Flipped Assignments

Flipped Assignment 1 (1/18)

Do the following problem by hand (upload a scanned pdf file). Consider the following data:

x	y
1	4
2	2
3	3

a. Plot the points on a graph by hand. Eyeball a straight line that you believe fits the points "best"
b. Fit a straight line to the data, f(x)=mx+b, using least squares. Sketch the fitted curve on the same graph created in part a. Comment on any similarities or dissimilarities between these lines.
c. Calculate SST
d. Calculate SSE
e. Calculate SSR
f. Sketch the errors (total, error, regression) for each of the data points on the graph.

Flipped Assignment 2 (1/25)

Consider the following data and assume that the response variable Y is Normally distributed. Upload a pdf of your handwritten work in response to this assignment

x	y
1	1.5
2	2
3	3
3	4

4	4.5
4.5	6

a. Fit a simple linear regression model to this data using least squares
b. What is your estimate of E[Y|X]?
c. What is your estimate of Var[Y|X] = Var[Y]?
d. Test the hypothesis that $\beta_1=0$ versus the alternative that it is not at the .05 level of significance
e. Interpret your findings, is the regression significant?
f. Calculate a 95% confidence interval on the mean response when x=3
g. Calculate a 95% prediction interval on a new observation when x=3

Flipped Assignment 3 (1/27)

Attached Files: Problem2_13.docx (557.919 KB)

Consider problem 2.13 from your book.

- Regress Days on Index using simple linear regression, what are the estimates of your fitted regression line?
- What is the value of R^2?
- Test for the signifiance of the regression at a 0.05 level of signficance assuming the response is Normally distributed, what is your conclusion?
- Regardless of whether you conclude that the regression is signficant above, make a scatterplot of the data showing the fitted regression line, confidence interval, and prediction interval
- Calculate a 95% confidence interval on the mean number of days the ozone level exceeds 20ppm when the meterological index is 17.0. Comment on the meaning of this interval?
- Calculate a 95% prediction interval on the mean number of days the ozone level exceeds 20ppm when the meterological index is 17.0. Comment on the meaning of this intervall? Compare the width of the prediction interval to that of the confidence interval and comment

Flipped Assignment 4 (2/8)

Attached Files: Problem2_13.docx (95.074 KB)

Consider problem 2.13 in your text book. You may use R to do your calculations, but you need to additionally generate a report in pdf/Word using the output from R with appropriate comments. Upload your RScript+Word Document or RDM+Word Document or RMD+pdf

a. Make a scatterplot of the data, Title the plot and label the axes appropriately
b. What are the Least Squares estimates of the parameters of a simple linear regression?
c. Add the least squares line to the scatterplot.
d. Check for model adequacy using diagnostic plots. Show the plots and comment on the assumptions of Constant Variance, Normality, and whether there appear to be outliers
e. Does the regression appear to be signficant? Why or why not?

Flipped Assignment 5 (2/10)

Consider problem 5.3 in your text book, in which the Number of Bacteria is regressed on Minutes of Exposure. The data is reported as follows.

Number of Bacteria	Minutes of Exposure
175	1
108	2
95	3
82	4
71	5
50	6
49	7
31	8
28	9
17	10
16	11
11	12

Upload a commented R file that answers the following questions.

a. Fit a simple linear regression model to the data. What Is the value of R^2?
b. Check for model adequacy (comment)
c. Use Box-Cox to perform a power transformation, transform the data as appropriate
d. Fit a simple linear regression model to the transformed data. What Is the value of R^2?
e. Check for model adequacy using the transformed data (comment)
f. Estimate the number of bacteria at 10 minutes of exposure, how does this compare with the observed value? (note, the transformed estimate from the fitted regression line must be transformed back)
g. Provide a 95% prediction interval on the number of bacteria at 10 minutes

of exposure.

Flipped Assignment 6 (2/15)

Attached Files: data-table-B8.csv (453 B)

Do problem 3.12 in the book **without using the LM function in R**. The corresponding data file is data-table-B8.csv is attached. Specifically, the problem is regressing clathrate formation (y - mass%) on the amount of surfactant (x1 - mass %) and time (x2 - min).

a. Write the model equation of clathrate formation regressed on both surfactant and time.
b. Fit the model in R without using the LM function. What is the dimensionality of Y? What is the dimensionality of X?
c. What are the least squares estimates of the regression parameters in the model?

Flipped Assignment 7 (2/17)

Attached Files: data-table-B8.csv (453 B)

Do problem 3.12 in the book **using the LM function in R**. The corresponding data file is data-table-B8.csv is attached. Specifically, the problem is regressing clathrate formation (y - mass%) on the amount of surfactant (x1 - mass %) and time (x2 - min).

a. Fit the model in R using the LM function.
b. What are the least squares estimates of the regression parameters in the model?

Flipped Assignment 8 (2/22)

Attached Files: data-table-B9.csv (1.548 KB)

Do problem 3.13 in the book **using the LM function in R**. Specifically, the problem studies the effect of four factors (x1 = superficial fluid velocity of the gas (cs/s), x2= kinematic velocity, x3= mesh opening (cm), and x4= dimensionless number relating the superficial velocity of the gas to the fluid) on the pressure drop in a screen plate bubble column (y). The corresponding data file data-table-B9.csv is attached.

a) Consider a first order multiple regression model with two-factor interactions.

Check for model adequacy and make any corrective actions if deemed necessary. Test for the signifcance of the full regression model, what do you conclude?
b) Test for the signifiance of all 2 factor interactions using a partial F-test. What are your findings?
c) Determine the best fitting model using partial F and/or t-tests. What is the final model?
d) Using the model from part c), calculate a 95% confidence interval on the mean response at the following points of interest (note, if a variable in the points of interest are not in the model, then it is omitted).
e) Using the model from part c), calculate a 95% prediction interval on the mean response at the points of interest (note, if a variable in the points of interest are not in the model, then it is omitted).

Points of Interest			
X1	**X2**	**X3**	**X4**
5.0	10.0	0.5	0.75
10.0	3.0	0.25	0.85