test whether there is a significant difference in the variances (95%).

```
n1 = 15
ssq1 = 0.1130
n2 = 12
ssq2 = 0.0152


test_stat = ssq1/ssq2

2*(1 - st.f.cdf(test_stat, n1-1, n2-1))
```

Together for Tomorrow! Enabling People
Education for Future Generations

Coding Exercise #0311

Follow practice steps on 'ex_0311.ipynb' file.

Coding Exercise #0312

Follow practice steps on 'ex_0312.ipynb' file.

End of chapter Quiz

Quiz #0301 ~ #0307

Duration : 5 Hours

End of Document

SAMSUNG

Together for Tomorrow!

Enabling People

Education for Future Generations