MTH375: Mathematical Statistics - Homework #4

Cason Konzer

February 10, 2022

Key Concepts: Rao-Blackwell Theorem, Exponential Families of Distributions, UMVUEs, Lehman-Sheffé Theorem.

1. Let $X_1, \ldots, X_n$ be a sample of *iid* $Binomial(k = 1, p)$ random variables.

(a) Write out the likelihood function $L(p|x_1, \ldots x_n)$, and find a sufficient statistic for p.

Solution:

We know that for each X_i its *pmf* is $f(x_i; p) = p^{x_i}(1-p)^{1-x_i}$.

With this we can compute the likelihood function $L(p|x_1, \ldots, x_n) = f(x_1; p) \cdots f(x_n; p)$.

- $L(p) = p^{x_1}(1-p)^{1-x_1} \cdots p^{x_n}(1-p)^{1-x_n}$.
- $L(p) = p^{x_1+.+.+x_n}(1-p)^{n-x_1-.-.-x_n}$.
- $L(p) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$.

We can now see that $T = \sum_{i=1}^{n} x_i$ is a sufficient statistic for p.

(b) Use your answer to (a) to find an UMVUE for p.

Solution:

For each X_i we have $E(X_i) = kp = p$.

Thus ...

- $E(T) = nE(X_i) = np$.
- $np \cdot \dfrac{1}{n} = p$.
- $T \cdot \dfrac{1}{n} = \overline{X}$.
- $E(\overline{X}) = E(T) \cdot \dfrac{1}{n} = np \cdot \dfrac{1}{n} = p$.

We can now see that $\overline{X}$ is an UMVUE for p.

(c) Evaluate $E(\overline{X}^2)$. Hint: Use $E(\overline{X})$ and $V(\overline{X})$.

Solution:

We know $V(\overline{X}) = E(\overline{X}^2) - E(\overline{X})^2$.

For each X_i we have $V(X_i) = kpq = p(1-p)$.

Thus ...

- $V(\overline{X}) = V\left(\dfrac{T}{n}\right) = \dfrac{V(T)}{n^2} = \dfrac{nV(X_i)}{n^2} = \dfrac{p(1-p)}{n}$.
- $E(\overline{X})^2 = p^2$.
- $E(\overline{X}^2) = V(\overline{X}) + E(\overline{X})^2 = \dfrac{p(1-p)}{n} + p^2$.

(d) Use your answer to (c) to find an UMVUE for p^2.

Solution:

Some simplification implies ...

- $E(\overline{X}^2) = \dfrac{p(1-p)}{n} + p^2 = \dfrac{p - p^2 + np^2}{n} = \dfrac{p + p^2(n-1)}{n}$.
- $E(\overline{X}) + p^2(n-1) = nE(\overline{X}^2)$.
- $p^2(n-1) = nE(\overline{X}^2) - E(\overline{X})$.
- $p^2 = \dfrac{nE(\overline{X}^2) - E(\overline{X})}{n-1} = E\left(\dfrac{n\overline{X}^2 - \overline{X}}{n-1}\right)$.

We can now see that $\dfrac{n\overline{X}^2 - \overline{X}}{n-1}$ is an UMVUE for p^2.

2. Let $X_1, \ldots, X_n$ be a sample of *iid* $Normal(\mu, 1)$ random variables. We know $\overline{X}$ is UMVUE for μ.

(a) Find an UMVUE for μ^2.

Solution:

We know $E(X_i) = \mu$, $E(\overline{X}) = \dfrac{nE(X_i)}{n} = \mu$, $V(X_i) = 1^2$ and $V(\overline{X}) = \dfrac{nV(X_i)}{n^2} = \dfrac{1}{n}$.

Thus $\overline{X} \sim N(\mu, \dfrac{1}{n})$.

From above $E(\overline{X}^2) = V(\overline{X}) + E(\overline{X})^2$.

- $E(\overline{X}^2) = \dfrac{1}{n} + \mu^2$.
- $E(\overline{X}^2) - \dfrac{1}{n} = \mu^2$.
- $E\left(\overline{X}^2 - \dfrac{1}{n}\right) = \mu^2$.

We can now see that $\overline{X}^2 - \dfrac{1}{n}$ is an UMVUE for μ^2.

(b) Use the MGF of $\overline{X}$ to compute $E(\overline{X}^3)$.

Solution:

We need to solve for the third moment of $\overline{X}$.

