COVER PAGE

STAT 608 Homework 04, Summer 2017

Please TYPE your name and email address below, then convert to PDF and attach as the first page of your homework upload.

NAME:

EMAIL:

STATISTICS 608
Homework 608 S17 04
Due: *11:59 PM, July 3, 2017*

Question 1 [**2+4=6**]

Suppose we have a linear model

$$y_i = \alpha_1 x_{i1} + \alpha_2 x_{i2} + e_i, i = 1, ..., n$$

with two dummy variables

$$x_{i1} = \begin{cases} 1, & i = 1, \dots m \\ 0, & i = m+1, \dots, n \end{cases} ; \quad x_{i2} = \begin{cases} 0, & i = 1, \dots m \\ 1, & i = m+1, \dots, n \end{cases} .$$

There are m people in the first group, and $n - m$ people in the second group.

1.1 Interpret the parameters α_1 and α_2 in the context of the problem.

1.2 Use the formula $\hat{\boldsymbol{\alpha}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ to obtain explicit expressions for α_1 and α_2 in terms of m, n and $y_1, \dots, y_n$.

Question 2 [**6+4=10**]

Suppose we have an ordinary household scale such as might be used in a kitchen. When an object is placed on the scale, the reading is the sum of the true weight and a random error. You have two coins of unknown weights β_1 and β_2. To estimate the weights of the coins, you take four observations:

- Put coin 1 on the scale and observe y_1.
- Put coin 2 on the scale and observe y_2.
- Put both coins on the scale and observe y_3.
- Put both coins on the scale again and observe y_4.

Suppose the random errors are independent and identically distributed with mean 0 and variance σ^2.

2.1 Write a linear model in matrix form and find explicit expressions in terms of $y_1, \dots, y_4$ for the least-squares estimates of the coin weights.

2.2 Explain in words why these estimates make intuitive sense.

Question 3 [**4+2=6**]

Consider the linear model

$$\mathbf{Y} = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \mathbf{e}$$

in which the columns $\mathbf{x}_1$ and $\mathbf{x}_2$ of the design matrix have mean 0 and length 1. That is, $\mathbf{x}_1'\mathbf{x}_1 = 1$, $\mathbf{x}_1'\boldsymbol{J} = \mathbf{0}$, where $\boldsymbol{J}$ is a column consisting entirely of ones and the same is true of $\mathbf{x}_2$. Let ρ be the Pearson correlation coefficient between $\mathbf{x}_1$ and $\mathbf{x}_2$.

3.1 *Show* that

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} n & 0 & 0 \\ 0 & 1 & \rho \\ 0 & \rho & 1 \end{bmatrix}$$

and *verify* that

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} \frac{1}{n} & 0 & 0 \\ 0 & \frac{1}{1-\rho^2} & \frac{-\rho}{1-\rho^2} \\ 0 & \frac{-\rho}{1-\rho^2} & \frac{1}{1-\rho^2} \end{bmatrix}.$$

3.2 Determine what values of ρ will make the variance of $\hat{\beta}_1$ and $\hat{\beta}_2$ larger than $5\sigma^2$.

Question 4 [**4+6=10**]

In a study on weight gain in rabbits, researchers randomly assigned 6 rabbits to $1, 2$ or 3 *mg.* of one of dietary supplement A or B (one rabbit to each level of each supplement). Consider the linear model $\mathbf{Y} = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \mathbf{e}$, where $\mathbf{x}_1$ is the dosage level of the supplement, and $\mathbf{x}_2$ is a dummy variable indicating the type of supplement used.

4.1 Compute the variance inflation factor for the covariate $\mathbf{x}_1$.

4.2 Now suppose the researcher used instead 1, 2 and 3 *mg.* for supplement A, and 2, 3 and 4 *mg.* for supplement B. What is the variance inflation factor for the covariate $\mathbf{x}_1$ in this case? Explain why it is larger or smaller than in **4.1** above.

Question 5 [**2+2+3+3**]

Work Exercise 1 on page 252 of our textbook.