

Solution

1. Let \((X_1, \ldots, X_n)\) be a simple random sample from a distribution with probability density function given by

\[
f(x; \theta, \beta) = \frac{1}{\theta^\beta} \left( \frac{x}{\theta} \right)^{1-\beta}, \quad 0 \leq x \leq \theta, \quad \theta > 0, \beta < 1.
\]

(i) Find a minimal sufficient statistic for \((\theta, \beta)\) and give your justification.

The joint pdf of \(X = (X_1, \ldots, X_n)\) is given by

\[
f_n(X; \beta, \theta) = \frac{1}{\theta^{n/\beta} \beta^n} \left( \prod_{i=1}^n X_i \right)^{1-\beta} I_{[0,\theta]}(X_{(n)}).
\]

By the factorization theorem, \((\prod_{i=1}^n X_i, X_{(n)})\) is sufficient. Further, if \(X = (X_1, \ldots, X_n)\) and \(Y = (Y_1, \ldots, Y_n)\) such that \(f_n(X; \beta, \theta) = f_n(Y; \beta, \theta)\) for all \(\beta\) and \(\theta\), we then have

\[
\left( \prod_{i=1}^n X_i \right)^{1-\beta} I_{[0,\theta]}(X_{(n)}) = \left( \prod_{i=1}^n Y_i \right)^{1-\beta} I_{[0,\theta]}(Y_{(n)})
\]

for all \(\beta\) and \(\theta\), which implies \((\prod_{i=1}^n X_i, X_{(n)}) = (\prod_{i=1}^n Y_i, Y_{(n)})\). Hence \((\prod_{i=1}^n X_i, X_{(n)})\) is minimal sufficient.

(ii) Suppose that \(\beta\) is known. Show that \(X_{(n)}\) is sufficient and complete for \(\theta\), where \(X_{(n)}\) is the largest order statistic, i.e., \(X_{(n)} = \max_{1 \leq i \leq n} X_i\).

By the factorization theorem, \(X_{(n)}\) is sufficient. The pdf of \(X_{(n)}\) is given by

\[
f_{X_{(n)}}(x) = \frac{n}{\theta^\beta} \left( \frac{x}{\theta} \right)^{n-1} I_{[0,\theta]}(x).
\]

Let \(h\) be any Borel function such that \(E h(X_{(n)}) = 0\). Then we have

\[
\int_0^\theta h(x) x^{n-1} dx = 0,
\]
for all $\theta$. Differentiating both sides w.r.t. $\theta$ yields

$$h(\theta)\theta^{n/\beta-1} = 0 \text{ for all } \theta > 0.$$ 

This implies $h \equiv 0$. Therefore $X_{(n)}$ is complete.

(iii) Suppose that $\theta$ is known. Find a sufficient and complete statistic for $\beta$ and give your justification.

The joint pdf of $X$ can be expressed as

$$\exp\left\{\frac{1 - \beta}{\beta} \sum_{i=1}^{n} \ln X_i - \xi(\beta)\right\} h(X)$$

for some function $\xi$ and $h$. Hence, it is an exponential family with natural sufficient statistic $\sum_{i=1}^{n} \ln X_i$. Hence $\sum_{i=1}^{n} \ln X_i$ is sufficient and complete for $\eta = (1 - \beta)/\beta$. Since $\eta$ and $\beta$ are one-to-one, $\sum_{i=1}^{n} \ln X_i$ is also sufficient and complete for $\beta$.

(iv) Suppose that $\theta$ is known. Drive the uniformly minimum variance unbiased estimator (UMVUE) for $\beta$.

We can compute

$$E \ln X_i = \frac{1}{\theta^{1/\beta}} \int_{0}^{\theta} \ln(x)x^{1/\beta-1}dx = \ln \theta - \beta.$$ 

Take $\hat{\beta} = \ln \theta - \frac{1}{n} \sum_{i=1}^{n} \ln X_i$. Since $\hat{\beta}$ is a function of the S&C statistic and is unbiased for $\beta$, it is the UMVUE of $\beta$.

