BST 140.752 Midterm exam

Notes:

- You may not use a calculator for this exam.
- Please be neat and write legibly. Use the back of the pages if necessary.
- Good luck!

____

printed name

1. Let $Y = X_1\beta_1 + X_2\beta_2 + \epsilon$ for X_1 an $n \times p_1$ matrix and X_2 an $n \times p_2$ so that $X_1'X_2 = 0$ where σ^2 is known.

   A. Argue that the least squares estimate of β_1 doesn't depend on whether X_2 is included in the model or omitted.

   B. Derive a Chi-squared test of $H_0 : K'\beta_1 = m$ verus $H_a : K'\beta_1 \neq m$ and show the Chi-squared statistics does not depend on whether X_2 is included in the model or omitted.

   C. Suppose σ^2 was unkown and an F test was performed. Does the denominator depend on whether X_2 was included in the model or omitted? (Just give an argument, no formal proof needed.)

2. Consider the true model $Y = X_1\beta_1 + X_2\beta_2 + \epsilon$, where, unlike the previous problem, we are no longer assuming X_1 and X_2 are orthogonal. Conisder thethat we fit an incorrect model $Y = X_1\beta_1 + \epsilon$ (i.e. omitted X_2 errantly).

   For an estimator $\hat{\beta}$ of β, define the mean squared error to be $MSE(\hat{\beta}) = E[(\hat{\beta} - \beta)'(\hat{\beta} - \beta)]$ and the bias to be $B(\hat{\beta}) = E[\hat{\beta}] - \beta$.

   A. Show that $MSE(\hat{\beta}) = tr\{Var(\hat{\beta})\} + B(\hat{\beta})'B(\hat{\beta})$.

   B. Let $\hat{\beta}_1$ be the estimate of β_1 using the model that excludes X_2. Derive the bias, variance and mean squared error of this estimate.

3. Let $Y_{ij} = \beta_i + \epsilon_{ij}$ for $i = 1, 2$ and and $j = 1, \ldots J$ and $\epsilon_{ij} \sim N(0, \sigma^2)$.

   A. Derive the F test for $\beta_1 = \beta_2$ and demonstrate how it is related to the variation between groups to the variation within groups.

   B. Argue that the estimate of σ^2 is the average of the within group variances.

   C. Derive a 95% lower confidence bound for $\beta_1 - \beta_2$.

4. Let $Y = X\beta + \epsilon$ where $\epsilon \sim N(0, \sigma^2 I)$ and X contains an intercept column. Let 1 be a vector of ones. Show the following

   A. $1 = X(X'X)^{-1}X'1$

   B. (Assume the previous problem.) Show that $X(X'X)^{-1}X' - 1(1'1)^{-1}1'$ is idempotent.

   C. Let $\hat{Y} = X(X'X)^{-1}X'Y$ and $\bar{Y} = 1(1'1)^{-1}1'Y$. Show that

   $$||Y - \bar{Y}||^2 = ||Y - \hat{Y}||^2 + ||\bar{Y} - \hat{Y}||^2$$

   (i.e. that the variation in Y decomposes into error variation and regression variation.)