TA section 3

JERRY C.

108354501@nccu.edu.tw

April 1, 2024

Review: Completeness

完備性

• Definition

設 $T := T(\boldsymbol{X})$ 為一統計量,T 之 pdf 為 $f(t;\theta)$, $\theta \in \Omega$. 當 T 為一完備統計量, 對任一函數 g, 若

$$\mathbb{E}[g(T)|\theta] = 0, \ \forall \theta \in \Omega,$$

則

$$\mathbb{P}(g(T) = 0) = 1, \ \forall \theta \in \Omega,$$

i.e., $g(T) = 0$ almost surely。

定理 3.2

• Theorem

令 $X_1, \cdots, X_n$ 為一組由 k 個參數之指數族分佈所產生之隨機樣本, 其 pdf 可表示成:

$$f(x;\theta) = h(x)c(\theta)\exp\left(\sum_{j=1}^{k} w_j(\theta)t_j(x)\right),$$

其中 $C := \{w_1(\theta), \cdots, w_k(\theta)\} \subset \mathbb{R}^k$ 其值域包含一非空開矩形 (nonempty open set in $\mathbb{R}^k$), 則統計量 $T(\boldsymbol{X}) = \left(\sum_{i=1}^{n} t_1(X_i), \cdots, \sum_{i=1}^{n} t_k(X_i)\right)$ 為一完備充份統計量。

定理 3.1

Theorem

若 $T(\boldsymbol{X})$ 為一完備、充份統計量 (C.S.S.), 則 $T(\boldsymbol{X})$ 與每一個輔助統計量 (A.S.) 獨立。

Example

18. 設$X_1, \cdots, X_n$爲一組由$\mathcal{U}(\theta, 2\theta)$分佈所產生之隨機樣本, $\theta > 0$。試求θ之一最小充分統計量, 並問此統計量是否具有完備性。

- 證 $T = (X_{(1)}, X_{(n)})$ is not complete.
- 利用定義找反例。Consider $g(T) = R - \mathbb{E}_\theta[R]$, with $R = X_{(n)} - X_{(1)}$ which is an A.S.
- $\mathbb{E}_\theta[g(T)] = \mathbb{E}_\theta[R] - \mathbb{E}_\theta[R] = 0$, but $\mathbb{P}(g(T) \neq 0) > 0$. So, T is not complete.
- 或可直接利用定理 3.1, R is not independent of T, 得到 T is not complete。

- 另解: 因為
$$f(x|\theta) = \frac{1}{\theta}\boldsymbol{I}(\theta < x < 2\theta) = h(x/\theta)/\theta,$$
where $h(x/\theta) = \boldsymbol{I}(1 < x/\theta < 2)$, which is free of θ. So, $\boldsymbol{X} \in$ scale family with the scale parameter θ. 可得知 $U := X_{(n)}/X_{(1)}$ is scale invariant, and is thus an A.S.

  [*proof*] Let $X_i = \theta Z_i \Rightarrow X_{(n)} = \theta Z_{(n)}$ and $X_{(1)} = \theta Z_{(1)}$, where $Z \sim f(z)$ which is free of θ. So, $\mathbb{P}(U(\boldsymbol{X}) \leq u) = \mathbb{P}(X_{(n)}/X_{(1)} \leq u) = \mathbb{P}(Z_{(n)}/Z_{(1)} \leq u) = \mathbb{P}(U(\boldsymbol{Z}) \leq u)$, which is free θ. $\square$
- 令 $g(T) := U - \mathbb{E}_\theta[U]$, 其分佈也與 θ 無關。則 $\mathbb{E}_\theta[g(T)] = 0$ but $\mathbb{P}(g(T) \neq 0) > 0$. Thus, T is not complete.
- 或利用定理 3.1,U is not independent of T, 得到 T is not complete。

注意

- 定理 3.2 只能來找完備充份統計量, 不能用來反證不完備性。(以往很多同學常犯的錯誤)
- $P \to Q \iff \neg Q \to \neg P$. (若 P 則 $Q \iff$ 若非 Q 則非 P)
- but NOT: $\neg P \to \neg Q$.

Example: (課本例 3.12)

- $X_1, \cdots, X_n \sim i.i.d. \mathcal{N}(\theta, \theta^2)$. 可知 $T := (T_1, T_2) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ 是 θ 之一最小充份統計量。
- (curved exponential family) $C := (w_1(\theta), w_2(\theta)) = (1/\theta, -1/(2\theta^2))$, 其圖形為二維平面上的曲線, 未包含二維開矩形, 不能用定理 3.2.
- 依定義: 取 $g(T) = 2T_1^2 - (n+1)T_2$, 故知 $\mathbb{E}[g(T)|\theta] = 0\ \forall \theta \in \mathbb{R}$. 但, $\mathbb{P}(g(T) \neq 0) > 0$. 即 T is not complete.