2. Let $(X_1, \ldots, X_n)$ be a random sample from the exponential distribution $\mathcal{E}(a, \theta)$ with parameter $-\infty < a < \infty$, $\theta > 0$. The probability density function of the distribution is given by

$$f(x; a, \theta) = \frac{1}{\theta} e^{-\frac{x-a}{\theta}} I_{(a, \infty)}(x).$$

Let $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ and $X_{(1)} = \min_{1 \leq i \leq n}\{X_i\}.$
(i) Show that \( \sqrt{n}(X_{(1)} - a) \to 0 \) in probability.

\( X_{(1)} \) is distributed as \( E(a, \theta/n) \) with mean \( a + \theta/n \) and variance \((\theta/n)^2\). By Chebychev’s inequality,

\[
P(\sqrt{n}|X_{(1)} - EX_{(1)}| > \epsilon) \leq \frac{n\text{Var}(X_{(1)})}{\epsilon^2} = \frac{\theta^2}{n\epsilon^2} \to 0.
\]

Hence \( \sqrt{n}(X_{(1)} - EX_{(1)}) \to_p 0 \). Note that \( \sqrt{n}(EX_{(1)} - a) = \frac{\theta}{\sqrt{n}} \to 0 \). Then, by continuous mapping theorem, \( \sqrt{n}(X_{(1)} - a) \to_p 0 \).

(ii) Show that \( \bar{X} - X_{(1)} \to \theta \) in probability.

By the LLN, \( \bar{X} \to_p a + \theta \). From (i) \( X_{(1)} \to_p a \). By continuous mapping theorem, \( \bar{X} - X_{(1)} \to \theta \).

(iii) Show that \( \frac{\sqrt{n}(X_{(1)} - a - \theta)}{\theta} \to Z \) in distribution, where \( Z \) is a standard normal random variable.

It follows directly from CLT since \( E\bar{X} = a + \theta \) and \( \text{Var}(\bar{X}) = \theta^2/n \).

(iv) Show that \( \frac{\sqrt{n}(\bar{X} - X_{(1)} - a - \theta)}{X_{(1)} - \bar{X}} \to Z \) in distribution as well.

From (i), \( \sqrt{n}(X_{(1)} - a)/\theta \to_p 0 \). By (iii) and Slutsky’s theorem

\[
\frac{\sqrt{n}(\bar{X} - X_{(1)} - \theta)}{\theta} = \frac{\sqrt{n}(\bar{X} - a - \theta)}{\theta} - \frac{\sqrt{n}(X_{(1)} - a)}{\theta} \to_d Z.
\]

From (ii), \( X_{(1)} - \bar{X} \to_p \theta \). By Slutsky’s theorem again,

\[
\frac{\sqrt{n}(\bar{X} - X_{(1)} - \theta)}{X_{(1)} - \bar{X}} \to Z.
\]

3. Let \((X_1, \ldots, X_n)\) be a random sample from the Gamma distribution \( \Gamma(\alpha, \gamma) \) where \( \alpha \) is known. The probability density function of the distribution is given by

\[
f(x; a, \theta) = \frac{1}{\Gamma(\alpha)\gamma^\alpha x^{\alpha-1}e^{-x/\gamma} I_{(0,\infty)}(x)}.
\]
(i) Derive the UMVUE of $\gamma$.

When $\alpha$ is known, the family is an exponential family and $T = \sum_{i=1}^{n} X_i$ is sufficient and complete for $\gamma$. Since $E(X_i) = \alpha \gamma$, $E(T) = n\alpha \gamma$. Hence $\hat{\gamma} = \frac{1}{n\alpha} T$ is the UMVUE of $\gamma$.

(ii) Let $F(t; \alpha, \gamma)$ be the cumulative distribution function of $\Gamma(\alpha, \gamma)$. Derive the UMVUE of $F(c; \alpha, \gamma)$ where $c > 0$ is fixed.

The UMVUE of $F(c; \alpha, \gamma)$ is given by

$$P(X_1 \leq c|T) = P(\frac{X_1}{T} \leq \frac{c}{T}) = P(\frac{X_1}{T} \leq \frac{c}{t}),$$

since $X_1/T$ is ancilary and hence independent from $T$, the sufficient and complete statistic. Note that, by the property of the Gamma distribution, $X_1/T \sim \text{Beta}(\alpha, (n-1)\alpha)$). Hence, the UMVUE is given by

$$\hat{F}(c; \alpha, \gamma) = \frac{\Gamma(n\alpha)}{\Gamma(\alpha)\Gamma((n-1)\alpha)} \int_{0}^{c/T} x^{\alpha-1}(1-x)^{(n-1)\alpha-1} dx.$$

(iii) Derive the UMVUE of $\frac{dF(t; \alpha, \gamma)}{dt}$, the probability density function, at a fixed $t > 0$.