- $M_{\overline{X}}(t) = e^{\mu t + t^2/2}$.
- $M'_{\overline{X}}(t) = (\mu + t/2)e^{\mu t + t^2/2} = \mu e^{\mu t + t^2} + \dfrac{t}{2}e^{\mu t + t^2}$.
- $M^2_{\overline{X}}(t) = \mu(\mu + t/2)e^{\mu t + t^2} + \dfrac{1}{2}e^{\mu t + t^2} + \dfrac{t}{2}(\mu + t/2)e^{\mu t + t^2}$.
- $M^2_{\overline{X}}(t) = \left(\mu^2 + \dfrac{\mu t}{2}\right)e^{\mu t + t^2} + \dfrac{1}{2}e^{\mu t + t^2}\left(\dfrac{\mu t}{2} + \dfrac{t^2}{2}\right)e^{\mu t + t^2}$.
- $M^2_{\overline{X}}(t) = \mu^2 e^{\mu t + t^2} + \dfrac{\mu t}{2}e^{\mu t + t^2} + \dfrac{1}{2}e^{\mu t + t^2}\dfrac{\mu t}{2}e^{\mu t + t^2} + \dfrac{t^2}{2}e^{\mu t + t^2}$.
- $M^2_{\overline{X}}(t) = \mu^2 e^{\mu t + t^2} + \mu t e^{\mu t + t^2} + \dfrac{1}{2}e^{\mu t + t^2} + \dfrac{t^2}{2}e^{\mu t + t^2}$.
- $M^3_{\overline{X}}(t) = \mu^2(\mu + t/2)e^{\mu t + t^2} + \mu e^{\mu t + t^2} + \mu t(\mu + t/2)e^{\mu t + t^2} + \dfrac{1}{2}(\mu + t/2)e^{\mu t + t^2} + \dfrac{2t}{2}e^{\mu t + t^2} + \dfrac{t^2}{2}(\mu + t/2)e^{\mu t + t^2}$.
- $M^3_{\overline{X}}(t) = \left(\mu^3 + \dfrac{\mu^2 t}{2}\right)e^{\mu t + t^2} + \mu e^{\mu t + t^2} + \left(\mu^2 t + \dfrac{\mu t^2}{2}\right)e^{\mu t + t^2} + \left(\dfrac{\mu}{2} + \dfrac{t}{4}\right)e^{\mu t + t^2} + \dfrac{2t}{2}e^{\mu t + t^2} + \left(\dfrac{\mu t^2}{2} + \dfrac{t^3}{4}\right)e^{\mu t + t^2}$.
- $M^3_{\overline{X}}(0) = \mu^3 + \mu + \dfrac{\mu}{2} = \mu^3 + \dfrac{3\mu}{2} = E(\overline{X}^3)$.

(c) Use (b) to find an UMVUE for μ^3.

Solution:

We have $\mu^3 = E(\overline{X}^3) - \dfrac{3\mu}{2}$

Thus ...

- $\mu^3 = E(\overline{X}^3) - \dfrac{3E(\overline{X})}{2} = E\left(\overline{X}^3 - \dfrac{3\overline{X}}{2}\right)$

We can now see that $\overline{X}^3 - \dfrac{3\overline{X}}{2}$ is an UMVUE for μ^3.

3. Let $X_1, \ldots, X_n$ be a sample of *iid* random variables with common pdf $f(x_i;\theta)$.

$$f(x_i;\theta) = \frac{3x_i^2}{\theta^3} \quad \text{for} \quad 0 \le x_i \le \theta.$$

(a) the likelihood function $L(\theta|x_1, \ldots x_n)$, and find a sufficient statistic T for θ.

Solution:

We can compute the likelihood function $L(\theta|x_1, \ldots, x_n) = f(x_1;\theta) \cdots f(x_n;\theta)$.

- $L(\theta) = \dfrac{3x_1^2}{\theta^3} \cdots \dfrac{3x_n^2}{\theta^3} = \left(\dfrac{3}{\theta^3}\right)^n \displaystyle\prod_{i=1}^{n} x_i^2 \quad \text{for} \quad 0 \le x_i \le \theta.$

We can now see that $T = \max\{X_i, ..., X_n\}$ is a sufficient statistic for θ, as T provides the lower bound on θ.

(b) Find the pdf of T, $E(T)$, and an UMVUE for θ.

Solution:

We will first solve for the $CDFs$ of X_i and T then the requested ...

