TA section 1

JERRY C.

108354501@nccu.edu.tw
jerryc520.github.io/teach/MS.html

March 18, 2024

Homework 1 (part I)

§5.2 #6, #8, #9, #16, #19

6. 設$X_1, \cdots, X_n$爲一組隨機樣本, $n \geq 2$。試證統計量

   (i) $(\overline{X}_n, S_n^2)$與$(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$等價;

   (ii) $(X_{(1)}, X_{(n)})$與$(X_{(n)} - X_{(1)}, (X_{(n)} + X_{(1)})/2)$等價。

- 「統計量 (statistic)」$T(\boldsymbol{X})$: 樣本 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 的函數, 跟未知參數 θ 無關。用以對樣本資訊的描述性測量 (measure)。
- 「估計量/估計式 (estimator)」$\widehat{\theta}$: 樣本 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 的函數, 用以推估感興趣的未知參數 θ 的統計量或統計量的函數式。例如: $\overline{X}_n$ 是 母體平均數 μ 的一個好的估計量, 而 S_n^2 不會是 μ 一個好的估計量。
- 兩個統計量 T_1 與 T_2「等價」: $T_2(\boldsymbol{X})$ 與 $T_1(\boldsymbol{X})$ 一對一轉換 (函數)。
- 例如: (affine 轉換)
  $$T_2 = AT_1 + b,$$
  A is nonzero and is of full rank (linear independent, or nonsingular if A is square)。
- 符號可表示為: $T_1 \sim T_2$ (等價關係, 即 T_1, T_2 具反身性, 對稱性, 遞移性)。

(i)

- Let $Y_i = X_i - \overline{X}_n$;
- $T_1 = (\overline{X}_n, S_n^2) = (\overline{X}_n, \sum_{i=1}^n Y_i^2/(n-1))$;
  and $T_2 = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2) = (n\overline{X}_n, \sum_{i=1}^n Y_i^2 + n\overline{X}^2)$.
- Consider

$$A = \begin{bmatrix} 1/n & 0 \\ 0 & 1/(n-1) \end{bmatrix}$$

  and

$$b = \begin{bmatrix} 0 \\ -\frac{n}{n-1}\overline{X}_n^2 \end{bmatrix}$$

- Thus,

$$T_1 = AT_2 + b.$$

(ii)

- $T_1 = (X_{(1)}, X_{(n)})$ and $T_2 = ((X_{(n)} - X_{(1)})/2, (X_{(n)} + X_{(1)})/2)$.
- Consider
$$A = \begin{bmatrix} -1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$
and
$$b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
- Thus,
$$T_2 = AT_1 + b.$$
- (i), (ii) 的轉換方式不唯一。

§5.2 #6, #8, #9, #16, #19

8. 設$X_1, \cdots, X_n$爲一組由$\mathcal{E}(\theta)$分佈所產生之隨機樣本, $\theta > 0$。試分別以下述三法, 判定$T(\boldsymbol{X}) = \sum_{i=1}^n X_i$爲$\theta$之一充分統計量。

   (i) 試證給定$T = t$, $X_1, \cdots, X_n$之條件p.d.f.爲

   $$f_{X_1,\cdots,X_n|T}(x_1, \cdots, x_n|T = t) = \frac{(n-1)!}{t^{n-1}}, \sum_{i=1}^n x_i = t;$$

   (ii) 定理2.2;

   (iii) 定理2.3。

• Definition

設 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 的分佈與參數 θ 有關, 且 joint pdf 為 $f(\boldsymbol{x}|\theta)$。對任一統計量 $T(\boldsymbol{X})$, 若給定 $T(\boldsymbol{X})$, 其 $\boldsymbol{X}$ 之條件分佈, 即

$$f(\boldsymbol{x}|T = t; \theta) = \frac{f(\boldsymbol{x}, t|\theta)}{f_T(t)}$$

與 θ 無關, 則 $T(\boldsymbol{X})$ 為 θ 之一充份統計量。

定理 2.2.

Theorem (分解定理)

令 $f(\boldsymbol{x}|\theta)$ 表 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 的 joint pdf。則 $T(\boldsymbol{X})$ 為 θ 之一充份統計量, 若且唯若存在函數 $g(t|\theta)$ 和 $h(\boldsymbol{x})$, 使得對所有樣本點 $\boldsymbol{x}$ 及參數 θ,

$$f(\boldsymbol{x}|\theta) = g(T(\boldsymbol{x}) = t|\theta)h(\boldsymbol{x}).$$

定理 2.3.

