Statistical Inference I

Final exam: 2012/1/9(Mon)

Q1
Q2
Q3
Q4
Q5

YOUR NAME________________________

NOTE1: Please write down the derivation of your answer very clearly for all questions. The score will be reduced when you only write answer. Also, the score will be reduced if the derivation is not clear. The score will be added even when your answer is incorrect but the derivation is correct.

1. Two sample problem

Consider a linear model

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i = 1,2;\ j = 1,\ldots,n_i .$$

where $\beta = (\mu, \alpha_1, \alpha_2)^T$ is a vector of unknown parameters.

a) Find the form of $X = Z\beta + \varepsilon$ and $Z^T Z$.

b) Find a generalized inverse $(Z^T Z)^-$ and show $(Z^T Z)^-$ is not unique.

c) Consider estimation of $l^T \beta$ where $l^T = (1, 1, 0)$. Show that there exist a vector ξ such that $l = Z^T Z \xi$ [i.e., $l \in \Re(Z^T Z)$].

d) Show that the LSE $l^T \hat{\beta}$ is unique [does not depend on the choice of $(Z^T Z)^-$].

2. BLUE / Gauss-Markov Theorem

Consider a linear model $X = Z\beta + \varepsilon$, where $E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2 I_n$, σ^2 is unknown, and the design matrix Z is not necessary of full rank. Define a class of linear and unbiased estimators

$$\Im_{LU} = \{ c^T X ; E[c^T X] = l^T \beta \}.$$

a) Find $E[XX^T]$.

b) Let $l \in \Re(Z) = \Re(Z^T Z)$. Show the LSE $l^T \hat{\beta}$ is the best linear unbiased estimator (BLUE) (with detailed calculations).

3. UMP test / Simple hypothesis

Consider a problem of testing a goodness-of-fit based on X .

$$H_0 : X \text{ has a p.d.f. } f_0(x) = \exp(-x^2/2)/\sqrt{2\pi},$$
$$H_1 : X \text{ has a p.d.f. } f_1(x) = \exp(-|x|/2)/4.$$

1) Derive a UMP test with level $\alpha = 0.05$.

2) Calculate the power. Is this unbiased test?

4. UMP test / Two-sided hypothesis

Let $Y \sim Bin(n = 5, p)$.

1) Derive a UMP test for $H_0 : p = 1/3$ vs. $H_1 : p = 2/3$ of level $\alpha = 0.05$.

2) Is the test in 1) is UMP for $H_0 : p = 1/3$ vs. $H_1 : p > 1/3$? (with explanation).

3) Is the test in 1) UMP for $H_0 : p \le 1/3$ vs. $H_1 : p > 1/3$? (with explanation).

5. UMPU test / MLE

Let $X = (X_1, ..., X_n) \sim_{iid} N(\mu, \sigma^2)$ with $\theta = (\mu, \sigma^2)$ being unknown.

1) Derive a UMPU test $T(X)$ for $H_0 : \mu \le 0$ vs. $H_1 : \mu > 0$ of level $\alpha \le 0.5$.

2) Consider the following two-stage estimation scheme:

   If $H_0 : \mu \le 0$ is accepted, then find a MLE $\hat{\theta}_-$ with restricted parameter space $\Theta_- = \{(\mu, \sigma^2);\ \mu < 0, \sigma^2 > 0\}$. If $H_0 : \mu \le 0$ is rejected, then find a MLE $\hat{\theta}_+$ with restricted parameter space $\Theta_+ = \{(\mu, \sigma^2);\ \mu > 0, \sigma^2 > 0\}$.

   Find the formula of the two-stage estimator $\hat{\theta}_{TEST} = \hat{\theta}_-(1 - T(X)) + \hat{\theta}_+ T(X)$.

   (this type of estimator is called "testimator"). What if $\alpha = 0.5$?

3) Please show that the "mean part" of $\hat{\theta}_{TEST} = (\hat{\mu}_{TEST}, \sigma^2_{TEST})$ has the risk

$$E(\hat{\mu}_{TEST} - \mu)^2 = \frac{\sigma^2}{n} - E[(\overline{X} - \mu)^2 I(0 \le \overline{X} \le S \cdot t_{n-1}(\alpha))] + \mu^2 P(0 \le \overline{X} \le S \cdot t_{n-1}(\alpha)).$$

   What if $\alpha = 0.5$?

Answers: This is a simplified answer. In the exam, you need to write more detailed calculations.

