

# ST552 Midterm

*Winter 2016*

Answer the questions in the spaces provided on this exam.

Name: \_\_\_\_\_

- You have 50 minutes to complete the exam.
- There are 3 questions. Answer all of the questions.
- Please
  - do not look at the exam until I tell you and
  - stop writing when I announce that the exam is over.
- There is one page of statistical tables at the end of the exam. You may remove the page of tables if you desire.

| Question | Points | Score |
|----------|--------|-------|
| 1        | 15     |       |
| 2        | 8      |       |
| 3        | 17     |       |
| Total:   | 40     |       |

1. (a) Derive the variance-covariance matrix of the least squares estimates of  $\beta$ . You should begin by stating the multiple linear regression model in matrix form, along with any assumptions you require. (10)

- (b) Imagine the errors have non-zero mean. For example, that  $E(\boldsymbol{\epsilon}) = c$ , where  $c$  is a constant  $n \times 1$  vector. Are the least squares estimates still unbiased? Justify your answer. (5)

2. An observational study was conducted to explore the relationship between the brain size of mammals and their gestation length and lifetime, after accounting for body size.

The following models were fit to a dataset that contains data on 55 mammals:

$$\log(\text{Brain size})_i = \beta_0 + \beta_1 \log(\text{Body Size})_i + \beta_2 \log(\text{Gestation Length})_i + \beta_3 \log(\text{Lifetime})_i + \epsilon_i$$

$$\log(\text{Brain size})_i = \beta_0 + \beta_1 \log(\text{Body Size})_i + \epsilon_i$$

The estimates of  $\sigma^2$  are 0.31 and 0.51 for the two models respectively.

- (a) Conduct an F-test to compare the two models. (6)

- (b) An 95% confidence region is estimated for  $(\beta_2, \beta_3)$ , in the larger model. Would you expect the point  $(0, 0)$  to be inside or outside of the confidence region? Justify your answer. (2)

3. An experiment was conducted to explore the relationship between the *lifetime* (measured in days) and sexual activity of fruitflies.

125 fruit flies were divided randomly into 5 treatment groups, each of 25 flies. Each treatment was designed to simulate a different level of sexual activity, with levels: *none*, *one*, *many*, *low* and *high*.

The *thorax length* of each male was also measured as this was known to affect lifetime.

One observation in the *many* group was lost.

The following model was fit:

$$\text{Lifetime}_i = \beta_0 + \beta_1 \text{Thorax Length}_i + \beta_2 \text{one}_i + \beta_3 \text{low}_i + \beta_4 \text{many}_i + \beta_5 \text{high}_i + \epsilon_i$$

where *one*, *low*, *many*, and *high* are indicator variables for the respective treatment groups.

Results from the least squares fit are given below.

$$\hat{\beta} = \begin{pmatrix} -48.8 \\ 134.3 \\ 2.6 \\ -7.0 \\ 4.1 \\ -20.0 \end{pmatrix}, \quad (X^T X)^{-1} = \begin{pmatrix} 1.06 & -1.22 & -0.05 & -0.04 & -0.07 & -0.08 \\ -1.22 & 1.46 & 0.02 & -0.00 & 0.03 & 0.05 \\ -0.05 & 0.02 & 0.08 & 0.04 & 0.04 & 0.04 \\ -0.04 & -0.00 & 0.04 & 0.08 & 0.04 & 0.04 \\ -0.07 & 0.03 & 0.04 & 0.04 & 0.08 & 0.04 \\ -0.08 & 0.05 & 0.04 & 0.04 & 0.04 & 0.08 \end{pmatrix}$$

$$\hat{\sigma} = 10.54,$$

- (a) Conduct a t-test of the null hypothesis that  $\beta_1 = 0$ . (6)

- (b) For two flies with the same thorax length, show the difference between the expected lifetime for a fly in the *one* treatment group and the expected lifetime for a fly in the *low* treatment group is  $\beta_2 - \beta_3$ . (3)

- (c) Construct a 95% confidence interval for  $\beta_2 - \beta_3$ . (6)

(d) Interpret your interval from (c) in the context of the study.

(2)