Homework#1 Statistical Inference I

Name: Shih Jia-Han

Example 6.23

(i) Theorem 6.22 proves the completeness of

(a) X for the binominal family $\{ b(p,n), 0<p<1 \}$.

(b) X for the Poisson family $\{ P(\lambda), \lambda>0 \}$.

Solution:

(a)

For $X \sim b(p,n)$, $0<p<1$.

$$
\begin{aligned}
P(X=x) &= \binom{n}{x} p^x (1-p)^{n-x} \\
&= \binom{n}{x} \left(\frac{p}{1-p} \right)^x (1-p)^n \\
&= \binom{n}{x} \exp\left\{ x \log\left(\frac{p}{1-p} \right) + n\log(1-p) \right\}.
\end{aligned}
$$

Let

$$
\eta = \log\left(\frac{p}{1-p} \right),\ T(x)=x,\ A(\eta) = n\log(1+e^{\eta}) \text{ and } h(x) = \binom{n}{x}.
$$

Hence

$$
p(x|\eta) = \exp\{ \eta T(x) - A(\eta) \} h(x)
$$

is the canonical form of an one-dimensional exponential family with

$$
\chi = \{ 0,1,\cdots,n \} \text{ and } \Theta = \{ \eta; \eta \in (-\infty,\infty) \}.
$$

Since $\Theta = \{ \eta; \eta \in (-\infty,\infty) \}$ contains an one-dimensional open rectangle (e.g., $(0,1)$ is an one-dimensional open rectangle contained in Θ). Hence it is full rank. By Theorem 6.22, $T(x)=x$ is complete.

(b)

For $X \sim P(\lambda)$, $\lambda > 0$.

$$P(X = x) = \frac{1}{x!}\lambda^x e^{-\lambda}$$
$$= \frac{1}{x!}\exp\{x\log\lambda - \lambda\}.$$

Let

$$\eta = \log\lambda,\ T(x) = x,\ A(\eta) = e^{\eta} \text{ and } h(x) = \frac{1}{x!}.$$

Hence

$$p(x\,|\,\eta) = \exp\{\eta T(x) - A(\eta)\}h(x)$$

is the canonical form of an one-dimensional exponential family with

$$\chi = \{0, 1, \cdots\} \text{ and } \Theta = \{\eta; \eta > 0\}.$$

Since $\Theta = \{\eta; \eta > 0\}$ contains an one-dimensional open rectangle (e.g., $(1, 2)$ is an one-dimensional open rectangle contained in Θ). Hence it is full rank. By Theorem 6.22, $T(x) = x$ is complete.

(ii) *Uniform*. Let $X_1, \cdots, X_n$ be iid according to the uniform distribution $U(0, \theta)$, $\theta > 0$.

Solution:

For $X_1, \cdots, X_n \sim U(0,1)$.

$$
\begin{aligned}
P_\theta(x) &= \prod_{i=1}^{n} P_\theta(x_i) = \prod_{i=1}^{n} \frac{1}{\theta} I_{(0<x_i<\theta)} \\
&= \frac{1}{\theta^n} I_{(0<x_1, \cdots, x_n<\theta)} = \frac{1}{\theta^n} I_{(0<x_{(1)}, x_{(n)}<\theta)} \\
&= \frac{I_{(x_{(n)}<\theta)}}{\theta^n} I_{(0<x_{(1)})}.
\end{aligned}
$$

Hence $T = X_{(n)}$ is sufficient by the factorization criterion with $g_\theta(t) = \frac{I_{(t<\theta)}}{\theta^n}$ and $h(x) = I_{(0<x_{(1)})}$.

The cumulative distribution function of $T = X_{(n)}$ is

$$
\begin{aligned}
F_T(t) &= P(T \le t) = P(X_{(n)} \le t) \\
&= P(X_1, \cdots, X_n \le t) \\
&= P(X_1 \le t)^n \\
&= \left(\frac{t}{\theta}\right)^n,
\end{aligned}
$$

with $0 < t < \theta$.

The probability density function is

$$
p_\theta(t) = \frac{d}{dt} F_T(t) = \frac{d}{dt}\left(\frac{t}{\theta}\right)^n = \frac{nt^{n-1}}{\theta^n},
$$

with $0 < t < \theta$.