Answer 1

a) $X = Z\beta + \varepsilon$, where $X = (X_{11},..., X_{1n_1}, X_{21},..., X_{2n_2})^T$

$$\varepsilon = (\varepsilon_{11},..., \varepsilon_{1n_1}, \varepsilon_{21},..., \varepsilon_{2n_2})^T, \quad Z = \begin{bmatrix} 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \end{bmatrix}, \text{ and } Z^T Z = \begin{bmatrix} n & n_1 & n_2 \\ n_1 & n_1 & 0 \\ n_2 & 0 & n_2 \end{bmatrix}.$$

b) Set $(Z^T Z)^- = \begin{bmatrix} 1/n & 0 & 0 \\ -1/n & 1/n_1 & 0 \\ -1/n & 0 & 1/n_2 \end{bmatrix}$.

Indeed, it holds that $(Z^T Z)(Z^T Z)^-(Z^T Z) = (Z^T Z)$.

Take $(Z^T Z)^- = \begin{bmatrix} a/n & 0 & 0 \\ -a/n & 1/n_1 & 0 \\ -a/n & 0 & 1/n_2 \end{bmatrix}, \quad \forall a \in \mathbf{R}$.

c) Set $\xi = \begin{bmatrix} 0 \\ 1/n_1 \\ 0 \end{bmatrix}$. Then, $l = Z^T Z \xi \quad \Leftrightarrow \quad \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} n & n_1 & n_2 \\ n_1 & n_1 & 0 \\ n_2 & 0 & n_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1/n_1 \\ 0 \end{bmatrix}$.

d) Let $\hat{\beta}$ and $\tilde{\beta}$ be LSE. Then, it holds that $Z^T Z\hat{\beta} = Z^T X$ and $Z^T Z\tilde{\beta} = Z^T X$.

Hence, $l^T \hat{\beta} - l^T \tilde{\beta} = \xi^T Z^T Z(\hat{\beta} - \tilde{\beta}) = \xi^T (ZX - ZX) = 0$.

Answer 2

See the proof of Thorem 3.9 of J. Shao.

Answer 3

1) Based on the Neyman-Peason Lemma, a UMP test is $T(X) = I(|X| < 1-t \text{ or } |X| > t)$ for some $t > 1/2$. If $1/2 < t \le 1$, then $P_0(|X| > t) \ge P_0(|X| > 1) = 0.337 > \alpha$. To be $\alpha = 0.05$, it must be $t > 1$. Hence, $T(X) = I(|X| > t)$ for some $t > 1$. Solving $0.05 = P_0(|X| > t) = 2\{1 - \Phi(t)\}$, we have $t = \Phi^{-1}(1 - 0.05/2) = \Phi^{-1}(0.975) = 1.96$. Therefore, $T(X) = I(|X| > 1.96)$.

2) Unbiased:

$$P_1(|X| > 1.96) = 2\int_{1.96}^{\infty} \exp(-x/2)/4dx = -\exp(-x/2)\Big|_{1.96}^{\infty}$$
$$= \exp(-1.96/2) = \exp(-0.98) > \exp(-1) > 0.05$$

Answer 5

1) $T(X) = I(\bar{X}/S > t_{n-1}(\alpha))$

2) $\hat{\theta}_+ = \begin{cases} (\bar{X}, \hat{\sigma}^2) & \bar{X} > 0 \\ (0, \tilde{\sigma}^2) & \bar{X} \le 0 \end{cases}, \quad \hat{\theta}_- = \begin{cases} (\bar{X}, \hat{\sigma}^2) & \bar{X} < 0 \\ (0, \tilde{\sigma}^2) & \bar{X} \ge 0 \end{cases},$

where $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ and $\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^{\,2}$.

$$\hat{\theta} = \left\{ \begin{pmatrix} \bar{X} \\ \hat{\sigma}^2 \end{pmatrix} I(\bar{X} < 0) + \begin{pmatrix} 0 \\ \tilde{\sigma}^2 \end{pmatrix} I(\bar{X} \ge 0) \right\} I(\bar{X}/S \le t_{n-1}(\alpha))$$
$$+ \left\{ \begin{pmatrix} \bar{X} \\ \hat{\sigma}^2 \end{pmatrix} I(\bar{X} > 0) + \begin{pmatrix} 0 \\ \tilde{\sigma}^2 \end{pmatrix} I(\bar{X} \le 0) \right\} I(\bar{X}/S > t_{n-1}(\alpha))$$
$$= \begin{pmatrix} 0 \\ \tilde{\sigma}^2 \end{pmatrix} I(0 \le \bar{X}/S \le t_{n-1}(\alpha)) + \begin{pmatrix} \bar{X} \\ \hat{\sigma}^2 \end{pmatrix} I(\bar{X} < 0, \text{ or } \bar{X}/S > t_{n-1}(\alpha)).$$

3) Omit.