Biostatistics 533

Classical Theory of Linear Models

Spring 2007

Midterm

Name: ______________________________

Problems do not have equal value and some problems will take more time than others. Spend your time wisely. This test has six pages including this title page.

Problem	1	2	3	4	5	6	Total
Possible Points	20	5	10	20	10	15	80
Score							

All problems pertain to the linear model

$$\mathbf{Y}_{n\times 1} = \mathbf{X}_{n\times p}\boldsymbol{\beta}_{p\times 1} + \boldsymbol{\varepsilon}_{n\times 1}$$

with $E[\boldsymbol{\varepsilon}] = \mathbf{0}$, $\text{cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$.

1. (20 points) Let $\mathbf{P}$ be the projection operator onto $\mathcal{R}(\mathbf{X})$. For least-squares estimation, recall that $\hat{\boldsymbol{\varepsilon}} = (\mathbf{I} - \mathbf{P})\mathbf{Y}$. Derive

(a) $E(\hat{\boldsymbol{\varepsilon}})$

(b) $\mathrm{cov}(\hat{\boldsymbol{\varepsilon}})$

(c) $\mathrm{cov}(\hat{\boldsymbol{\varepsilon}}, \mathbf{PY})$

(d) $E[RSS]$

2. (5 points) Suppose $\hat{\boldsymbol{\beta}}_1 \neq \hat{\boldsymbol{\beta}}_2$ are two different least-squares estimates of $\boldsymbol{\beta}$. Show there are infinitely many least-squares estimates of $\boldsymbol{\beta}$.

3. (10 points) Suppose $\text{rank}(\mathbf{X}) < p$. Show $\boldsymbol{\beta}$ is not estimable. That is, show there is no matrix $\mathbf{C}$ such that $\mathbf{CY}$ is an unbiased estimate of $\boldsymbol{\beta}$. (Equivalently, show that if $\boldsymbol{\beta}$ is estimable then $\mathbf{X}$ has full rank.)

4. (20 points)

(a) What does BLUE stand for?

(b) What does BLUE mean?

(c) We proved in class that for the least squares estimator $\hat{\boldsymbol{\theta}}$ of the mean vector of $\mathbf{Y}$, $\mathbf{c}'\hat{\boldsymbol{\theta}}$ is the BLUE of $\mathbf{c}'\boldsymbol{\theta}$ for any $\mathbf{c}$. Using this fact in the case $\text{rank}(\mathbf{X}) = p$, prove that $\mathbf{d}'\hat{\boldsymbol{\beta}}$ is the BLUE of $\mathbf{d}'\boldsymbol{\beta}$.

5. (10 points) Since $\text{cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$, one might wonder whether the fitted residuals can also be uncorrelated with the same variance. That is, one might wonder whether $\text{cov}(\hat{\boldsymbol{\varepsilon}})$ can have the form $\tau^2\mathbf{I}$. Prove that $\text{cov}(\hat{\boldsymbol{\varepsilon}}) = \tau^2\mathbf{I}$ for some $\tau^2 \geq 0$ if and only if $\hat{\mathbf{Y}} = \mathbf{Y}$.

6. (15 points) Circle true or false after each statement

$\hat{\mathbf{Y}}$, the least-squares estimate, is always unique

TRUE FALSE

$\hat{\boldsymbol{\beta}}$, the least-squares estimate, is always unique

TRUE FALSE

The result that $\hat{\mathbf{Y}}$ is the BLUE of $E[\mathbf{Y}]$ requires the assumption that $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$

TRUE FALSE