Suppose $E_\theta f(T) = 0$ for all θ.

$$\begin{aligned}
&E_\theta f(T) = 0 \\
&\Rightarrow \int_0^\theta \frac{nt^{n-1}}{\theta^n} f(t)dt = 0 \\
&\Rightarrow \int_0^\theta t^{n-1} f(t)dt = 0.
\end{aligned}$$

Let f^+ and f^- be the positive and negative part of f, respectively. Therefore,

$$\begin{aligned}
&\int_0^\theta t^{n-1} f(t)dt = 0 \\
&\Rightarrow \int_0^\theta t^{n-1}\{ f^+(t) - f^-(t) \}dt = 0 \\
&\Rightarrow \int_0^\theta t^{n-1} f^+(t)dt = \int_0^\theta t^{n-1} f^-(t)dt \\
&\Rightarrow \frac{d}{dt}\int_0^\theta t^{n-1} f^+(t)dt = \frac{d}{dt}\int_0^\theta t^{n-1} f^-(t)dt \\
&\Rightarrow \theta^{n-1} f^+(\theta) - 0 = \theta^{n-1} f^-(\theta) - 0 \\
&\Rightarrow f^+(\theta) = f^-(\theta) \\
&\Rightarrow f^+(t) = f^-(t).
\end{aligned}$$

for all t. This implies $f = 0$, for all t.

Therefore, $T = X_{(n)}$ satisfied (6.12), hence T is complete for θ.

(iii) *Exponential*. Let $Y_1, \cdots, Y_n$ be iid according to the exponential distribution $E(\eta, 1)$.

Solution:

For $Y_1, \cdots, Y_n \sim E(\eta, 1)$. The probability density function of Y_i is

$$f_{Y_i}(y) = e^{-(y-\eta)} I_{(\eta < y < \infty)}, \quad i = 1, 2, \cdots, n.$$

If $X_i = e^{-Y_i}$ and $\theta = e^{-\eta}$. We have

$$x_i = e^{-y_i} \Rightarrow y_i = -\log x_i \Rightarrow |J| = 1/x_i.$$

Then

$$\begin{aligned} f_{X_i}(x_i) &= f_{Y_i}(-\log x_i)/x_i = \frac{e^{\log x_i + \eta}}{x_i} I_{(\eta < -\log x_i < \infty)} \\ &= e^{\eta} I_{(-\eta > \log x_i > -\infty)} = \frac{1}{e^{-\eta}} I_{(e^{-\eta} > x_i > 0)} \\ &= \frac{1}{\theta} I_{(\theta > x_i > 0)}. \end{aligned}$$

for $i = 1, 2, \cdots, n$.

Therefore, $X_1, \cdots, X_n \sim U(0, \theta)$. By (ii), $T = X_{(n)}$ is sufficient and complete for θ.

Since $f(x) = -\log x$ is monotone and decrease in x. Therefore, $Y_{(1)} = -\log X_{(n)}$.

Hence $Y_{(1)}$ is also sufficient and complete for η.

Exercise 6.18

Show that the statistic $X_{(1)}$ and $\sum[X_i - X_{(1)}]$ of Problem 6.17(c) are independently distributed as $E(a, b/n)$ and $bGamma(n-1, 1)$, respectively.

Solution:

For $X_1, \cdots, X_n \sim E(a, b)$, the cumulative distribution function of $T = X_{(1)}$ is

$$
\begin{aligned}
F_T(t) &= P(T \le t) = P(X_{(1)} \le t) = 1 - P(X_{(1)} > t) \\
&= 1 - P(X_1, \cdots, X_n > t) = 1 - P(X_1 > t)^n \\
&= 1 - \left(e^{\frac{-(x-a)}{b}} \right)^n = 1 - e^{\frac{-n(x-a)}{b}} \\
&= 1 - e^{\frac{-(x-a)}{b/n}} .
\end{aligned}
$$

Therefore, the distribution of $T = X_{(1)}$ is $E(a, b/n)$.

Since exponential distribution is location-scale family. We have

$$X_i = a + bX_i',$$

where $X_i' \sim E(0, 1)$.