- $F_{X_i} = \displaystyle\int_0^{x_i} \frac{3x^2}{\theta^3}\, dx = \frac{x^3}{\theta^3}\Big|_0^{x_i} = \frac{x_i^3}{\theta^3}.$

- $F_T = P(T \le t) = P(X_1 \le t \;\& \cdots \&\; Xn \le t) = \Big(F_{X_i}\Big)^n = \Big(\dfrac{m^3}{\theta^3}\Big)^n = \dfrac{m^{3n}}{\theta^{3n}}.$

- $f_T = \dfrac{d}{dx}\Big[\dfrac{m^{3n}}{\theta^{3n}}\Big] = \dfrac{3nm^{3n-1}}{\theta^{3n}}.$

- $E(T) = \displaystyle\int_0^{\theta} \frac{3nm^{3n}}{\theta^{3n}}\, dm = \frac{3nm^{3n+1}}{(3n+1)\theta^{3n}}\Big|_0^{\theta} = \frac{3n\theta^{3n+1}}{(3n+1)\theta^{3n}} - \frac{3n(0)^{3n+1}}{(3n+1)\theta^{3n}} = \frac{3n\theta}{3n+1}.$

- $T^* = \dfrac{3n+1}{3n}T;\; E(T^*) = E\Big(\dfrac{3n+1}{3n}T\Big) = \dfrac{3n+1}{3n}E(T) = \theta$

We can now see that $T^* = \max\{X_i, ..., X_n\}\dfrac{3n+1}{3n}$ is an UMVUE for θ.

4. Let $X_1, \ldots, X_n$ be a sample of *iid* random variables with common pdf $f(x_i; \theta_1, \theta_2)$.

$$f(x_i; \theta_1, \theta_2) = \frac{1}{\theta_1} e^{-(x_i - \theta_2)/\theta_1} \quad \text{for} \quad x_i > \theta_2.$$

(a) Show that the pair $\big(\sum_{k=1}^{n} X_k, X_{(1)}\big)$ is sufficient for (θ_1, θ_2). We use the notation $X_{(1)} = \min\{X1, \ldots, Xn\}$, the 1^{st} order statistic.

Solution:

We can compute the likelihood function $L(\theta_1, \theta_2 | x_1, \ldots, x_n) = f(x_1; \theta_1, \theta_2) \cdots f(x_n; \theta_1, \theta_2)$.

- $L(\theta_1, \theta_2) = \frac{1}{\theta_1} e^{-(x_1 - \theta_2)/\theta_1} \cdots \frac{1}{\theta_1} e^{-(x_n - \theta_2)/\theta_1} = \frac{1}{\theta_1^n} \Big(e^{\theta_2/\theta_1 - x_1/\theta_1} \cdots e^{\theta_2/\theta_1 - x_n/\theta_1} \Big)$.
- $L(\theta_1, \theta_2) = \frac{1}{\theta_1^n} \Big(e^{n\theta_2/\theta_1 - \sum_{k=1}^{n} x_k/\theta_1} \Big) \quad \text{for} \quad x_i > \theta_2.$

As this likelihood function is an exponential family of ditributions, it is straightforward to see that the pair $\big(\sum_{k=1}^{n} X_k, X_{(1)}\big)$ is sufficient for (θ_1, θ_2).

For clarity, $X_{(1)}$ is in this pair as θ_2 has an upperbound of $X_{(1)}$ and appears in the term $\frac{n\theta_2}{\theta_1}$ within our family.

(b) Find a pair (T_1, T_2) which is an UMVUE for (θ_1, θ_2).

Solution:

We will first solve for $F(x_i; \theta_1, \theta_2)$.

- $F_{X_i} = \displaystyle\int_{\theta_2}^{x_i} \frac{1}{\theta_1} e^{\theta_2/\theta_1 - x/\theta_1}\, dx = -e^{\theta_2/\theta_1 - x/\theta_1}\Big|_{\theta_2}^{x_i} = e^{\theta_2/\theta_1 - \theta_2/\theta_1} - e^{\theta_2/\theta_1 - x_i/\theta_1} = 1 - e^{\theta_2/\theta_1 - x_i/\theta_1}$.

Next we will need $F_{X_{(1)}}$.

- $F_{X_{(1)}}(m) = P(\min\{X1, \ldots, Xn\} \leq m) = 1 - P(\min\{X1, \ldots, Xn\} > m)$.
- $F_{X_{(1)}}(m) = 1 - [P(X_i > m)]^n = 1 - [1 - F_{X_i}(m)]^n = 1 - (e^{\theta_2/\theta_1 - m/\theta_1})^n = 1 - e^{n\theta_2/\theta_1 - nm/\theta_1}$.

Following $f_{X_{(1)}}$ ...

- $f_{X_{(1)}}(m) = \dfrac{d}{dm}[F_{X_{(1)}}(m)] = \dfrac{d}{dm}\big[1 - e^{n\theta_2/\theta_1 - nm/\theta_1}\big] = \dfrac{n}{\theta_1} e^{n\theta_2/\theta_1 - nm/\theta_1}$.

Continuing, we will now solve for $E(X_{(1)})$.

