UNIVERSITY OF MASSACHUSETTS
Department of Mathematics and Statistics
Advanced Exam - Linear Models
Friday, January 17, 2020

Work all problems. 60 points are needed to pass at the Masters Level and 75 to pass at the Ph.D. level.

1. (20pts) Suppose Y_1, Y_2, and Y_3 are measurements of the angles of a triangle subject to error. The information is given as a linear model $Y_i = \theta_i + \epsilon_i$, where θ_i's are the true angles, $i = 1, 2, 3$. Assume that $E(\epsilon) = 0$, and $Var(\epsilon) = \sigma^2$. Obtain the least squares estimates of θ_i (subjet to the constraint $\sum_{i=1}^{3} \theta_i = 180$).

2. If we test two independent hypotheses, each at level α.

   (a) (14pts) Show that the probability of rejecting at least one, even when the null hypotheses are true, is given by $1 - (1 - \alpha)^2$, which is less than 2α.

   (b) (5pts) Show how the above inequality is used in the Bonferroni correction procedure for the weak control of the family-wise error rate in multiple testing.

3. (a) (7pts) Let $\mathbf{x} = (X_1, \cdots, X_k)^T \sim N_k(\mu, \mathbf{D})$, where μ is a $k \times 1$ vector and $\mathbf{D} = diag\{\sigma_1^2, \cdots, \sigma_k^2\}$, $r(\mathbf{D}) = k$. Find the mean and variance of the random variable $U = \mathbf{x}^T \mathbf{D}^{-1} \mathbf{x}$.

   (b) (7pts) Let $\mathbf{x} = (X_1, \cdots, X_k)^T \sim N_k(\mu, \mathbf{\Sigma})$, where μ is a $k \times 1$ vector and $r(\mathbf{\Sigma}) = k$. What is the distribution of $U = (\mathbf{x} - \mu)^T \mathbf{\Sigma}^{-1} (\mathbf{x} - \mu)$?

   (c) (7pts) Suppose $\mathbf{A} = \mathbf{D}^{-1} - (\mathbf{D}^{-1}\mathbf{1}_k\mathbf{1}_k^T\mathbf{D}^{-1})/\mathbf{1}_k^T\mathbf{D}^{-1}\mathbf{1}_k$. Assume that $\mathbf{x} \sim N_k(\mu, \mathbf{D})$. Find the distribution of $\mathbf{x}^T\mathbf{A}\mathbf{x}$.

4. Consider the following model:

$$
\begin{aligned}
Y_1 &= \tau_1 + \tau_2 + \tau_3 + \epsilon_1 \\
Y_2 &= \tau_1 \phantom{{}+\tau_2} + \tau_3 + \epsilon_2 \\
Y_3 &= + \tau_2 \phantom{{}+\tau_3} + \epsilon_3
\end{aligned}
$$

   (a) (5pts) Write out the model in matrix form. What is the rank of the design matrix?

   (b) (5pts) Is $\tau_i, i = 1, 2, 3$ estimable? Is $\tau_1 - 2\tau_2 + \tau_3$ estimable? Explain.

   (c) (5pts) Find two different linear unbiased estimate of $\tau_1 - 2\tau_2 + \tau_3$.

   (d) (5pts) Find the BLUE (Best Linear Unbiased Estimator) of $\tau_1 - 2\tau_2 + \tau_3$.

5. Suppose in truth the model is

$$Y = X\beta + Z\eta + \epsilon, E(\epsilon) = 0, Cov(\epsilon) = \sigma^2 I,$$

but we fit the smaller model

$$Y = X\beta + \epsilon,$$

and let $\hat{\beta}$ be the corresponding least square fit. Assume X has full rank.

(a) (15pts) Find $E(\hat{\beta})$ and an expression for the bias of $\hat{\beta}$.

(b) (5pts) Show that the bias is zero if X and Z are orthogonal.