TA section 2

JERRY C.

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

March 28, 2024

Homework 1 (part II)

§5.3 #1, #2, #11, #18, #19

1. 設X_1, X_2爲由p.d.f.$f(x|\alpha) = \alpha x^{\alpha-1} e^{-x^\alpha}$, $x > 0$, $\alpha > 0$, 所產生之隨機樣本。試證$\log X_1 / \log X_2$爲一輔助統計量。

Definition

A statistic $A(\boldsymbol{X})$ is ancillary if the distribution of $A(\boldsymbol{X})$ does not depend on the unknown parameter θ.

- Let $Y = \log X$. $\mathbb{P}(Y \le y) = \mathbb{P}(\log X \le y) = \mathbb{P}(X \le e^y)$. Thus,

$$f_Y(y|\alpha) = f_X(e^y|\alpha)e^y = \alpha(e^y)^{\alpha-1}e^{-(e^y)^\alpha}e^y = \alpha e^{\alpha y - e^{\alpha y}}, y \in \mathbb{R}.$$

- $$f_Y(y|\alpha) = \frac{1}{1/\alpha}\exp\left[\frac{y}{1/\alpha} - e^{y/(1/\alpha)}\right],$$

  belongs to a scale family with scale parameter $1/\alpha$.

- Let $Y_i := (1/\alpha)Z_i$, where the distribution of Z_i has the form of $f(z) \propto \exp(z - e^z)$ which is free of α.

  $T := \dfrac{\log X_1}{\log X_2} = \dfrac{Y_1}{Y_2} = \dfrac{(1/\alpha)Z_1}{(1/\alpha)Z_2} = \dfrac{Z_1}{Z_2}$, whose distribution does not involve α.

- 另種做法: 令 $Y_i = \alpha \log X_i$, 可求得聯合 pdf:

  $$f_{Y_1,Y_2}(y_1, y_2) = e^{y_1+y_2} e^{-(e^{y_1}+e^{y_2})}$$

  與 α 無關。So, the distribution of $T = \alpha \log X_1 / \alpha \log X_2$ does not depend on α.

§5.3 #1, #2, #11, #18, #19

2. 設$X_1, \cdots, X_n$爲一組由一位置族分佈所產生之隨機樣本。令M表樣本中位數。試證$M - \overline{X}_n$爲一輔助統計量。

- $X \in$ location family, such that $X = Z + \mu$, i.e., $F_X(x) = F_Z(x - \mu)$ or $f_X(x) = f_Z(x - \mu)$, where μ is the location parameter.
- Fact: $\boldsymbol{X} \in$ location family, and if the statistic $S(\boldsymbol{X})$ is location invariant, such that $S(\boldsymbol{X}) = S(\boldsymbol{X} + c)$, then $S(\boldsymbol{X})$ is an A.S..

  [*proof*] $\mathbb{P}(S(\boldsymbol{X}) \le x) = \mathbb{P}(S(\boldsymbol{Z} + \mu) \le x) = \mathbb{P}(S(\boldsymbol{Z}) \le x)$, which is free of μ. $\square$
- $X \in$ scale family, such that $X = \theta Z$, i.e., $F_X(x) = F_Z(x/\theta)$ or $f_X(x) = f_Z(x/\theta)/\theta$, where θ is the scale parameter.
- Fact: $\boldsymbol{X} \in$ scale family, and if the statistic $S(\boldsymbol{X})$ is scale invariant, such that $S(\boldsymbol{X}) = S(c\boldsymbol{X})$, then $S(\boldsymbol{X})$ is an A.S..

- Given $X_i = Z_i + \mu \Rightarrow \bar{X} = \bar{Z} + \mu$ and $M(\boldsymbol{X}) = M(\boldsymbol{Z}) + \mu$.

$$
\begin{aligned}
\mathbb{P}(S(\boldsymbol{X}) \le x) &= \mathbb{P}(M(\boldsymbol{X}) - \bar{X} \le x) \\
&= \mathbb{P}(S(\boldsymbol{Z} + \mu) \le x) \\
&= \mathbb{P}(M(\boldsymbol{Z}) + \mu - (\bar{Z} + \mu) \le x) = \mathbb{P}(M(\boldsymbol{Z}) - \bar{Z} \le x) \\
&= \mathbb{P}(S(\boldsymbol{Z}) \le x), \text{ is free of } \mu.
\end{aligned}
$$

So, $S(\boldsymbol{X}) = M(\boldsymbol{X}) - \bar{X}$ is an A.S..

§5.3 #1, #2, #11, #18, #19

11. 設$X_1, \cdots, X_n$爲一組由$\mathcal{G}e(\theta)$分佈所產生之隨機樣本, $0 < \theta < 1$, 令$\boldsymbol{X} = (X_1, \cdots, X_n)$。試證$T(\boldsymbol{X}) = \sum_{i=1}^n X_i$爲$\theta$之一充分統計量。又試判定$T$是否有完備性。

- $$f(x;\theta) = \theta \exp(x \log(1-\theta)) =: h(x)c(\theta)\exp(t(x)w(\theta)),$$
  belongs to the 1-dimensional exponential family, where $h(x) = 1$, $x = 0, 1, 2, \cdots$; $c(\theta) = \theta/(1-\theta)$; $t(x) = x$; $w(\theta) = \log(1-\theta)$.
- $T(\boldsymbol{X}) = \sum_{i=1}^n X_i$ is a S.S.
- $C := \{\log(1-\theta), \theta \in (0,1)\} \subset \mathbb{R}$, contains an open interval in $\mathbb{R}$. So, $T(\boldsymbol{X})$ is a C.S.S. by 課本 定理 **3.2**.