Review: Estimation

找估計量的方法

- 動差法 (method of moments estimator, MME)
- 最大概似法 (maximum likelihood estimator, MLE)
- 貝氏估計法 (bayesian estimator, BE)

估計量的優劣評估

- 不偏性
- 效率性
  - 相對效率性
  - 均方差 (MSE)
  - 最佳不偏估計量 (best unbiased estimator, BUE)
  - 一致最小變異不偏估計量 (uniformly minimum variance unbiased estimator, UMVUE)
  – [Rao-Blackwell Theorem]: (充份 $+$ 不偏 $\Rightarrow$ 效率)
  – [Lehmann-Scheffé (-Rao-Blackwell) Theorem]: (完備 $+$ 充份 $+$ 不偏 $\Rightarrow$ 效率)
- 一致性
- 漸近不偏性
- 漸近效率性
- 漸近常態性
- 最佳漸近常態性

Homework 2 (part I)

§6.2 #2, #5, #6

2. 設$X_1, \cdots, X_n$為一組由$\mathcal{P}(\theta)$分佈所產生之隨機樣本, $\theta > 0$。

(i) 試以動差法求兩種θ之估計量;

(ii) 試利用(i)給出$P(X \neq 0)$之兩種動差估計量。

- Method of Moment Estimator Idea: "Matching the moments".
- $\mathbb{E}[X] = \mathrm{Var}[X] = \lambda$. Let $m_k = n^{-1} \sum_{i=1}^n X_i^k$, for $k \geq 1$.
- (1) $m_1 = \mathbb{E}[X] \Rightarrow \widehat{E}[X] = m_1$, i.e., $\widehat{\lambda}_1 = \overline{X}_n$.
- (2) matching: $m_1 = \mathbb{E}[X]$ and $m_2 = \mathbb{E}[X^2] \Rightarrow \widehat{E}[X] = m_1$ and $m_2 = \widehat{\mathbb{E}}[X^2] = \widehat{\mathrm{Var}}[X] + \widehat{\mathbb{E}}[X]^2 = \widehat{\lambda} + m_1^2$. Thus, $\widehat{\lambda}_2 = m_2 - m_1^2$.
- $\widehat{\mathbb{P}(X \neq 0)} = 1 - \exp(-\widehat{\lambda})$, with plugging in (1) and (2).

5. 設$X_1, \cdots, X_n$爲一組由$\mathcal{U}(-\theta, \theta)$分佈所產生之隨機樣本, $\theta > 0$。試求一θ之最低次的動差估計量。

- $m_1 = \mathbb{E}\,X = 0$ and $m_2 = \mathbb{E}\,X^2 = \theta^2/3$.
- $$\widehat{\theta}^2 = \frac{3}{n}\sum_{i=1}^{n} X_i^2,$$

  or

  $$\widehat{\theta} = +\sqrt{\frac{3}{n}\sum_{i=1}^{n} X_i^2}\ (\because \theta > 0).$$

6. 設$X_1, \cdots, X_n$爲一組由$\mathcal{B}e(\alpha, \beta)$分佈所產生之隨機樣本, $\alpha, \beta > 0$ 。試求α, β之動差估計量 。

- matching: $m_1 = \mathbb{E}[X] = \alpha/(\alpha+\beta)$ and $m_2 = \mathbb{E}[X^2] = \alpha(\alpha+1)/(\alpha+\beta)(\alpha+\beta+1)$.
- first note that
$$\alpha = \frac{m_1 \beta}{(1-m_1)} \tag{1}$$
and
$$\frac{m_1}{m_2} = \frac{\alpha+1}{\alpha+\beta+1}. \tag{2}$$
- Thus, by (1) & (2), solve for
$$\widehat{\alpha} = \frac{m_1(m_1 - m_2)}{m_2 - m_1^2}$$
and
$$\widehat{\beta} = \frac{(1-m_1)(m_1-m_2)}{m_2 - m_1^2}.$$

§6.3 #6, #15, #21, #31, #35

6. 設$X_1, \cdots, X_n$爲一組由對數常態分佈所產生之隨機樣本, 即設$\log X$有$\mathcal{N}(\mu, \sigma^2)$分佈, $\mu \in R$, $\sigma > 0$。試求μ及σ^2之MLE。

