# Correlation Analysis
# Hypothesis Testing

LEC 4


**DR.ALAA MOHAMMED**

3RD LEVEL, $2^{ND}$ SEMESTER, BIOMEDICAL INSTRUMENTATION AND BIOMECHANICS BRANCHES, BIOMEDICAL ENG. DEPARTMENT

# Hypothesis Testing

## PART3

# Hypothesis testing

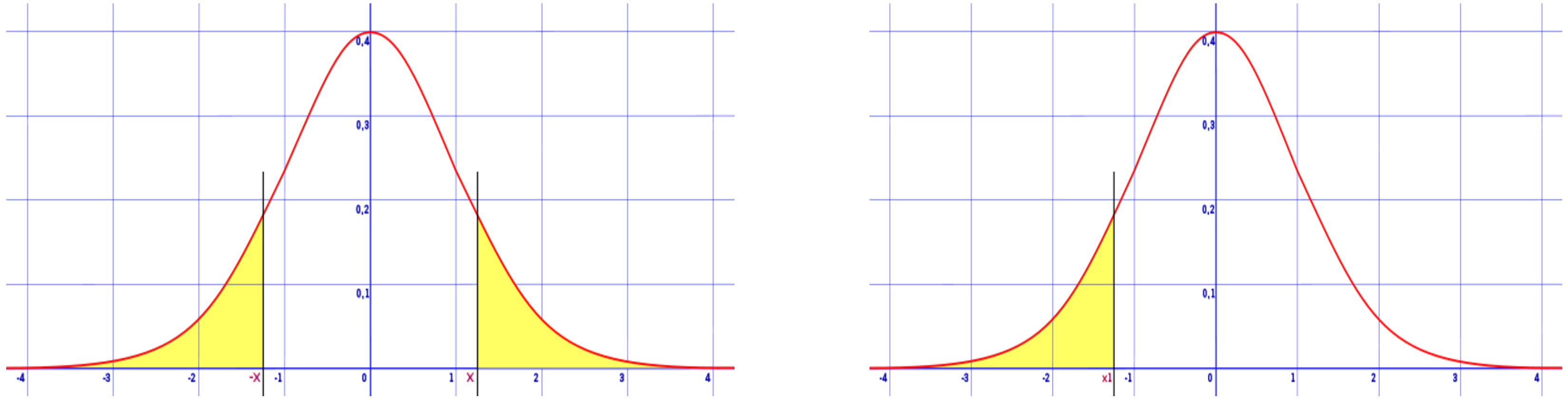
**Hypothesis testing**: is the statement or an assumption about relationship between two or more variables.

Based on sample evidence and probability theory, Hypothesis is the procedure which enable researcher to decide whether to accept or reject hypothesis.

A hypothesis test can also be used to determine whether the sample correlation coefficient ***r*** provides enough evidence to conclude that the population correlation coefficient ***ρ*** is significant at a specified level of significance.

Hypothesis can be categorized into null hypothesis ($H_0$) and alternative hypothesis ($H_a$)

A hypothesis test can be one tailed or two tailed




(a) (b)

FIGURE (1): (A) TWO-TAILED TEST, (B) ONE-TAILED TEST

# The hypothesis test for the Correlation Coefficient $\rho$

**Procedure:**

1. State the null and alternative hypothesis (State $H_0$ and $H_a$).
2. Specify the level of significance (Identify $\alpha$).
3. Identify the degrees of freedom (d.f. = $n$ – 2).
4. Determine the critical value(s) and rejection region(s) (**t-Distribution Table**).

# The hypothesis test for the Correlation Coefficient $\rho$

**Procedure (cont.):**

5. Find the standardized test statistic

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

6. Make a decision to reject or fail to reject the null hypothesis (If $t$ is in the rejection region, reject $H_0$. Otherwise fail to reject $H_0$).

7. Interpret the decision in the context of the original claim

# Hypothesis testing for ρ

1. State the null hypothesis ($H_0$) and the alternative hypothesis ($H_a$)

A hypothesis test can be one tailed or two tailed

$$\begin{cases} H_0\colon \rho \geq 0 & \text{(no significant negative correlation)} \\ H_a\colon \rho < 0 & \text{(significant negative correlation)} \end{cases}$$

Left-tailed test

$$\begin{cases} H_0\colon \rho \leq 0 & \text{(no significant positive correlation)} \\ H_a\colon \rho > 0 & \text{(significant positive correlation)} \end{cases}$$

Right-tailed test

$$\begin{cases} H_0\colon \rho = 0 & \text{(no significant correlation)} \\ H_a\colon \rho \neq 0 & \text{(significant correlation)} \end{cases}$$

Two-tailed test

# Hypothesis Testing for $\rho$

2. *Chose the levels of significance* $\alpha$ , and the sample size $n$

**The $t$-Test for the Correlation Coefficient**

A ***t*-test** can be used to test whether the correlation between two variables is significant. The **test statistic** is $r$ and the **standardized test statistic is $t$**

$$t = \frac{r}{\sigma_r} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

follows a $t$-distribution with $n - 2$ degrees of freedom.

In this text, only two-tailed hypothesis tests for $\rho$ are considered.

