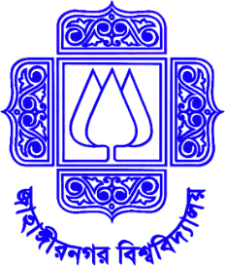
**Assignment 04**

**Title: Assignment of regression analysis using excel and minitab**

*Course title: Statistics*

*Course code: CSE-251*

*2nd Year 2nd Semester 2022*



**Date of Submission**: 10/01/2024

###### 

###### **Submitted to-**

***Sabina Yeasmin***

*Lecturer*

*Department of Statistics and Data Science, Jahangirnagar University*

*Savar, Dhaka-1342*

|  |  |  |  |
| --- | --- | --- | --- |
| **Sl** | Class Roll | Exam Roll | Name |
| 01 | 379 | 210899 | Shah sultan |

**Question :**

**Problem 1:**

The article “Size effect in Shear Strength of Large Beams-Behavior and Finite

Element Modelling” (Mag. Of Concrete Res., 2005:497-509) reported on a study of various

characteristics of large reinforced concrete deep and shallow beams tested until failure. Consider

the following data on x = cube strength and y = cylinder strength (both in MPa):

A table with numbers and a few black text

Description automatically generated with medium confidence

a) Estimate the correlation coefficient.

b) Fit a regression model relating cube strength to cylinder strength.

c) Test for significance of regression.

d) Find a 95% confidence interval for the slope.

e) What will be the cylinder strength if the cube strength is 100 MPa?

f) What percentage of the variability in cylinder strength is explained by the cube strength?

**Using MS Excel and minitab for Solution :**



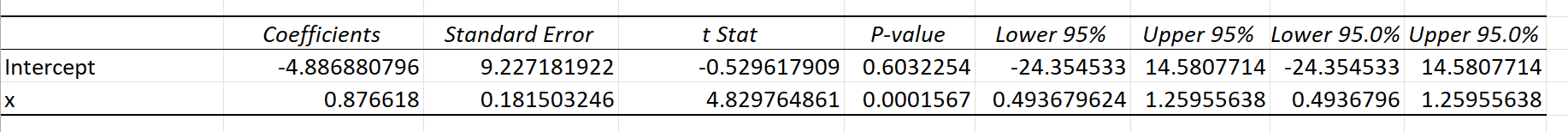
A white sheet with black text

Description automatically generated

The correlation coefficient is, rxy= 0.760553956.

**Comment :**

Here, the correlation coefficient contains 0.6<= rxy <= 0.8. So, we can say that there is a strong positive linear correlation between cube strength and cylinder strength of concrete beams.

1. 

A screenshot of a computer

Description automatically generated

A close-up of a white background

Description automatically generated

**Comment :**

For a unit change of cube strength, cylindrical strength will be increased by 0.877 MPa.

In the absence of cube strength, cylindrical strength will be zero as strength cannot be a negative value.



Let us define two hypotheses.

H0 : B1 = 0, against alternative hypothesis.

HA : B1 != 0.

We assume that significance At 5% significance level, we will test the hypothesis.

Here, a = 5% = 0.05

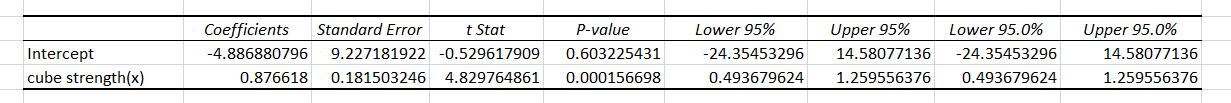
A close up of numbers

Description automatically generated

Significance F = 0.000156698 < a .

So, null hypothesis rejected indicating the significant relationship between cube strength and cylindrical strength.





**Comment :**

The 95% confidence interval for the coefficient of cylindrical strength is (0.493679624, 1.259556376). That implies for a unit change in cube strength, we are 95% confident that cylindrical strength change will be in the interval (0.37160752, 0.948106016).

Similarly, the 95% confidence interval for the coefficient of intercept is (-24.35453296, 14.58077136). That implies if the cube strength is zero, we are 95% confident that cylindrical strength will be in the interval (-24.35453296, 14.58077136).

1. cylindrical strength = y

cube strength = x

The regression equation is:

y= -4.89 + 0.877(x)

Given, cube strength is 100 MPa, So, cylindrical strength will be :

y = -4.89 + 0.877 (100) = 82.81 MPa.



A white sheet with black text

Description automatically generated

Here, coefficient of determination , R2 = 0.760553956 . It interprets that on an average, 76.06% variation in cylindrical strength can be explained by the variation in cube strength of concrete beams.