- Maximum Likelihood Estimator Idea: "Maximizing the likelihoods (probs.)"
- (Lognormal distribution) let $\theta := (\mu, \sigma^2)$.

$$f_X(x|\theta) = (2\pi\sigma^2)^{-1/2} x^{-1} \exp\Big(- (\log(x) - \mu)^2 / 2\sigma^2 \Big).$$

- log-likelihood function:
  $\log L(\theta) = -\frac{n}{2}\log(2\pi\sigma^2) - \sum_{i=1}^n \log x_i - \sum_{i=1}^n (\log x_i - \mu)^2 / 2\sigma^2$.
- F.O.C.: $\partial \log L / \partial\mu \triangleq 0$, and $\partial \log L / \partial\sigma^2 \triangleq 0$. Sove for

$$\widehat{\mu} = \sum_{i=1}^n \log x_i / n$$

and

$$\widehat{\sigma}^2 = \sum_{i=1}^n (\log x_i - \widehat{\mu})^2 / n.$$

- S.O.C.: Hessian matrix

$$H(\widehat{\theta}) := \begin{pmatrix} \partial^2 \log L/\partial\mu^2 & \partial^2 \log L/\partial\mu\partial\sigma^2 \\ \partial^2 \log L/\partial\sigma^2\partial\mu & \partial^2 \log L/\partial(\sigma^2)^2 \end{pmatrix} |_{\mu=\widehat{\mu},\sigma^2=\widehat{\sigma}^2}$$
$$= \begin{pmatrix} -n/\widehat{\sigma}^2 & 0 \\ 0 & -\sum_{i=1}^n (\log x_i - \widehat{\mu})^2/2(\widehat{\sigma}^2)^3 \end{pmatrix}$$

is negative-definite $(H(1,1) < 0, |H| > 0)$, providing a strictly local maximum.

15. 設X_1, X_2爲由$\mathcal{N}(\mu, \sigma^2)$分佈所產生之隨機樣本, 其中$\mu$爲已知, $\sigma^2 > 0$。若只記錄$Y = X_1 - X_2$, 試求σ^2之MLE。

- (two-points sample) $Y \sim \mathcal{N}(0, 2\sigma^2)$.

$$f_Y(y|\sigma^2) = (4\pi\sigma^2)^{-1/2} \exp(-y^2/4\sigma^2).$$

- $\log L(\sigma^2) = -\frac{1}{2}\log(4\pi\sigma^2) - y^2/4\sigma^2$.
- F.O.C.:

$$d\log L/d\sigma^2 = -1/2\sigma^2 + y^2/4\sigma^4 \overset{\triangle}{=} 0 \Rightarrow \widehat{\sigma}^2 = y^2/2 = (x_1 - x_2)^2/2.$$

- S.O.C.:

$$d^2 \log L/d(\sigma^2)^2|_{\sigma^2=\widehat{\sigma}^2} = 1/2\sigma^4 - y^2/2(\sigma^2)^3|_{\sigma^2=\widehat{\sigma}^2} = -2/y^4 < 0.$$

21. 設$X_1, \cdots, X_n$爲一組由p.d.f.$f(x|\theta) = \theta^x(1-\theta)^{1-x}$, $x = 0, 1$, $0 \le \theta \le 1/2$, 所產生之隨機樣本。試分別求θ之動差估計量及MLE。

- Matching: $m_1 = \mathbb{E}[X] = \theta$. So, $\widehat{\theta}_{MME} = m_1 = \bar{X}_n$.
- log-likelihood function:
  $\log L(\theta) = \sum_{i=1}^n x_i \log\theta + (n - \sum_{i=1}^n x_i)\log(1-\theta), 0 \le \theta \le 1/2$. Then, F.O.C.:

  $$d\log L/d\theta = \sum_{i=1}^n x_i/\theta - (n - \sum_{i=1}^n x_i)/(1-\theta) \overset{\triangle}{=} 0,\ 0 \le \theta \le 1/2,$$

  so

  $$\widehat{\theta}_{MLE} = \bar{X}_n \boldsymbol{I}(0 \le \widehat{\theta} \le 1/2) = (\bar{X}_n \wedge 1/2),$$

  or $\min\{\bar{X}_n, 1/2\}$, i.e., when $\bar{X}_n \le 1/2$, $\widehat{\theta}_{MLE} = \bar{X}_n$; when $\bar{X}_n > 1/2$, $\widehat{\theta}_{MLE} = 1/2$.
- S.O.C.: $\partial^2 \log L/\partial\theta^2|_{\theta=\widehat{\theta}} < 0$.

