Statistical Inference
Midterm Examination I

2017/10/16

Time: 1:20 pm – 5:00 pm

1. Consider a linear regression model

$$y_i = \beta_0 + \sum_{j=1}^{K} \beta_j x_{ij} + \epsilon_i, \quad i = 1, \ldots, n, \tag{$*$}$$

where ϵ_i's are independent and identically distributed (i.i.d.) with mean 0 and positive variance σ^2. Let $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_K)'$. We are interesting in estimating $\boldsymbol{\beta}'\boldsymbol{\beta}/\sigma^2$ (which can be regarded as a signal-to-noise ratio), and a natural way to estimate it is

$$\frac{\hat{\boldsymbol{\beta}}'\hat{\boldsymbol{\beta}}}{\hat{\sigma}^2}$$

where $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$,

$$\boldsymbol{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1K} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nK} \end{pmatrix}_{n\times(K+1)},$$

$\boldsymbol{y} = (y_1, \ldots, y_n)'$, $\hat{\sigma}^2 = (n-K-1)^{-1}\boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{M})\boldsymbol{y}$, $\boldsymbol{I}_n$ is a $n \times n$ identity matrix and $\boldsymbol{M} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ is the orthogonal projection matrix onto $C(\boldsymbol{X})$, the column space of $\boldsymbol{X}$.

   (1) Assuming ϵ_1 follows a normal distribution, compute $E(\hat{\boldsymbol{\beta}}'\hat{\boldsymbol{\beta}}/\hat{\sigma}^2)$. Is $\hat{\boldsymbol{\beta}}'\hat{\boldsymbol{\beta}}/\hat{\sigma}^2$ an unbiased estimator of $\boldsymbol{\beta}'\boldsymbol{\beta}/\sigma^2$? If no, please find an unbiased estimator for $\boldsymbol{\beta}'\boldsymbol{\beta}/\sigma^2$.

   (2) Does $\hat{\boldsymbol{\beta}}'\hat{\boldsymbol{\beta}}/\hat{\sigma}^2$ converge in probability to $\boldsymbol{\beta}'\boldsymbol{\beta}/\sigma^2$ when there is no normality assumption on ϵ_1? Why?

2. Consider model ($*$) with $K = 1$ and

$$x_{i1} = \begin{cases} 1, & \text{if } i = 1; \\ i^{-1/2}, & \text{if } i = 2, , \ldots, n. \end{cases}$$

Does $\hat{\boldsymbol{\beta}}$ converge in probability to $\boldsymbol{\beta}$? Why?

3. Consider model ($*$) without normality assumption on the noise.

   (1) Find an asymptotic level 5% test of $H_0 : \beta_0\beta_1 + \beta_1^2 = d$ versus $H_a : \sim H_0$.

   (2) Let $\boldsymbol{x}^*$ be a $(K+1)$-dimensional known vector. Assuming ϵ_1 follows a normal distribution, find a 95% confidence interval for $\boldsymbol{x}^{*\prime}\boldsymbol{\beta}$.

   (3) Construct a 95% confidence interval for $(\boldsymbol{x}^{*\prime}\boldsymbol{\beta})^2$ when the normality assumption of the noise holds true. If you can not do that, please find an asymptotic 95% confidence interval for $(\boldsymbol{x}^{*\prime}\boldsymbol{\beta})^2$.

4. If $X_n \sim \chi^2(n)$ and $y_n \sim \chi^2(n^2)$, prove that

$$n^{1/2}\left(\frac{X_n/n}{y_n/n^2} - 1\right) \xrightarrow{d} \mathrm{N}(0,2).$$

5. Let $(x_i, y_i)'$, $i = 1, \ldots, n$, be i.i.d. bivariate normal with mean $(\mu_x, \mu_y)'$ and covariance matrix

$$\begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix}$$

where $\sigma_x > 0$ and $\sigma_y > 0$. Define

$$\hat{\rho} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}.$$

(1) Find the limiting distribution of $\sqrt{n}(\hat{\rho} - \rho)$ where $\rho = \sigma_{xy}/(\sigma_x\sigma_y)$.

(2) Use the result of part (1) to find the limiting distribution of

$$\sqrt{n}\left(\ln\left(\frac{1+\hat{\rho}}{1-\hat{\rho}}\right) - \ln\left(\frac{1+\rho}{1-\rho}\right)\right)$$

where $\rho \neq \pm 1$.

(3) Use the result of part (2) to establish an asymptotic level 5% test of $H_0 : \rho = 0$ versus $H_a : \rho \neq 0$.

6. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with common probability density function $f_\theta(x)$. Find an maximum likelihood estimate (MLE) for θ in each of the following cases:

(a) $f_\theta(x) = \frac{1}{2}e^{-|x-\theta|}$, $-\infty < x < \infty$.

(b) $f_\theta(x) = e^{-x+\theta}$, $\theta \leq x < \infty$.

(c) $f_\theta(x) = (\theta\alpha)x^{\alpha-1}e^{-\theta x^\alpha}$, $x > 0$, and α known.

7. Find an MLE, if it exists, in each of the following cases:

(a) $X_1, X_2, \ldots, X_n \sim \mathrm{N}(\theta, \theta^2)$, $\theta \in \mathbb{R}$.

(b) $X_1, X_2, \ldots, X_n$ is a sample from

$$P(X = y_1) = \frac{1-\theta}{2}, \quad P(X = y_2) = \frac{1}{2}, \quad P(X = y_3) = \frac{\theta}{2}\,(0 < \theta < 1).$$

(c) $X_1, X_2, \ldots, X_n \sim \mathrm{N}(\theta, \theta)$, $0 < \theta < \infty$.

8. Let $X_1, X_2, \ldots, X_n$ be a sample from exponential density

$$f_\theta(x) = \theta e^{-\theta x},\ x \geq 0,\ \theta > 0.$$

Find the MLE of θ, and show that it is consistent and asymptotically normal.