The joint probability density function of order statistics $X_{(1)}', \cdots, X_{(n)}'$ is

$$
\begin{aligned}
f_{X_{(1)}' \cdots X_{(n)}'}(x_{(1)}', \cdots, x_{(n)}') &= n!\, f_{X'}(x_{(1)}') \cdots f_{X'}(x_{(n)}') \\
&= n! e^{-(x_{(1)}' + \cdots + x_{(n)}')} .
\end{aligned}
$$

Let

$$
\begin{aligned}
y_1 &= n x_{(1)}' \\
y_2 &= (n-1)(x_{(2)}' - x_{(1)}') \\
&\vdots \\
y_n &= x_{(n)}' - x_{(n-1)}'
\end{aligned}
$$

Thus,

$$y_1 + \cdots + y_n = x_{(1)}' + \cdots + x_{(n)}'$$

and

$$J=\begin{vmatrix} n & 0 & 0 & 0 & \cdots & 0 \\ n-1 & n-1 & 0 & 0 & \cdots & 0 \\ 0 & n-2 & n-2 & 0 & \cdots & 0 \\ 0 & 0 & n-3 & n-3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{vmatrix}=n!$$

Therefore, the joint probability density function of $Y_1,\cdots,Y_n$ is

$$\begin{aligned} f_{Y_1\cdots Y_n}(y_1,\cdots,y_n) &= n!e^{-(y_1+\cdots+y_n)}\frac{1}{n!} \\ &= e^{-(y_1+\cdots+y_n)} \\ &= e^{-y_1}\times e^{-y_2}\times\cdots\times e^{-y_n}. \end{aligned}$$

Hence we obtained that $Y_i=(n-i+1)(X'_{(i)}-X'_{(i-1)})$ are identically independent distributed as $E(0,1)$, for $i=1,\cdots,n$ and define $X'_{(0)}=0$.

Then

$$\begin{aligned} &\sum_{i=2}^{n}(n-i+1)[X_{(i)}-X_{(i-1)}] \\ &=\{(n-1)[X_{(2)}-X_{(1)}]+(n-2)[X_{(3)}-X_{(2)}]+\cdots+[X_{(n)}-X_{(1)}]\} \\ &=\{X_{(n)}+\cdots+X_{(2)}-(n-1)X_{(1)}\}=\{X_{(n)}+\cdots+X_{(2)}+X_{(1)}-nX_{(1)}\} \\ &=\sum_{i=1}^{n}[X_i-X_{(1)}] \end{aligned}$$

and

$$\begin{aligned} (n-i+1)[X_{(i)}-X_{(i-1)}] &= (n-i+1)[a+bX'_{(i)}-\{a+bX'_{(i-1)}\}] \\ &= b(n-i+1)[X'_{(i)}-X'_{(i-1)}] \\ &= bY_i. \end{aligned}$$

Therefore,

$$\begin{aligned} \sum_{i=1}^{n}[X_i-X_{(1)}] &= \sum_{i=2}^{n}(n-i+1)[X_{(i)}-X_{(i-1)}] \\ &= b\sum_{i=2}^{n}Y_i. \end{aligned}$$

Since

$$Y_i \sim E(0,1) \Rightarrow \sum_{i=2}^{n} Y_i \sim Gamma(n-1,1).$$

Hence

$$\sum_{i=1}^{n} [X_i - X_{(1)}] = b\sum_{i=2}^{n} Y_i \sim bGamma(n-1,1).$$

Proof of independence

Since $Y_i = (n-i+1)(X'_{(i)} - X'_{(i-1)})$ for $i = 1, \cdots, n$ are independent.

Therefore,

$$a + \frac{b}{n} Y_1 = a + bX'_{(1)} = X_{(1)}$$

and

$$b\sum_{i=2}^{n} Y_i = \sum_{i=1}^{n} [X_i - X_{(1)}]$$

are also independent.

Thus, we proved that $X_{(1)}$ and $\sum [X_i - X_{(1)}]$ are independent.

Problem 6.34

Suppose that $X_1, \cdots, X_n$ are an iid sample from a location-scale family with distribution function $F((x-a)/b)$.