• Theorem (指數族充份統計量)

給定 $X_1, \cdots, X_n$ 為一組由 pdf $f(x|\theta)$ 所產生的隨機樣本。假設 $f(x|\theta)$ 有以下形式:

$$f(x|\theta) = h(x)c(\theta)\exp\{\sum_{j=1}^{k} w_j(\theta)t_j(x)\}\boldsymbol{I}_A(x),$$

其中 $\theta = (\theta_1, \cdots, \theta_d), d \leq k$, 且 $h(x) \geq 0, c(\theta) \geq 0$。則, 令 $\boldsymbol{X} := (X_1, \cdots, X_n)$,

$$T(\boldsymbol{X}) = (\sum_{i=1}^{n} t_1(X_i), \cdots, \sum_{i=1}^{n} t_k(X_i))$$

為 θ 之一充份統計量。

(i)

- Want to show: $f(\boldsymbol{x}|T=t;\theta) = f(\boldsymbol{x},t|\theta)/f_T(t)$, which is free of θ.
- $T = \sum_{i=1}^{n} X_i$, $X_i \sim i.i.d.\mathcal{E}(\theta) = \Gamma(1, 1/\theta)$; then $T \sim \Gamma(n, 1/\theta)$.
- $$f(\boldsymbol{x}|T=t;\theta) = \frac{\theta^n e^{-\theta t}}{\theta^n t^{n-1} e^{-\theta t}/(n-1)!} = \frac{(n-1)!}{t^{n-1}}$$

  which does not depend on θ.
- Thus, $T(\boldsymbol{X}) = \sum_{i=1}^{n} X_i$ is a S.S. for θ.
- Recall: $X \sim \Gamma(\alpha, \beta)$,

  $$f_X(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}, \ x > 0.$$

  MGF: $M_X(t) = (1-\beta t)^{-\alpha}$, $t < 1/\beta$.

(ii)

- $$f(\boldsymbol{x}|\theta) = \theta^n \exp(-\theta \sum_{i=1}^{n} x_i)\boldsymbol{I}(x_i > 0) =: g(T(\boldsymbol{X}) = t|\theta)h(\boldsymbol{x}),$$

  where $g(t|\theta) = \theta^n \exp(-\theta t)$ and $h(\boldsymbol{x}) = 1, x_i > 0$ (or $\boldsymbol{I}(x_i > 0)$).
- Thus, $T(\boldsymbol{X}) = \sum_{i=1}^{n} X_i$ is a S.S. for θ by Factorization theorem.

(iii)

- $$f(x|\theta) = \theta \exp(-\theta x) \boldsymbol{I}(x > 0) =: h(x)c(\theta) \exp(w(\theta)t(x)),$$

  belongs to an one-dimensional exponential family, where $h(x) = \boldsymbol{I}(x > 0); c(\theta) = \theta; w(\theta) = -\theta$ and $t(x) = x$.
- Thus,

  $$T(\boldsymbol{X}) = \sum_{i=1}^{n} t(X_i) = \sum_{i=1}^{n} X_i$$

  is a S.S. for θ.

§5.2 #6, #8, #9, #16, #19

9. 設$X_1, \cdots, X_n$爲一組由$\mathcal{U}(0,\theta)$分佈所產生隨機樣本, $\theta > 0$。試分別以下述二法, 判定$T(\boldsymbol{X}) = X_{(n)}$是否爲$\theta$之一充分統計量。

   (i) 求在給定$X_{(n)} = t$之下, $X_1, \cdots, X_n$之條件分佈;

   (ii) 利用定理2.2。

(i)

- $\mathbb{P}(T \leq t) = \mathbb{P}(X_1 \leq t, \cdots, X_n \leq t) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq t) = \theta^{-n} t^n$, then
$$f_T(t) = \frac{nt^{n-1}}{\theta^n} \boldsymbol{I}(0 \leq t \leq \theta).$$
- $$f(\boldsymbol{x}|T = t, \theta) = \frac{\theta^{-n} \boldsymbol{I}(0 \leq t \leq \theta)}{\frac{nt^{n-1}}{\theta^n} \boldsymbol{I}(0 \leq t \leq \theta)} = \frac{1}{nt^{n-1}},$$
which is free of θ.
- Thus, $T(\boldsymbol{X}) = X_{(n)}$ is a S.S. for θ.