# Example 4

The following data represents the number of hours 12 different students watched television during the weekend and the scores of each student who took a test the following Monday.

| Hours, $x$ | 0 | 1 | 2 | 3 | 3 | 5 | 5 | 5 | 6 | 7 | 7 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Test score, $y$ | 96 | 85 | 82 | 74 | 95 | 68 | 76 | 84 | 58 | 65 | 75 | 50 |

The correlation coefficient $r \approx -0.831$.

Test the significance of this correlation coefficient significant at $\alpha = 0.01$?

# Example 4: Hypothesis Testing for $\rho$

$H_0$: $\rho = 0$   (no correlation) , $H_a$: $\rho \neq 0$  (significant correlation)

The level of significance is $\alpha = 0.01$.

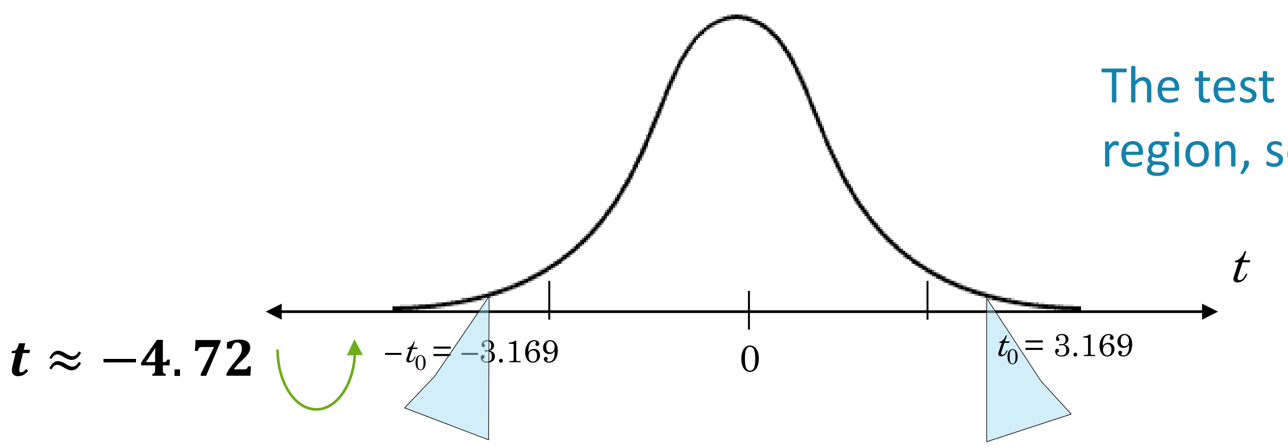
Degrees of freedom are d.f. = 12 – 2 = 10.

The critical values are $-t_0 = -3.169$ and $t_0 = 3.169$ (two tailed)

The standardized test statistic is

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{-0.831}{\sqrt{\frac{1-(-0.831)^2}{12-2}}} \approx -4.72.$$

# Example 4: Hypothesis Testing for $\rho$




The test statistic falls in the rejection region, so $H_0$ is rejected

At the 1% level of significance, there is enough evidence to conclude that there is a significant linear correlation between the number of hours of TV watched over the weekend and the test scores on Monday morning.

## *t* Distribution: Critical Values of *t*

| | | *Significance level* | | | | | |
|---|---|---|---|---|---|---|---|
| *Degrees of freedom* | *Two-tailed test:* | 10% | 5% | 2% | 1% | 0.2% | 0.1% |
| | *One-tailed test:* | 5% | 2.5% | 1% | 0.5% | 0.1% | 0.05% |
| **1** | | 6.314 | 12.706 | 31.821 | 63.657 | 318.309 | 636.619 |
| **2** | | 2.920 | 4.303 | 6.965 | 9.925 | 22.327 | 31.599 |
| **3** | | 2.353 | 3.182 | 4.541 | 5.841 | 10.215 | 12.924 |
| **4** | | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 | 8.610 |
| **5** | | 2.015 | 2.571 | 3.365 | 4.032 | 5.893 | 6.869 |
| **6** | | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 | 5.959 |
| **7** | | 1.894 | 2.365 | 2.998 | 3.499 | 4.785 | 5.408 |
| **8** | | 1.860 | 2.306 | 2.896 | 3.355 | 4.501 | 5.041 |
| **9** | | 1.833 | 2.262 | 2.821 | 3.250 | 4.297 | 4.781 |
| **10** | | 1.812 | 2.228 | 2.764 | 3.169 | 4.144 | 4.587 |
| **11** | | 1.796 | 2.201 | 2.718 | 3.106 | 4.025 | 4.437 |
| **12** | | 1.782 | 2.179 | 2.681 | 3.055 | 3.930 | 4.318 |
| **13** | | 1.771 | 2.160 | 2.650 | 3.012 | 3.852 | 4.221 |
| **14** | | 1.761 | 2.145 | 2.624 | 2.977 | 3.787 | 4.140 |
| **15** | | 1.753 | 2.131 | 2.602 | 2.947 | 3.733 | 4.073 |
| **16** | | 1.746 | 2.120 | 2.583 | 2.921 | 3.686 | 4.015 |
| **17** | | 1.740 | 2.110 | 2.567 | 2.898 | 3.646 | 3.965 |
| **18** | | 1.734 | 2.101 | 2.552 | 2.878 | 3.610 | 3.922 |
| **19** | | 1.729 | 2.093 | 2.539 | 2.861 | 3.579 | 3.883 |
| **20** | | 1.725 | 2.086 | 2.528 | 2.845 | 3.552 | 3.850 |