GR5205 Section 4 - Homework 4:

Due date: 10/10 by 7:25pm EST.

1. Chapter 5, page 212 5.25.

No need to work on b., i.e., only work on a and c below

*5.25. Refer to **Airfreight breakage** Problems 1.21 and 5.6.

a. Using matrix methods, obtain the following: (1) $(\mathbf{X}'\mathbf{X})^{-1}$, (2) $\mathbf{b}$, (3) $\mathbf{e}$, (4) $\mathbf{H}$, (5) SSE, (6) $\mathbf{s}^2\{\mathbf{b}\}$, (7) $\hat{Y}_h$ when $X_h = 2$, (8) $s^2\{\hat{Y}_h\}$ when $X_h = 2$.

b. From part (a6), obtain the following: (1) $s^2\{b_1\}$; (2) $s\{b_0, b_1\}$; (3) $s\{b_0\}$.

c. Find the matrix of the quadratic form for SSR.

Chapter 6, page 251 6.15.

Only need to work on b, c, d, e, f, g

You can use lm directly

*6.15. **Patient satisfaction.** A hospital administrator wished to study the relation between patient satisfaction (Y) and patient's age (X_1, in years), severity of illness (X_2, an index), and anxiety

level (X_3, an index). The administrator randomly selected 46 patients and collected the data presented below, where larger values of Y, X_2, and X_3 are, respectively, associated with more satisfaction, increased severity of illness, and more anxiety.

i:	1	2	3	...	44	45	46
X_{i1}:	50	36	40	...	45	37	28
X_{i2}:	51	46	48	...	51	53	46
X_{i3}:	2.3	2.3	2.2	...	2.2	2.1	1.8
Y_i:	48	57	66	...	68	59	92

a. Prepare a stem-and-leaf plot for each of the predictor variables. Are any noteworthy features revealed by these plots?

b. Obtain the scatter plot matrix and the correlation matrix. Interpret these and state your principal findings.

c. Fit regression model (6.5) for three predictor variables to the data and state the estimated regression function. How is b_2 interpreted here?

d. Obtain the residuals and prepare a box plot of the residuals. Do there appear to be any outliers?

e. Plot the residuals against $\hat{Y}$, each of the predictor variables, and each two-factor interaction term on separate graphs. Also prepare a normal probability plot. Interpret your plots and summarize your findings.

f. Can you conduct a formal test for lack of fit here?

g. Conduct the Breusch-Pagan test for constancy of the error variance, assuming $\log \sigma_i^2 = \gamma_0 + \gamma_1 X_{i1} + \gamma_2 X_{i2} + \gamma_3 X_{i3}$; use $\alpha = .01$. State the alternatives, decision rule, and conclusion.