Take-Home Final Exam (200 points): Due on 5:00 PM, Friday, December 11, 2020

- *All work on this exam must be your own;*
- *You can not discuss the exam with anyone;*
- *Please submit your PDF file made by LATEX, MS Word etc.. if possible.*

1. (20 points) Consider a trivariate normal random vector with a nonsingular covariance matrix

$$\begin{pmatrix} y \\ x_1 \\ x_2 \end{pmatrix} \sim N_3\left(\begin{pmatrix} \mu_y \\ \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_y^2 & \sigma_{y1} & \sigma_{y2} \\ \sigma_{y1} & \sigma_1^2 & \sigma_{12} \\ \sigma_{y2} & \sigma_{12} & \sigma_2^2 \end{pmatrix} \right).$$

Is it possible to construct an example so that

$$y \perp\!\!\!\perp x_1 | x_2 \quad \text{and} \quad y \perp\!\!\!\perp x_2 | x_1$$

and yet y is depenedent on $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$?

2. (20 points) Assume

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_n\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right).$$

(2-1) (10 points) Show that

$$\min_{\mathbf{c}} \mathbf{E}[(X_1 - \mathbf{c})^T (X_1 - \mathbf{c}) = \text{tr}(\Sigma_{11})]$$

is attained at $\mathbf{c} = \mu_1$.

(2-2) (10 points) Show that

$$\min_{\mathbf{C},\mathbf{d}} \mathbf{E}[(X_1 - (\mathbf{C}X_2 + \mathbf{d}))^T (X_1 - (\mathbf{C}X_2 + \mathbf{d}))] = \text{tr}(\Sigma_{11\cdot 2})$$

is attained at $\mathbf{C} = \Sigma_{12}\Sigma_{22}^-$ and $\mathbf{d} = \mu_1 - \Sigma_{12}\Sigma_{22}^-\mu_2$ where $\mu_{1\cdot 2} = \mu_1 + \Sigma_{12}\Sigma_{22}^-(X_2 - \mu_2)$ and $\Sigma_{11\cdot 2} = \mathbf{E}[(X_1 - \mu_{1\cdot 2})(X_1 - \mu_{1\cdot 2})^T]$.

3. (40 points) Consider the linear model

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

where $E(\epsilon) = 0$ and $\text{Cov}(\epsilon) = \sigma^2 V$, where V is a known positive definite matrix and X has rank r. Consider estimating $\rho^T X\beta$. Show that

$$\rho^T AY = \rho^T MY$$

if and only if $C(VX) = C(X)$, where $M = X(X^T X)^- X^T$ and $A = X(X^T V^{-1} X)^- X^T V^{-1}$.

4. (40 points) Consider a full rank linear model: $X = (X_0, X_1)$ where $X_0 : n \times d$ and $X_1 : n \times (p-d)$.

$$Y = X_0\beta_0 + X_1\beta_1 + \epsilon \qquad (1)$$

Let $Y_* \equiv e(Y|X_0)$ and $X_* \equiv e(X_1|X_0)$ denote the residuals from fitting $Y = X_0\alpha + \epsilon$ and $X_1 = X_0\gamma + \epsilon$ respectively. Let's consider

$$Y_* = X_*\delta + \epsilon. \qquad (2)$$

(4-1) (15 points) Prove or disprove that $\hat{\beta}_1 = \hat{\delta}$.

(4-2) (10 points) Let SSE(1) and SSE(2) denote SSE from models (1) and (2) respectively. Discuss whether SSE(1) = SSE(2) or not.

(4-3) (15 points) Show that

$$\hat{\alpha} = \hat{\beta}_0 + \hat{\gamma}\hat{\beta}_1.$$

5. (40 points) Let Y_{ij} be independent normal random variables with a common variance and suppose that

$$Y_{ij} = \eta_i + \beta_j + \epsilon_{ij} \qquad \text{for } i, j = 1, \ldots, K.$$

Find, as explicitly as you can, the F-test of the hypothesis

$$\mathrm{H}_0 : \eta_i - \beta_i = 2 \quad \text{for } i = 1, \ldots, K$$

6. (40 points) Consider the general ANCOVA model as the following

$$Y = X\beta + Z\gamma + \epsilon$$

where X is an $n \times p$ experimental design matrix for an ANOVA, Z is an $n \times s$ matrix of concomitant variables and $\epsilon \sim N_n(0, \sigma^2 I_n)$.

Derive the test for the full model $Y = X\beta + Z\gamma + \epsilon$ versus the reduced model $Y = X\beta + Z_0\gamma_0 + \epsilon$, where $\mathcal{C}(Z_0) \subset \mathcal{C}(Z)$. Describe how the procedure would work for testing $H_0 : \gamma_2 = 0$ in the model

$$y_{ij} = \mu + \alpha_i + \eta_j + \gamma_1 z_{ij1} + \gamma_2 z_{ij2} + \epsilon_{ij}, \ i = 1, \ldots, a, \ j = 1, \ldots, b.$$