(a) If b is known, show that the difference $(X_1 - X_i)/b$, $i = 2, \cdots, n$ are ancillary.

(b) If a is known, show that the difference $(X_1 - a)/(X_i - a)$, $i = 2, \cdots, n$ are ancillary.

(c) If neither a or b are known, show that the quantities $(X_1 - X_i)/(X_2 - X_i)$, $i = 3, \cdots, n$ are ancillary.

Solution:

(a)

Since $X_1, \cdots, X_n$ are an iid sample from a location-scale family with distribution function. Let

$$X_i = a + bY_i, \quad i = 1, 2, \cdots, n$$

The distribution function of Y_i are

$$\begin{aligned} F_{Y_i}(y_i) &= P(Y_i \le y_i) = P\left(\frac{X_i - a}{b} \le y_i\right) \\ &= P(X_i \le a + by_i) = F\left(\frac{a + by_i - a}{b}\right) \\ &= F(y). \end{aligned}$$

Therefore, the distribution of Y_i does not depend on a, b for $i = 1, 2, \cdots, n$

Let $T = (X_1 - X_i)/b$, the distribution function is

$$\begin{aligned} F_T(t) &= P(T \le t) = P\left(\frac{X_1 - X_i}{b} \le t\right) \\ &= P\left(\frac{a + bY_1 - a - bY_i}{b} \le t\right) = P(Y_1 - Y_i \le t) \\ &= F_{Y_1 - Y_i}(t). \end{aligned}$$

Therefore, $(X_1 - X_i)/b = Y_1 - Y_i$.

Since Y_i does not depend on a, b for $i = 1, 2, \cdots, n$. Hence, $Y_1 - Y_i$ also does not depend on a, b for $i = 2, \cdots, n$. Hence $(X_1 - X_i)/b$, $i = 2, \cdots, n$ are ancillary.

(b)

Let $S = (X_1 - a)/(X_i - a)$, the distribution function is

$$
\begin{aligned}
F_S(s) = P(S \le s) &= P\left(\frac{X_1 - a}{X_i - a} \le s\right) \\
&= P\left(\frac{a + bY_1 - a}{a + bY_i - a} \le s\right) = P\left(\frac{Y_1}{Y_i} \le s\right) \\
&= F_{Y_1/Y_i}(s).
\end{aligned}
$$

Therefore, $(X_1 - a)/(X_i - a) = Y_1 / Y_i$.

Since Y_i does not depend on a, b for $i = 1, 2, \cdots, n$. Hence, Y_1 / Y_i also does not depend on a, b for $i = 2, \cdots, n$. Hence $(X_1 - a)/(X_i - a)$, $i = 2, \cdots, n$ are ancillary.

(c)

Let $K = (X_1 - X_i)/(X_2 - X_i)$, the distribution function is

$$
\begin{aligned}
F_K(k) = P(K \le k) &= P\left(\frac{X_1 - X_i}{X_2 - X_i} \le k\right) \\
&= P\left(\frac{a + bY_1 - a - bY_i}{a + bY_2 - a - bY_i} \le k\right) = P\left(\frac{Y_1 - Y_i}{Y_2 - Y_i} \le k\right) \\
&= F_{(Y_1 - Y_i)/(Y_2 - Y_i)}(k).
\end{aligned}
$$

Therefore, $(X_1 - X_i)/(X_2 - X_i) = (Y_1 - Y_i)/(Y_2 - Y_i)$.

Since Y_i does not depend on a, b for $i = 1, 2, \cdots, n$. Hence, $(Y_1 - Y_i)/(Y_2 - Y_i)$ also does not depend on a, b for $i = 3, \cdots, n$. Hence $(X_1 - X_i)/(X_2 - X_i)$, $i = 3, \cdots, n$ are ancillary.

