線性模式（資格考）

90年3月

1. Let $Y_1,\ldots,Y_n$ be independent and $Y_i \sim N(\mu_i,\sigma^2), i=1,\ldots,n$. Consider the model given by

$$\mu_i = \beta_1 + \beta_2 x_i, \qquad i=1,\ldots,n-1$$

and

$$\mu_n = \beta_3,$$

where $x_1,\ldots,x_{n-1}$ are constants.

(1) Show that the model is linear. Find the design matrix, and find an orthogonal basis for the linear subspace of the model.

(2) Find the estimators $\hat\beta_1, \hat\beta_2$ and $\hat\beta_3$ and their distributions.

(3) Find the estimator of $\hat\sigma^2$ and its distribution.

(4) Find a confidence interval for β_3.

(5) Let c be a given number. Derive a test for the hypothesis $H_0 : \beta_3 = \beta_1 + c\beta_2$.

(6) Derive the F-statistic for the hypothesis $H_0 : \beta_1 = \beta_3$.

2. For the one-way classification model

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, i=1,\ldots,a, j=1,\ldots,n_i, \text{ and } \varepsilon_{ij} \sim \text{i.i.d.} N(0,\sigma^2)$$

Derive an F-test for the following hypothesis:

$$H_0 : \frac{\mu+\alpha_1}{c_1} = \frac{\mu+\alpha_2}{c_2} = \cdots = \frac{\mu+\alpha_a}{c_a}$$

where $c_1,\ldots,c_a$ are some given constants.

1. (15%) Consider the model $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. The fitted values are $\hat{\boldsymbol{Y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}$ with LS estimate and the residuals can be expressed as $\boldsymbol{e} = \boldsymbol{Y} - \hat{\boldsymbol{Y}}$. Assume $E(\boldsymbol{\epsilon}) = \boldsymbol{0}$ and $E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T) = \boldsymbol{I}\sigma^2$.

(1) Please find $Var(\boldsymbol{e})$ and verify the correlation between e_i and e_j, where e_i is the ith element of $\boldsymbol{e}$ and $i = 1, 2, \cdots, n$.

(2) Please show the correlation between the vectors $\boldsymbol{e}$ and $\boldsymbol{Y}$.

(3) Please also show the correlation between the vectors $\boldsymbol{e}$ and $\hat{\boldsymbol{Y}}$.

2. (25%) Consider the two models $\boldsymbol{Y}_1 = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \epsilon_1$ and $\boldsymbol{Y}_2 = \boldsymbol{X}_2\boldsymbol{\beta}_2 + \epsilon_2$, where the $\boldsymbol{X}_i$'s are $n_i \times p$ matrices and $i = 1, 2$. Suppose that $\epsilon_i \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$ and that ϵ_1 and ϵ_2 are independent.

(1) Obtain the least squares estimator for $\boldsymbol{\beta}^T = (\boldsymbol{\beta}_1 \ \ \boldsymbol{\beta}_2)$ and its covariance matrix.

(2) Assuming that σ^2 is known, obtain a test for the null hypothesis $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$.

3. (15%) Let $\boldsymbol{\Omega}$ be the space of arrays $\begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ & y_{32} & \end{bmatrix}$. Let $\boldsymbol{C}_j$ be the indicator of column j, and $\boldsymbol{Y} = \mu_1\boldsymbol{C}_1 + \mu_2\boldsymbol{C}_2 + \mu_3\boldsymbol{C}_3 + \epsilon$, where $E(\epsilon) = \boldsymbol{0}$ and $Var(\epsilon) = \boldsymbol{I}_7$. Please find the BLUE for $\eta = 2\mu_1 - \mu_2 - \mu_3$ and determine its variance.

4. (25%) Suppose that the true model is $\boldsymbol{\theta} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\gamma$, but the postulated model is $\boldsymbol{\theta} = \boldsymbol{X}\boldsymbol{\beta}$. What would this have any effects upon the estimates of parameters and the expectation of the residual mean square value (i.e. MSE)?

5. (20%) For the analysis of covariance model $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma x_{ijk} + \epsilon_{ijk}$ where $i = 1, 2$, $j = 1, 2$ and $k = 1, 2$. Please find $Var(\hat{\alpha}_1 - \hat{\alpha}_2)$, both for the models with and without the γx_{ijk} term. Also use the results to obtain 95% confidence interval on $\alpha_1 - \alpha_2$ for both models.

1. Consider a regular linear model $\vec{Y} = X\vec{\beta} + \vec{\varepsilon}$ that contains an intercept term, where $\vec{Y}$ and $\vec{\varepsilon}$ are two vectors of length n, $\vec{\beta}$ is another vector of length p, and X is a $n \times p$ matrix. Let H denote the hat matrix of this model.

   a. Show that the value of the leverage h_{ii} may not be greater than one and less than $\frac{1}{n}$.

   b. Find the sum of all leverages.

   c. Furthermore, find the sum of all the elements in H.

2. Let $X_1, X_2, \cdots, X_n$ be i.i.d. random variables with mean μ and variance σ^2, find the expected value of $(X_1 - X_2)^2 + (X_2 - X_3)^2 + \cdots + (X_{n-1} - X_n)^2$.

3. An experiment is performed comparing a strains of rice and their cross-breedings. A rice plant has both a male parent and a female parent. The same strain could be both the male and female parent. Each possible breeding of males and females was replicated n times.

   Let Y_{ijk} be the yield of the kth replicate of the breeding of the ith male to the jth female. The effect of a particular strain is known to be the same whether it is a male parent or female parent. Therefore, the mean μ_{ij} for the cell with the ith male parent and the jth female parent should be the same as the mean μ_{ji} for the cell with the jth male parent and the ith female parent. Hence, a sensible model for this experiment would be

$$Y_{ijk} = \theta + \alpha_i + \alpha_j + \gamma_{ij} + \varepsilon_{ijk}, \quad i, j = 1, \cdots, a, \quad k = 1, \cdots, n,$$

   where, as usual, ε_{ijk} are independent normal random variables with mean zero and variance σ^2, and θ, α_i, γ_{ij}, and σ^2 are unobserved parameters such that

$$\sigma^2 > 0, \quad \sum_i \alpha_i = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0, \quad \gamma_{ij} = \gamma_{ji}.$$

   a. Find the ordinary least squares estimators for α_i, γ_{ij}.

   b. Find the F-test for testing that $\gamma_{ij} = 0$.

   c. Find the simultaneous confidence intervals for contrasts in the α_i.