Note that

$$\frac{d}{dc} P(X_1 \leq c) = \frac{d}{dc} EP(X_1 \leq c|T) = E\frac{d}{dc} P(X_1 \leq c|T).$$

The UMVUE is given by

$$\frac{d}{dc} P(X_1 \leq c|T) = \frac{d}{dc} \frac{\Gamma(n\alpha)}{\Gamma(\alpha)\Gamma((n-1)\alpha)} \int_{0}^{c/T} x^{\alpha-1}(1-x)^{(n-1)\alpha-1} dx$$

$$= \begin{cases} \frac{1}{T} \frac{\Gamma(n\alpha)}{\Gamma(\alpha)\Gamma((n-1)\alpha)} \left(\frac{c}{T}\right)^{\alpha-1}(1 - \frac{c}{T})^{(n-1)\alpha-1}, & c < T, \\ 0, & c \geq T. \end{cases}$$

4. Let $X = (X_1, \ldots, X_n)$ be a random sample from a double exponential distribution $DE(0, \theta)$ with probability density function given by

$$f(x; \theta) = \frac{1}{2\theta} e^{-|x|/\theta} I_{(-\infty, \infty)}(x).$$
(i) Derive the variance of the double exponential distribution without using neither integration nor moment generating function. You can use the fact that $E|X_i| = \theta$.

The pdf of the distribution can be expressed in the canonical exponential family form:

$$f(x; \eta) = e^{-\frac{|x|}{\eta} + \ln(\eta)} \frac{1}{2} I_{(-\infty, \infty)}(x).$$

where $\eta = 1/\theta$. Hence $\text{Var}(|X_i|) = \text{Var}(-|X_i|) = \frac{d^2}{d\eta^2}(-\ln \eta) = 1/\eta^2 = \theta^2$. The distribution is symmetric about 0 and hence $EX_i = 0$ and $\text{Var}(X_i) = EX_i^2 = E|X_i|^2 = \text{Var}(|X_i|) + (E|X_i|)^2 = 2\theta^2$.

(ii) Find the UMVUE of $\theta$. Show that the UMVUE of $\theta$ achieves the Cramér-Rao lower bound.

Since $T = \sum_{i=1}^n |X_i|$ is sufficient and complete for $\theta$ and $ET = n\theta$, the UMVUE of $\theta$ is given by $\hat{\theta} = T/n$. From the solution of (i), $\text{Var}(T/n) = \theta^2/n$. On the other hand, the Fisher information number about $\theta$ is

$$I(\theta) = \text{Var} \left( \frac{d}{d\theta} \ln f_n(X, \theta) \right) = \text{Var} \left( \frac{T}{\theta^2} - \frac{1}{\theta} \right) = \frac{n\theta^2}{\theta^4} = \frac{n}{\theta^2}.$$

Hence $\text{Var}(\hat{\theta}) = I^{-1}(\theta)$.

(iii) Let $\xi = P_\theta(X_1 > t)$ where $t > 0$ is fixed. Derive the Fisher information number $I(\xi)$. Find the Cramér-Rao lower bound for unbiased estimators of $\xi$.

Note that $\xi = \xi(\theta) = \frac{1}{2} e^{-t/\theta}$. The Fisher information $I(\xi)$ about $\xi$ satisfies

$$I(\theta) = [\xi'(\theta)]^2 I(\xi) = \left[ \frac{t}{2\theta^2} e^{-t/\theta} \right]^2 I(\xi).$$

Hence

$$I(\xi) = \left[ \frac{t}{2\theta^2} e^{-t/\theta} \right]^2 \frac{n}{\theta^2} = \frac{4n\theta^2}{t^2} e^{2t/\theta} = \frac{n}{[\xi \ln(2\xi)]^2}.$$

The C-R lower bound for the unbiased estimators of $\xi$ is given by $I^{-1}(\xi) = [\xi'(\theta)]^2 I^{-1}(\theta)$. 
END OF PAPER