31. 設$X_1, \cdots, X_n$爲一組由$\mathcal{NB}(r, \theta)$分佈所產生之隨機樣本, 其中$r \geq 1$爲已知, $0 < \theta < 1$。試證θ之MLE爲$\hat{\theta} = nr/(nr + \sum_{i=1}^n X_i)$。

- log-likelihood function:
  $\log L(\theta) = \sum_{i=1}^n \log \binom{x_i+r-1}{x_i} + nr \log \theta + \sum_{i=1}^n x_i \log(1-\theta)$.
- F.O.C.:
  $$d \log L / d\theta = \frac{nr}{\theta} - \frac{\sum_{i=1}^n x_i}{1-\theta} \triangleq 0,$$
  solve for
  $$\widehat{\theta} = \frac{nr}{nr + \sum_{i=1}^n x_i}.$$
- S.O.C.: $\partial^2 \log L / \partial \theta^2 |_{\theta = \widehat{\theta}} < 0$.

35. 設$X_1, \cdots, X_n$爲一組由p.d.f.$f(x|\theta) = 2x/\theta^2$, $0 \le x \le \theta$, 所產生之隨機樣本, $\theta > 0$。試求X之中位數的MLE, 並證明此估計量爲一最小充分統計量。

- $L(\theta) = 2^n \theta^{-2n} \prod_{i=1}^n x_i \boldsymbol{I}(0 < x_i < \theta), i = 1, 2, \cdots, n.$
- log-likelihood function:
  $\log L(\theta) = n \log 2 - 2n \log \theta + \sum_{i=1}^n \log x_i$, $0 < x_i < \theta$.
- F.O.C.: $d \log L / d\theta = -2n/\theta < 0$ for all $\theta > x_{(n)}$.
- So, $\widehat{\theta}_{MLE} = x_{(n)}$.
- the median $M(\theta)$:

$$1/2 = \mathbb{P}(X \le M(\theta)) = \int_0^{M(\theta)} \frac{2x}{\theta^2} dx = M(\theta)^2/\theta^2.$$

So,

$$\widehat{M}(\theta)_{MLE} = M(\widehat{\theta}_{MLE}) = \widehat{\theta}_{MLE}/\sqrt{2} = x_{(n)}/\sqrt{2},$$

by the invariance principle.

- $$\frac{f(\boldsymbol{x}|\theta)}{f(\boldsymbol{y}|\theta)} = \frac{2^n\theta^{-2n}\prod_{i=1}^n x_i \boldsymbol{I}(x_{(n)} < \theta)}{2^n\theta^{-2n}\prod_{i=1}^n y_i \boldsymbol{I}(y_{(n)} < \theta)},$$

  which is free of θ iff $x_{(n)} = y_{(n)}$.
- 令 $T(\boldsymbol{x}) := x_{(n)}$ 且 $U(\boldsymbol{x}) := x_{(n)}/\sqrt{2}$ 是 $T(\boldsymbol{x})$ 之一等價統計量。Thus, $U(\boldsymbol{X}) = X_{(n)}/\sqrt{2}$ is also a M.S.S. for θ.

§6.4 #1, #3

1. 設$X_1, \cdots, X_n$爲一組由$\mathcal{N}(\theta, \sigma^2)$分佈所產生之隨機樣本, $\sigma > 0$爲已知。又設θ有$\mathcal{D}$-$\mathcal{E}(a)$分佈, 其中$a > 0$爲一常數。試求

   (i) 在給定$\boldsymbol{X} = \boldsymbol{x}$之下, θ之事後分佈;

   (ii) θ之貝氏估計量。

- Bayes Estimator in mean squares risk: "Finding the posterior mean" $\mathbb{E}[\theta|\boldsymbol{x}]$, where the posterior pdf:

$$\pi(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)\pi(\theta)}{m(\boldsymbol{x})},$$

$\pi(\theta)$ is a prior distribution, and $m(\boldsymbol{x})$ is the marginal pdf of $\boldsymbol{X}$.