- $E(X_{(1)}) = \displaystyle\int_{\theta_2}^{\infty} \frac{nm}{\theta_1} e^{n\theta_2/\theta_1 - nm/\theta_1}\, dm = -\frac{nm + \theta_1}{n} e^{n\theta_2/\theta_1 - nm/\theta_1}\Big|_{\theta_2}^{\infty}$.
- $E(X_{(1)}) = \dfrac{n\theta_2 + \theta_1}{n} e^{n\theta_2/\theta_1 - n\theta_2/\theta_1} - \infty e^{-\infty/\theta_1} = \dfrac{n\theta_2 + \theta_1}{n} e^{0} - \infty \dfrac{e^{1/\theta_1}}{e^{\infty}}$.
- $E(X_{(1)}) = \dfrac{n\theta_2 + \theta_1}{n} - \dfrac{\infty}{e^{\infty}} = \dfrac{n\theta_2 + \theta_1}{n} = \theta_2 + \dfrac{\theta_1}{n}$; As $\displaystyle\lim_{m \to \infty} \frac{m}{e^m} = 0$

We will now solve for $E(X_i)$.

- $E(X_i) = \displaystyle\int_{\theta_2}^{\infty} \frac{x_i}{\theta_1} e^{\theta_2/\theta_1 - x_i/\theta_1}\, dx = -(x_i + \theta_1) e^{\theta_2/\theta_1 - x_i/\theta_1}\Big|_{\theta_2}^{\infty}$.
- $E(X_i) = (\theta_2 + \theta_1) e^{\theta_2/\theta_1 - \theta_2/\theta_1} - (\infty + \theta_1) e^{\theta_2/\theta_1 - \infty/\theta_1} = (\theta_2 + \theta_1) e^{0} - \infty e^{\theta_2 - \infty/\theta_1}$.
- $E(X_i) = \theta_2 + \theta_1 - \infty e^{-\infty/\theta_1} = \theta_2 + \theta_1 - \dfrac{\infty e^{1/\theta_1}}{e^{\infty}} = \theta_2 + \theta_1 - \dfrac{\infty}{e^{\infty}}$.
- As $\displaystyle\lim_{x_i \to \infty} \frac{x_i}{e^{x_i}} = 0$; $E(X_i) = \theta_2 + \theta_1$.

We can now see $E\left(\displaystyle\sum_{k=1}^{n} X_k\right) = nE(X_k) = n\theta_2 + n\theta_1$ and $E(X_{(1)}) = \theta_2 + \dfrac{\theta_1}{n}$.

We will solve these two equations for solutions to θ_1 and θ_2 ...

- $n(\theta_1 + \theta_2) = E\Big(\sum_{k=1}^{n} X_k\Big)$; $\theta_1 + \theta_2 = E(\overline{X})$; $\theta_2 = E(\overline{X}) - \theta_1$.
- $\theta_2 + \dfrac{\theta_1}{n} = E(X_{(1)})$; $\theta_2 = E(X_{(1)}) - \dfrac{\theta_1}{n} = E(\overline{X}) - \theta_1 = \theta_2$.
- $\theta_1 - \dfrac{\theta_1}{n} = E(\overline{X}) - E(X_{(1)})$; $n\theta_1 - \theta_1 = nE(\overline{X}) - nE(X_{(1)})$.
- $\theta_1(n-1) = n\big(E(\overline{X}) - E(X_{(1)})\big)$; $\theta_1 = \dfrac{n\big(E(\overline{X}) - E(X_{(1)})\big)}{n-1}$.
- $\theta_1 = \dfrac{n\big(E(\overline{X} - X_{(1)})\big)}{n-1} = E\left(\dfrac{n(\overline{X} - X_{(1)})}{n-1}\right)$.
- $\theta_2 = E(\overline{X}) - \theta_1 = E(\overline{X}) - E\left(\dfrac{n(\overline{X} - X_{(1)})}{n-1}\right) = \dfrac{(n-1)E(\overline{X})}{n-1} - E\left(\dfrac{n(\overline{X} - X_{(1)})}{n-1}\right)$.
- $\theta_2 = E\left(\dfrac{(n-1)\overline{X}}{n-1}\right) - E\left(\dfrac{n(\overline{X} - X_{(1)})}{n-1}\right) = E\left(\dfrac{n\overline{X} - \overline{X} - n\overline{X} + nX_{(1)}}{n-1}\right)$.
- $\theta_2 = E\left(\dfrac{nX_{(1)} - \overline{X}}{n-1}\right)$.

We can now see that the pair $(T_1, T_2) = \left(\dfrac{n(\overline{X} - X_{(1)})}{n-1}, \dfrac{nX_{(1)} - \overline{X}}{n-1}\right)$ is an UMVUE for (θ_1, θ_2).