Stats 100C: HW3 due Friday, 4/17/09 (in class)

EXERCISE 1

Data have been collected for 19 observations of two variables, y and x, in order to run a regression of y on x. You are given that $s_y = 10$, $\sum_{i=1}^{19}(y_i - \hat{y}_i)^2 = 180$.

a. Compute the proportion of the variation in y that can be explained by x.

b. Compute an unbiased estimate of variance σ^2.

Exercise 2

Data on y and x were collected to run a regression of y on x. The intercept is included. You are given the following: $\bar{x} = 76, \bar{y} = 880, \sum_{i=1}^{n}(x_i - \bar{x})^2 = 6800, \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = 14200, r = 0.72, s = 20.13$.

a. What is the value of $\hat{\beta}_1$?

b. What is the value of $\hat{\beta}_0$?

c. What is the value of $\sum_{i=1}^{n}(y_i - \bar{y})^2$?

d. What is the sample size n?

EXERCISE 3

For $i = 1, \cdots n$ let $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ be the straight line regression model in which x_i are centered such that $\sum_{i=1}^{n} x_i = 0$, and ϵ_i are independently distributed with mean 0 and common variance σ^2.

a. Suppose the n points $x_1, \cdots, x_n$ are to be placed in the interval $[-1, 1]$. How should the x_i's be chosen in order to minimize var$(\hat{\beta}_1)$?

b. If we standardize both x and y so that they have mean 0 and standard deviation 1. Show that the fitted line equation has the form $\hat{y}_i = r x_i$, where r is the correlation coefficient between y and x.

EXERCISE 4

A company manufactures auto parts in lots that vary in size as demand fluctuates. The table below contains data on the number of man-hours of labor (y) and on lot size (x), for 10 recent production runs. You are also given:

$$n = 10, \sum_{i=1}^{10} y_i = 1100, \sum_{i=1}^{10} x_i = 500, \sum_{i=1}^{10} x_i y_i = 61800, \sum_{i=1}^{10} x_i^2 = 28400, \sum_{i=1}^{10} y_i^2 = 134660.$$

a. Construct an ANOVA table.

b. Perform an F test to test $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 \neq 0$. Use $\alpha = 0.05$.

c. How much of the variation in y can be explained by x?

d. Compute the correlation coefficient r.

EXERCISE 5

Let z_1, z_2, z_3 be random variables with mean vector and covariance matrix

$$\mu = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad V = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Define the new variables

$$y_1 = z_1 + 2z_3, \ y_2 = z_1 + z_2 - z_3, \ y_3 = 2z_1 + z_2 + z_3 - 7$$

a. Find the mean vector and covariance matrix of $y = (y_1, y_2, y_3)'$.

b. Find the mean and variance of $(y_1 + y_2 + y_3)/3$.