**Problem 2:**

Soil and sediment adsorption, the extent to which chemicals collect in a condensed

form of the surface, is an important characteristic influencing the effectiveness of pesticides and

various agricultural chemicals. The article “Adsorption of Phosphate, Arsenate, Methane

arsonate, and Cacodylate by Lake and Stream Sediments: Comparisons with Soils” (J. of

Environ. Qual., 1984: 499-504) gives the accompanying data on y = phosphate adsorption index,

x1 = amount of extractable iron, and x2 = amount of extractable aluminum.

|  |  |  |  |
| --- | --- | --- | --- |
| observation | x1 | x2 | y |
| 1 | 61 | 13 | 4 |
| 2 | 175 | 21 | 18 |
| 3 | 111 | 24 | 14 |
| 4 | 124 | 23 | 18 |
| 5 | 130 | 64 | 26 |
| 6 | 173 | 38 | 26 |
| 7 | 169 | 33 | 21 |
| 8 | 169 | 61 | 30 |
| 9 | 160 | 39 | 28 |
| 10 | 244 | 71 | 36 |
| 11 | 257 | 112 | 65 |
| 12 | 333 | 88 | 62 |
| 13 | 199 | 54 | 40 |

a) Construct a scatter plot for individual predictor variable with the response variable.

b) Fit a regression model.

c) Test for significance of regression.

d) Compute the t statistics for each independent variable. What conclusions can you draw?

e) Construct a 95% confidence interval estimate for .

f) What will be the cylinder strength if the cube strength is 100 MPa?

g) What percentage of the variability in y is explained by and ?

**Using MS Excel and Minitab for Solution:**



A graph with blue dots and numbers

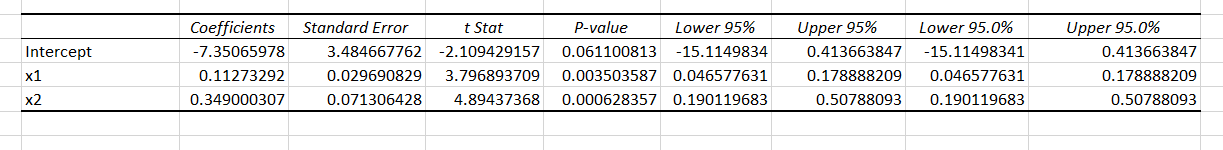
Description automatically generated

**Figure 1 :** Scatter plot for amount of extractable iron vs phosphate absorption index.

A graph with blue dots

Description automatically generated

**Figure 2 :** Scatter plot for amount of extractable aluminum vs phosphate absorption index.



A screenshot of a computer

Description automatically generated

A close-up of a white background

Description automatically generated

**Comment :**

For a unit change of the amount of extractable iron, phosphate absorption index will be changed by 0.1127 unit.

For a unit change of the amount of extractable aluminum, phosphate absorption index will be changed by 0.3490 unit.

In the absence of the amount of extractable iron and extractable aluminum, phosphate absorption index will be 0 (-7.35 unit) as absorption cannot be a negative value.

Let us define two hypotheses.

H0 : B1 = 0, B2 = 0 against alternative hypothesis.

HA : B1 != 0, B2 != 0.

We assume that significance At 5% significance level, we will test the hypothesis.

Here, a = 5% = 0.05

A close-up of a graph

Description automatically generated

Significance F = 3.63428E-07 < a .

Also, p- value for x1 = 0.003503587 < a and p- value for x2 = 0.000628357 < a.

So, null hypothesis rejected indicating the significant relationship between phosphate absorption index and amount of extractable iron and aluminum.

Thus, change in the amount of extractable iron and aluminum causes significant change in the amount of phosphate absorption index.

A screenshot of a computer

Description automatically generated

t-stat value for the amount of extractable iron : 3.796893709

t-stat value for the amount of extractable aluminum : 4.89437368

t-stat value is used in hypothesis to assess the significance of the relationship between variables. For a significance level of 5% and degree of freedom (n-2) , the critical value of t is 1.796. here, both t-stat value is greater than the critical value, rejecting the null hypothesis.

So, we can conclude that there is a significant relationship between phosphate absorption index and amount of extractable iron and aluminum.



A table with numbers and letters

Description automatically generated

**Comment :**

The 95% confidence interval for the coefficient of the amount of extractable iron is (0.046577631, 0.178888209) .That implies for a unit change in the amount of extractable iron, we are 95% confident that phosphate absorption index change will be in the interval (0.046577631, 0.178888209).

The 95% confidence interval for the coefficient of the amount of extractable aluminum is (0.190119683, 0.50788093) .That implies for a unit change in the amount of extractable aluminum, we are 95% confident that phosphate absorption index change will be in the interval (0.190119683, 0.50788093) .

Similarly, the 95% confidence interval for the coefficient of intercept is (-15.1149834, 0.413663847). That implies if the amount of extractable aluminum and iron is zero, we are 95% confident that phosphate absorption index will be in the interval (-15.1149834, 0.413663847).



cylindrical strength = y

cube strength = x

The regression equation is:

y= -4.89 + 0.877(x)

Given, cube strength is 100 MPa, So, cylindrical strength will be:

y = -4.89 + 0.877 (100) = 82.81 MPa.



A screenshot of a table

Description automatically generated

Here, coefficient of determination , R2 = 0.9484672 . It interprets that on an average, 94.85% variation in phosphate absorption index can be explained by the variation in the amount of extractable iron and aluminum.