(ii)

- $$\begin{aligned} f(\boldsymbol{x}|\theta) &= \frac{1}{\theta}\boldsymbol{I}(0 \le x_1 \le \theta)\cdots\frac{1}{\theta}\boldsymbol{I}(0 \le x_n \le \theta) \\ &= \boldsymbol{I}(0 \le x_{(1)})\theta^{-n}\boldsymbol{I}(x_{(n)} \le \theta) \\ &=: h(\boldsymbol{x})g(T = t|\theta), \end{aligned}$$

  where $h(\boldsymbol{x}) = \boldsymbol{I}(x_{(1)} \ge 0)$ and $g(t|\theta) = \theta^{-n}\boldsymbol{I}(t \le \theta)$.
- Thus, $T(\boldsymbol{X}) = X_{(n)}$ is a S.S. for θ by Factorization theorem.

§5.2 #6, #8, #9, #16, #19

16. 設$X_1, \cdots, X_n$爲一組由$\mathcal{P}(\theta)$分佈所產生之隨機樣本, $\theta > 0$。試求θ之一最小充分統計量。

定理 2.4.

• Theorem (最小充份統計量)

令 $\boldsymbol{X} := (X_1, \cdots, X_n)$ 之 joint pdf 為 $f(\boldsymbol{x}|\theta)$。假設存在一函數 $T(\boldsymbol{X})$, 使得對任二樣本點 $\boldsymbol{x}$ 及 $\boldsymbol{y}$,

$$\frac{f(\boldsymbol{x}|\theta)}{f(\boldsymbol{y}|\theta)}$$

與 θ 無關, 若且唯若

$$T(\boldsymbol{x}) = T(\boldsymbol{y}).$$

則 $T(\boldsymbol{X})$ 為 θ 之一最小充份統計量。

- Want to show:

$$f(\boldsymbol{x}|\theta)/f(\boldsymbol{y}|\theta),$$

which is free of θ if and only if

$$T(\boldsymbol{x}) = T(\boldsymbol{y}).$$

- $$\frac{f(\boldsymbol{x}|\theta)}{f(\boldsymbol{y}|\theta)} = \frac{\prod_{i=1}^n \frac{1}{x_i!} e^{-n\theta} \theta^{-\sum_{i=1}^n x_i}}{\prod_{i=1}^n \frac{1}{y_i!} e^{-n\theta} \theta^{-\sum_{i=1}^n y_i}} = \prod_{i=1}^n \frac{y_i!}{x_i!} \theta^{\sum_{i=1}^n y_i - \sum_{i=1}^n x_i}$$

is free of θ iff

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i.$$

Let $T(\boldsymbol{x}) := \sum_{i=1}^n x_i$ and $T(\boldsymbol{y}) := \sum_{i=1}^n y_i$.

- Thus,

$$T(\boldsymbol{X}) = \sum_{i=1}^n X_i$$

is a M.S.S. for θ.

§5.2 #6, #8, #9, #16, #19

19. 設$X_1, \cdots, X_n$為一組由p.d.f. $f(x|\theta) = \theta^{-2} e^{-(x-\theta)/\theta^2}$, $x \geq \theta$, $\theta > 0$, 所產生之隨機樣本。試求θ之一最小充分統計量。

- $$\frac{f(\boldsymbol{x}|\theta)}{f(\boldsymbol{y}|\theta)} = \frac{\theta^{-2n}e^{-\sum_{i=1}^{n}(x_i-\theta)/\theta^2}\boldsymbol{I}(x_i \geq \theta)}{\theta^{-2n}e^{-\sum_{i=1}^{n}(y_i-\theta)/\theta^2}\boldsymbol{I}(y_i \geq \theta)}, \ i = 1, 2, \cdots, n$$
$$= \exp\left(\sum_{i=1}^{n}(y_i - x_i)/\theta^2\right)\frac{\boldsymbol{I}(x_{(1)} \geq \theta)}{\boldsymbol{I}(y_{(1)} \geq \theta)},$$

  is free of θ iff

  $$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i, \quad x_{(1)} = y_{(1)}.$$

  Let $T(\boldsymbol{x}) := (\sum_{i=1}^{n} x_i, x_{(1)})$ and $T(\boldsymbol{y}) = (\sum_{i=1}^{n} y_i, y_{(1)})$.
- Thus, the two-dimensional statistic

  $$T(\boldsymbol{X}) = (\sum_{i=1}^{n} X_i, X_{(1)})$$

  is a M.S.S. for θ.