If $X_1, \cdots, X_n$ are random sample from $N(\mu, \sigma^2)$, then

$$\sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2 \sim \chi^2_{df=n-1}$$

Proof:

Without loss of generality, let $\mu = 0, \sigma^2 = 1$. It is easy to obtain

$$\overline{X}_{n+1} = \frac{1}{n+1}(X_{n+1} + n\overline{X}_n). \qquad (1)$$

Now, let

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2.$$

Then, use Equation (1)

$$\begin{aligned}
nS_{n+1}^2 &= \sum_{i=1}^{n+1}(X_i - \overline{X}_{n+1})^2 = \sum_{i=1}^{n+1}\left(X_i - \left(\frac{X_{n+1} + n\overline{X}_n}{n+1}\right)\right)^2 \\
&= \sum_{i=1}^{n+1}\left(X_i - \frac{n\overline{X}_n}{n+1} - \frac{X_{n+1}}{n+1}\right)^2 = \sum_{i=1}^{n+1}\left(X_i - \overline{X}_n + \frac{\overline{X}_n}{n+1} - \frac{X_{n+1}}{n+1}\right)^2 \\
&= \sum_{i=1}^{n+1}\left((X_i - \overline{X}_n)^2 + \left(\frac{\overline{X}_n}{n+1} - \frac{X_{n+1}}{n+1}\right)^2 + 2(X_i - \overline{X}_n)\left(\frac{\overline{X}_n}{n+1} - \frac{X_{n+1}}{n+1}\right)\right) \\
&= \sum_{i=1}^{n}(X_i - \overline{X}_n)^2 + (X_{n+1} - \overline{X}_n)^2 + \frac{1}{n+1}(\overline{X}_n - X_{n+1})^2 \\
&\qquad + 2\sum_{i=1}^{n+1}(X_i - \overline{X}_n)\left(\frac{\overline{X}_n - X_{n+1}}{n+1}\right) \\
&= \sum_{i=1}^{n}(X_i - \overline{X}_n)^2 + (\overline{X}_n - X_{n+1})^2 + \frac{1}{n+1}(\overline{X}_n - X_{n+1})^2 - \frac{2}{n+1}(\overline{X}_n - X_{n+1})^2 \\
&= (n-1)S_n^2 + \frac{n}{n+1}(\overline{X}_n - X_{n+1})^2.
\end{aligned}$$

Hence we have derived

$$nS_{n+1}^2 = (n-1)S_n^2 + \frac{n}{n+1}(\overline{X}_n - X_{n+1})^2.$$

It is the same as

$$(n-1)S_n^2 = (n-2)S_n^2 + \frac{n-1}{n}(\overline{X}_{n-1} - X_n)^2. \qquad (2)$$

Consider Equation (2) with $n = 2$, we have

$$S_2^2 = \left(\frac{X_1 - X_2}{\sqrt{2}} \right)^2 \sim \chi^2_{df=1} .$$

For $n = k$, assume

$$(k-1)S_k^2 \sim \chi^2_{df=k-1} .$$

For $n = k+1$,

$$kS_{k+1}^2 = (k-1)S_k^2 + \left(\frac{k}{k+1} \right)(\overline{X}_k - X_{k+1})^2 .$$

Since

$$(k-1)S_k^2 \sim \chi^2_{df=k-1} .$$

Hence we only need to check $\left(\frac{k}{k+1} \right)(\overline{X}_k - X_{k+1})^2 \sim \chi^2_{df=1}$, independent with S_k^2 then by mathematical induction, $kS_{k+1}^2 \sim \chi^2_{df=k}$ is proved.

Proof of independence:

Since the vector $(X_{k+1}, \overline{X}_k)$ is independent of S_k^2. Hence $(\overline{X}_k - X_{k+1})^2$ is also independent of S_k^2.

Proof of distribution:

Since $X_{k+1} \sim N(0,1)$ and $\overline{X}_k \sim N(0,1/k)$ are independent. Therefore,

$$\overline{X}_k - X_{k+1} \sim N(0,(k+1)/k) .$$

Hence

$$\left(\frac{k}{k+1} \right)(\overline{X}_k - X_{k+1})^2 \sim \chi^2_{df=1} .$$

is proved.

Therefore, by mathematical induction

$$(n-1)S_n^2 = \sum_{i=1}^{n}(X_i - \overline{X}_n)^2 \sim \chi^2_{df=n-1}.$$

Hence we have proved

$$\sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2 \sim \chi^2_{df=n-1}$$