Qualifying Exam (Linear Models)

9/28/04

1. Suppose that the model $Y_i = \beta_0 + \beta_1 x_{i1} + \varepsilon_i$ is fitted when the true model is actually $E(Y_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$. If e_i is the residual from the fitted model, prove that $E(e_i) = \beta_2 (x_{i2} + g x_{i1} + h)$, where g and h are functions of the x_{ij}.

2. Let $E\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}$, where $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_p)$. Let $\mathbf{x}_i^*$ be the projection of $\mathbf{x}_i$ onto the linear space spanned by all of the other columns of $\mathbf{X}$ and let $\mathbf{x}_i^\perp = \mathbf{x}_i - \mathbf{x}_i^*$.
   (a) Show that β_i is estimable if and only if $\mathbf{x}_i^\perp \neq \mathbf{0}$.
   (b) Show that $\operatorname{var}(\hat{\beta}_i) = \sigma^2 / |\mathbf{x}_i^\perp|^2$.

3. Let $Y_{ij} = \mu_i + \varepsilon_{ij}$, $i = 1, \ldots, 4$, $j = 1, \ldots, n$, where the ε_{ij} are independently distributed as $N(0, \sigma^2)$. Find a test statistic for testing the hypothesis that $\mu_1 = 2\mu_2 = 3\mu_3$.

第一份

1. Consider the simple linear model: $y_i = \alpha + \beta x_i + \varepsilon_i$, $i = 1,2,3$,

where $E(\varepsilon_i) = 0$, $V(\varepsilon) = \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$, $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$,

and $x_1 = -1$, $x_2 = 0$, $x_3 = 1$. Find the BLUE of α and β. (20%)

2. Consider the two models, $y_1 = X_1\beta_1 + \varepsilon_1$, and $y_2 = X_2\beta_2 + \varepsilon_2$, where X_i's are $n_i \times k$ full rank matrices, $\varepsilon_i \sim N(0, \sigma^2 I)$, $i=1,2$, and $\varepsilon_1, \varepsilon_2$ are independent.

(1) Combine the two models into a single model, and write it as $y = X\beta + \varepsilon$,

then $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$, $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$.

(1a) Derive an unbiased estimator for σ^2. (10%)

(1b) Obtain a test statistic for $H_0 : \beta_1 = \beta_2$, and find out its distribution. (25%)

(2) What is the estimator of β under the condition that $\beta_1 = \beta_2$? (15%)

3. Let the two-way model be $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$, where i=1,…,I, j=1,…,J, and $k = 1,\ldots,n$. Also, $\sum_i \alpha_i = \sum_j \beta_j = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0$.

(1) Derive the test statistic for the null hypothesis:

$\alpha_1 = \alpha_2 = \ldots = \alpha_I = 0$ and $\gamma_{11} = \gamma_{12} = \ldots = \gamma_{IJ} = 0$ (20%)

(2) Specify the distribution of the test statistic in (1) if H_0 is not true. (10%)

第二份

1. Let $\vec{Y}^T = (Y_1, Y_2, \cdots, Y_n)$ denote an n-dimensional random vector and has an n-dimensional normal distribution with mean vector $\vec{\mu}$ and variance-covariance matrix Σ. It is denoted by $\vec{Y} \sim N_n(\vec{\mu}, \Sigma)$.

(a) If $Z_1, Z_2, \cdots, Z_n$ are *i.i.d.* $N_1(0,1)$, show that the moment generating function (*m.g.f.*) of $\vec{Z}^T = (Z_1, Z_2, \cdots, Z_n)$ is $M_{\vec{Z}}(\vec{t}) = \exp(\frac{\vec{t}^T\vec{t}}{2})$, where $\vec{t} = (t_1, t_2, \cdots, t_n)^T \in R^n$.

(b) Using the result of (a) to show that the *m.g.f.* of $\vec{Y}$ is

$$M_{\vec{Y}}(\vec{t}) = \exp(\vec{\mu}^T\vec{t} + \frac{\vec{t}^T \Sigma \vec{t}}{2}) \quad \text{where } \vec{t} = (t_1, t_2, \cdots, t_n)^T \in R^n.$$

Let

$$\vec{Y} = \begin{pmatrix} \vec{Y}_1 \\ \vec{Y}_2 \end{pmatrix}, \ \vec{\mu} = \begin{pmatrix} \vec{\mu}_1 \\ \vec{\mu}_2 \end{pmatrix}, \ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

where $\vec{Y}_1$ and $\vec{\mu}_1$ are $p \times 1$, and Σ_{11} is $p \times p$.

(c) Show that $\vec{Y}_1 \sim N_p(\vec{\mu}_1, \Sigma_{11})$, and $\vec{Y}_2 \sim N_{n-p}(\vec{\mu}_2, \Sigma_{22})$ by the *m.g.f.* of $\vec{Y}$.

(d) If $\Sigma_{22} > 0$ and let $\vec{X} = \vec{Y}_1 - \Sigma_{12}\Sigma_{22}^{-1}\vec{Y}_2$. Find the joint distribution of $\vec{X}$ and $\vec{Y}_2$.

(e) Use the results of (d) to show that the conditional distribution of $\vec{Y}_1$, given $\vec{Y}_2$ is $\vec{Y}_1 \mid \vec{Y}_2 = \vec{y}_2 \sim N_p(\vec{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\vec{y}_2 - \vec{\mu}_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$.

第三份

2. Let $\vec{Y} = X\vec{\beta} + \vec{\varepsilon}$, where $\vec{Y}$ is $n \times 1$, X is $n \times p$ of full rank $p(< n)$, $\vec{\beta}$ is a $p \times 1$ vector of parameters, and $\vec{\varepsilon} \sim N_n(\vec{0}, \sigma^2 I_n)$ is an $n \times 1$ vector of random errors, where $\vec{0}$ is the $n \times 1$ zero vector and I_n is the identity matrix of size n.

Let $\hat{\beta}$ denote the ordinary least squares (OLS) estimator of $\vec{\beta}$

(a) Find $\hat{\beta}$.

(b) Suppose we wish to test the hypothesis $H_0 : A\vec{\beta} = \vec{c}$, where A is a known $q \times p$ matrix of rank q and $\vec{c}$ is a known $q \times 1$ vector. Show that the OLS estimator of $\vec{\beta}$ under H_0 is

$$\hat{\beta}_H = \hat{\beta} + (X^T X)^{-1} A^T [A(X^T X)^{-1} A^T]^{-1} (\vec{c} - A\hat{\beta}).$$

(c) Show that the difference between two residuals sum of squares (RSS) based on $\hat{\beta}$ and $\hat{\beta}_H$ is $RSS_H - RSS = (A\hat{\beta} - \vec{c})^T [A(X^T X)^{-1} A^T]^{-1} (A\hat{\beta} - \vec{c})$.

(d) If $E(\vec{W}) = \vec{\theta}$ and $Cov(\vec{W}) = \Sigma$, show that

$$E(\vec{W}^T B \vec{W}) = tr(B\Sigma) + \vec{\theta}^T B \vec{\theta}.$$

(e) Use the results of (d) to show that

$$E(RSS_H - RSS) = \sigma^2 q + (A\vec{\beta} - \vec{c})^T [A(X^T X)^{-1} A^T]^{-1} (A\vec{\beta} - \vec{c}).$$

(f) Let $F = \dfrac{(RSS_H - RSS)/q}{RSS/(n-p)}$. Show that F is distributed as $F_{q,n-p}$ if H_0 is true, where $F_{q,n-p}$ is the F-distribution with q and $n-p$ degrees of freedom, respectively.