Exam 1 Solutions

Math 430, Winter 2017

Problem 1

1.

The correlation coefficient, $r = 0.838$, is the most appropriate statistic to summarize the association between hemoglobin concentration and lymphocyte count. You are interested in the **association** between two observed quantities, not prediction.

If you wanted to use a regression slope, which slope is the correct one, i.e. should X be the H or the L variable?

2.

We expect hemoglobin concentration to increase by 1.98 for a 1 unit increase in the lymphocyte count.

3.

$$se(\widehat{\beta}_1) = \frac{S}{\sqrt{SXX}} = \frac{4.95}{1428} \approx 0.131$$

4.

To calculate t^*, you find the .96 quantile of a t distribution with 101 degrees of freedom.

5.

$\widehat{\beta}_1 \pm t^* se(\widehat{\beta}_1) = 1.98 \pm 1.77(0.131) = (1.734, 2.212)$

6.

Source	d.f.	SS	MS	F
Regression	1	5576	5576	227.6
Error	101	2474.2	24.5	
Total	102	8050.2		

7.

For a simple linear regression, the most precise prediction is at the mean of X. $L_i = 32$ is closest to the mean.

8.

$\widehat{L} = 30.81$, which is obtained by solving $5.4 = -55.6 + 1.98L$.

Problem 2

1.

For a 1 cent per gallon increase in the gasoline tax rate, we expect fuel consumption to decrease 4.2 gallons per 1,000 people over the age of 16, holding all other variables constant.

2.

$R^2 = 0.5105$

3.

$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4$

H_A : at least one $\beta_i \neq 0$ for $i = 1, 2, 3, 4$

4.

The p-value of the F-test is less than 0.0001, indicating that there is very strong evidence that at least one of the slope terms is not zero. In other words, there is very strong evidence that at least one of the predictor variables is useful in explaining variability in fuel consumption.

Problem 3

The normal Q-Q plot does not provide any evidence against the assumption of normality of residuals.

The standardized residual plot exhibits both curvature and non-constant variance; thus, the assumptions of linearity and constant error variance appear to be violated.

We cannot use one of the plots to assess the assumption of independent residuals.