- $$\theta \sim \pi(\theta) = (a/2)e^{-a|\theta|}$$

$$\Rightarrow f(\boldsymbol{x}, \theta) = f(\boldsymbol{x}|\theta)\pi(\theta) = (2\pi\sigma^2)^{-n/2} \exp(-\sum_{i=1}^{n}(x_i - \theta)^2/2\sigma^2) \times (a/2)e^{-a|\theta|}$$

$$\begin{aligned}
&\propto \exp(-\sum_{i=1}^{n}(x_i - \theta)^2/2\sigma^2 - a|\theta|) \\
&\propto \exp(-\frac{n\theta^2}{2\sigma^2} + \frac{n\bar{x}_n\theta}{\sigma^2} - a|\theta|) \\
&\propto \exp(-\frac{n}{2\sigma^2}(\theta - \xi_\pm(\boldsymbol{x}))^2), \text{ where } \xi_\pm(\boldsymbol{x}) := \bar{x}_n \pm \sigma^2 a/n.
\end{aligned}$$

- $m(\boldsymbol{x}) = \int_{\mathbb{R}} \exp(-\frac{n}{2\sigma^2}(\theta - \xi_\pm(\boldsymbol{x}))^2)d\theta$ and $\pi(\theta|\boldsymbol{x}) = f(\boldsymbol{x}, \theta)/m(\boldsymbol{x})$.
- $$m(\boldsymbol{x}) = \int_{\mathbb{R}} \exp(-\frac{n}{2\sigma^2}(\theta - \xi_\pm(\boldsymbol{x}))^2)d\theta = \int_0^\infty \exp(-\frac{n}{2\sigma^2}(\theta - \xi_\pm(\boldsymbol{x}))^2)d\theta$$
$$+ \int_{-\infty}^0 \exp(-\frac{n}{2\sigma^2}(\theta - \xi_\pm(\boldsymbol{x}))^2)d\theta = \sqrt{\frac{2\pi\sigma^2}{n}}.$$

- Thus,

$$\pi(\theta|\boldsymbol{x}) = \frac{\sqrt{n}}{\sqrt{2\pi\sigma^2}} \exp(-\frac{n}{2\sigma^2}(\theta - \xi_{\pm}(\boldsymbol{x}))^2).$$

i.e., $\theta|\boldsymbol{x} \sim \mathcal{N}(\xi_{\pm}(\boldsymbol{x}), \sigma^2/n)$.

-

$$\begin{aligned} \widehat{\theta}_{BE} &= \mathbb{E}[\theta|\boldsymbol{x}] = \mathbb{P}(\theta > 0)\,\mathbb{E}[\theta|\boldsymbol{x}, \theta \geq 0] + \mathbb{P}(\theta < 0)\,\mathbb{E}[\theta|\boldsymbol{x}, \theta < 0] \\ &= 0.5\xi_-(\boldsymbol{x}) + 0.5\xi_+(\boldsymbol{x}). \quad \square \end{aligned}$$

3. 設X有$\mathcal{U}(0,\theta)$分佈, θ之事前分佈爲$\mathcal{E}(1)$。試求

(i) 在給定$X=x$之下, θ之事後分佈;

(ii) θ之貝氏估計量。

- $X|\theta \sim f(x|\theta) = \theta^{-1}\boldsymbol{I}(0<x<\theta)$,
  $\theta \sim \pi(\theta) = e^{-\theta} \Rightarrow f(x,\theta) = f(x|\theta)\pi(\theta) = \theta^{-1}e^{-\theta}\boldsymbol{I}(0<x<\theta)$.
- $m(x) = \int_x^\infty \theta^{-1}e^{-\theta}d\theta = -\mathrm{E_i}(-x)$ (no closed form ?)
  ($\mathrm{E_i}(x) := \int_{-\infty}^x t^{-1}e^t dt$ 是「指數積分函數」)
  $\Rightarrow$ posterior $\pi(\theta|x) = f(x,\theta)/m(x) = \theta^{-1}e^{-\theta}\boldsymbol{I}(0<x<\theta)/m(x)$.
- $$\widehat{\theta}_{BE} = \mathbb{E}[\theta|x] = \int_x^\infty \frac{e^{-\theta}}{m(x)}d\theta = \frac{e^{-x}}{m(x)}.$$
- 題目可能有誤: 若 $\theta \sim \pi(\theta) = \theta e^{-\theta}, \theta > 0$. 則
  $m(x) = \int_x^\infty e^{-\theta}d\theta = e^{-x} \Rightarrow \pi(\theta|x) = e^{x-\theta}\boldsymbol{I}(0<x<\theta)$. 故,
  $\widehat{\theta}_{BE} = \mathbb{E}[\theta|x] = \int_x^\infty \theta e^{x-\theta}d\theta = x-1$.