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 $T(\boldsymbol{X}) = (\sum_{i=1}^{n} t_1(X_i), \cdots, \sum_{i=1}^{n} t_k(X_i))$ 為一完備充份統計量。

By definition...

Definition

設 $T := T(\boldsymbol{X})$ 為一統計量,T 之 pdf 為 $f(t;\theta)$, $\theta \in \Omega$. 對一函數 g, 若 $\mathbb{E}_\theta[g(T)] = 0, \ \ \forall\theta \in \Omega$, 則 $\mathbb{P}(g(T) = 0) = 1, \ \ \forall\theta \in \Omega$ i.e., $g(T) = 0$ almost surely。故稱 T 為一完備統計量。

- $T = \sum_{i=1}^{n} X_i \sim NB(n, \theta)$, i.e., $f_T(t|\theta) = \binom{t+n-1}{n-1}\theta^n(1-\theta)^t$, for $t = 0, 1, 2, \cdots$ (you can use MGF to prove it).
- $0 = \mathbb{E}_\theta[g(T)] = \sum_{t=0}^{\infty} g(t)\binom{t+n-1}{n-1}\theta^n(1-\theta)^t = \theta^n \sum_{t=0}^{\infty} a_t u^t < \infty \ \ \forall\theta$, where $a_t := g(t)\binom{t+n-1}{n-1}$ and $u := 1-\theta \in (0,1)$.
- $g(t)$ must be 0 for all $t \geq 0$ for the power series to sum to zero. That is, $\mathbb{P}(g(T) = 0) = 1 \ \forall\theta \in (0,1)$.
- So, T is complete.

§5.3 #1, #2, #11, #18, #19

18. 設$X_1, \cdots, X_n$為一組由$\mathcal{U}(\theta, 2\theta)$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。

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

  which is free of θ iff
  $$x_{(1)} = y_{(1)},\ \ x_{(n)} = y_{(n)}.$$
  Let $T(\boldsymbol{x}) = (x_{(1)}, x_{(n)})$ and $T(\boldsymbol{y}) = (y_{(1)}, y_{(n)})$.
- So, $T(\boldsymbol{X}) = (X_{(1)}, X_{(n)})$ is a M.S.S for θ.

- $T = (X_{(1)}, X_{(n)})$ is not complete. Consider $g(T) = R - \mathbb{E}_\theta[R]$, with $R = X_{(n)} - X_{(1)}$ which is an A.S. (and is not independent of T).
- $\mathbb{E}_\theta[g(T)] = \mathbb{E}_\theta[R] - \mathbb{E}_\theta[R] = 0$, but $\mathbb{P}(g(T) \neq 0) > 0$. So, T is not complete.
- 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, $\mathbb{P}(U(\boldsymbol{X}) \le u) = \mathbb{P}(X_{(n)}/X_{(1)} \le u) = \mathbb{P}(Z_{(n)}/Z_{(1)} \le u) = \mathbb{P}(U(\boldsymbol{Z}) \le u)$, 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.

§5.3 #1, #2, #11, #18, #19

19. 設$X_1, \cdots, X_n$爲一組由$\mathcal{P}(\theta)$分佈所產生之隨機樣本, $\theta = 1, 2$。試證此分佈族並無完備性(本題可與例3.6比較)。

- Need to find a counter-example, which is a function g such that $\mathbb{E}_\theta[g(T)] = 0$, but $g(T) \neq 0$ for some θ.

- For $\theta = 1$,

$$0 = \mathbb{E}[g(T)|\theta = 1] = \sum_{t=0}^{\infty} g(t) 1^t e^{-1}/t!$$

- For $\theta = 2$,

$$0 = \mathbb{E}[g(T)|\theta = 2] = \sum_{t=0}^{\infty} g(t) 2^t e^{-2}/t!$$

- Consider

$$g(t) = \begin{cases} 2, & t = 0, 2 \\ -3, & t = 1 \\ 0, & o.w. \end{cases}$$

Then,

$$\sum_{t=0}^{\infty} g(t)/t! = g(0)/0! + g(1)/1! + g(2)/2! = 2 - 3 + 1 = 0;$$

$$\sum_{t=0}^{\infty} 2^t g(t)/t! = g(0)/0! + 2g(1)/1! + 2^2 g(2)/2! = 2 - 6 + 4 = 0.$$

That is, $\mathbb{E}[g(T)|\theta = 1] = 0$ and $\mathbb{E}[g(T)|\theta = 2] = 0$. But, $g(t) \neq 0$ for $\theta \in \{1, 2\}$.