SIE 330R Homework, Spring 2023

**Agustin Espinoza**

**2/27/2023**

**HW 5 (Chapter 10)**

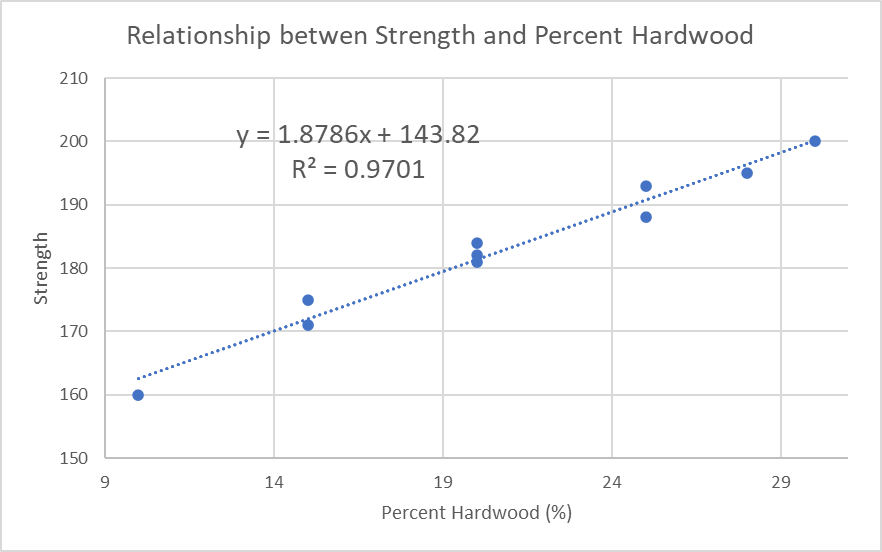
Homework must be readable! Do not just send in numbers or charts. You must explain the homework answers Preferred to receive homework in Word doc format with any excel or Minitab results pasted into word document. You may choose to use pdf which is also OK.

* Homework #5 10.1, 10.2, 10.5, 10.21
  1. The tensile strength of a paper product is related to the amount of hardwood in the pulp. Ten samples are produced in the pilot plant, and the data obtained are shown in the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| Strength | Percent Hardwood | Strength | Percent Hardwood |
| 160 | 10 | 181 | 20 |
| 171 | 15 | 188 | 25 |
| 175 | 15 | 193 | 25 |
| 182 | 20 | 195 | 28 |
| 184 | 20 | 200 | 30 |

1. Fit a linear regression model relating strength to percent hardwood.

The linear regression model relating the tensile strength of a paper product to the percentage of hardwood is illustrated below with Figure 10.1. The model displays the linear relationship between the Percent Hardwood explanatory variable and the Strength response variable.



*Figure 10.1 Relationship between Strength and Percent Hardwood.*

1. Test the model in part (a) for significance of regression.

To test the linear regression model in part (a), the correlation coefficient is calculated for the explanatory and response variables, equal to R, to evaluate the significance of the regression. The correlation coefficient (R) is 0.985. This R value signifies a strong linear relationship, 98.5%, between the Strength and the Percentage Hardwood variables of the regression model.

* 1. Plot the residuals from Problem 10.1 and comment on model adequacy.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |
| Intercept | 143.8243848 | 2.521528624 | 57.03857 | 9.91E-12 |
| Percent Hardwood | 1.878635347 | 0.116507923 | 16.12453 | 2.2E-07 |

The P-value from the residual output is less than 0.05. Therefore, we can conclude that the model adequacy is strong.

**10.5S.** The brake horsepower developed by an automobile engine on a dynamometer is thought to be a function of the engine speed in revolutions per minute (rpm), the road octane number of the fuel, and the engine compression. An experiment is run in the laboratory and the data that follow are collected.

|  |  |  |  |
| --- | --- | --- | --- |
| Brake Horsepower | rpm | Road Octane Number | Compression |
| 225 | 2000 | 90 | 100 |

|  |  |  |  |
| --- | --- | --- | --- |
| 212 | 1800 | 94 | 95 |
| 229 | 2400 | 88 | 110 |
| 222 | 1900 | 91 | 96 |
| 219 | 1600 | 86 | 100 |
| 278 | 2500 | 96 | 110 |
| 246 | 3000 | 94 | 98 |
| 237 | 3200 | 90 | 100 |
| 233 | 2800 | 88 | 105 |
| 224 | 3400 | 86 | 97 |
| 223 | 1800 | 90 | 100 |
| 230 | 2500 | 89 | 104 |

1. Fit a multiple linear regression model to the data.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Regression Statistics* | | | |  |  |  |
| Multiple R | | | 0.89806445 |  |  |  |
| R Square | | | 0.806519757 |  |  |  |
| Adjusted R Square | | | 0.733964665 |  |  |  |
| Standard Error | | | 8.812385189 |  |  |  |
| Observations | | | 12 |  |  |  |
|  | | |  |  |  |  |
| ANOVA | | |  |  |  |  |
|  | | | *df* | *SS* | *MS* | *F* |
| Regression | | | 3 | 2589.734938 | 863.2449794 | 11.11596364 |
| Residual | | | 8 | 621.2650618 | 77.65813272 |  |
| Total | | | 11 | 3211 |  |  |
| |  | | --- | | *Significance F* | | 0.003169979 | |  | | | |  |  |  |  |
|  | | | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |
| Intercept | | | -266.0312117 | 92.67365517 | -2.870623925 | 0.020808978 |
| rpm | | | 0.010713208 | 0.004483257 | 2.389603732 | 0.043882965 |
| Road Octane Number | | | 3.134806258 | 0.84443501 | 3.712312044 | 0.005936591 |
| Compression | | | 1.867409434 | 0.534525771 | 3.493581667 | 0.008155464 |
| *Lower 95%* | *Upper 95%* |
| -479.7370438 | -52.32537968 |
| 0.000374798 | 0.021051617 |
| 1.187535633 | 5.082076883 |
| 0.634790796 | 3.100028072 |

1. Test for significance of regression. What conclusions can you draw?

The significance of the multiple linear regression model is given by the R2 approximately equal to 80.7%. With this value we can conclude that the model has a moderately strong linear relationship between its variables.

1. Based on *t* tests, do you need all three regressor variables in the model?

Based on the t tests conducted for the regressor variables, all three regressor variables are needed in the model due to their p-values being less than 0.05.

**10.21.** The value of the adjusted *R*2 statistic always increases when a new regressor variable is added to the model.

True False

False, the adjusted *R*2 statistic always decreases as new regressor variables are added to the model and increases as the sample size is increased.