The University of Sydney
School of Mathematics and Statistics

Tutorial Week 6 Solution

STAT3023: Statistical Inference — Semester 2, 2022

1. (a) T is the sum of n iid Poisson(λ) distribution, hence $T \sim \text{Poisson}(n\lambda)$. Hence,

$$
\begin{aligned}
P(X_1 = x_1, \dots, X_n = x_n \mid T = t) &= \frac{P(X_1 = x_1, \dots, X_n = x_n, T = t)}{P(T = t)} \\
&= \begin{cases} 0 & \text{if } t \neq \sum_{i=1}^n x_i \\ \dfrac{P(X_1 = x_1, \dots, X_n = x_n)}{P(T = t)} & \text{if } t = \sum_{i=1}^n x_i \end{cases} \\
&= \begin{cases} 0 & \text{if } t \neq \sum_{i=1}^n x_i \\ \dfrac{\prod_{i=1}^n e^{-\lambda}\lambda^{x_i}/x_i!}{e^{-n\lambda}\lambda^t/t!} & \text{if } t = \sum_{i=1}^n x_i \end{cases} \\
&= \begin{cases} 0 & \text{if } t \neq \sum_{i=1}^n x_i \\ \dfrac{e^{-n\lambda}\lambda^t/\prod_{i=1}^n x_i!}{e^{-n\lambda}\lambda^t/t!} & \text{if } t = \sum_{i=1}^n x_i \end{cases} \\
&= \begin{cases} 0 & \text{if } t \neq \sum_{i=1}^n x_i \\ \dfrac{t!}{\prod_{i=1}^n x_i!} & \text{if } t = \sum_{i=1}^n x_i, \end{cases}
\end{aligned}
$$

which does not depend on the parameter λ.

(b) The likelihood is

$$
L(\lambda; \mathbf{X}) = \prod_{i=1}^n e^{-\lambda}\frac{\lambda^{X_i}}{X_i} = e^{-n\lambda}\frac{\lambda^{\sum_{i=1}^n X_i}}{\prod_{i=1}^n X_i!} = e^{-n\lambda}\lambda^T\frac{1}{\prod_{i=1}^n X_i!}.
$$

This is factored into $g(T, \lambda)h(\mathbf{X})$, so by the Neyman factorization theorem, $T = \sum_{i=1}^n X_i$ is a sufficient statistic for λ.

2. The likelihood of θ is given by

$$
\begin{aligned}
L(\theta; X_1, X_2) = f(X_1; \theta)f(X_2; \theta) &= \binom{n_1}{X_1}\theta^{X_1}(1-\theta)^{n_1 - X_1}\binom{n_2}{X_2}\theta^{X_2}(1-\theta)^{n_2 - X_2} \\
&= \binom{n_1}{X_1}\binom{n_2}{X_2}\theta^{(X_1 + X_2)}(1-\theta)^{n_1 + n_2 - X_1 - X_2} \\
&= \binom{n_1}{X_1}\binom{n_2}{X_2}\theta^{(n_1 + n_2)T_1}(1-\theta)^{n_1 + n_2(1 - T_1)}
\end{aligned}
$$

The joint likelihood is factored into the form $g(T_1, \theta)h(\mathbf{X})$, with $\mathbf{X} = (X_1, X_2)$, so T_1 is a sufficient statistic for θ.

Regarding T_2, the joint likelihood can't be factored into the form $g(T_2, \theta)h(\mathbf{X})$, with $\mathbf{X} = (X_1, X_2)$, so T_2 is not a sufficient statistic for θ.

3. Let S denote the number that is drawn. The estimator can be written in the form

$$T_n = \begin{cases} \bar{X}, & \text{if } S \in \{2, \ldots, n\} \\ n^2, & \text{if } S = 1 \end{cases}$$

(a) To show consistency, for any $\varepsilon > 0$, by the total probability formula, we have

$$\begin{aligned}
&P\left(|T_n - \mu| > \varepsilon\right) \\
&= P\left(|T_n - \mu| > \varepsilon \mid S \in \{2, \ldots, n\}\right) P(S \in \{2, \ldots, n\}) + P\left(|T_n - \mu| > \varepsilon \mid S = 1\right) P(S = 1) \\
&= P(|\bar{X}_n - \mu| > \varepsilon) \times \frac{n-1}{n} + P(|n^2 - \mu| > \varepsilon) \times \frac{1}{n}
\end{aligned}$$

By the Chebyshev inequality, we have

$$P(|\bar{X}_n - \mu| > \varepsilon) \leq \frac{\sigma^2}{n\varepsilon},$$

and also

$$P(|n^2 - \mu| > \varepsilon) \leq 1.$$

As a result,

$$P\left(|T_n - \mu| > \varepsilon\right) \leq \frac{\sigma^2(n-1)}{n^2\varepsilon} + \frac{1}{n},$$

As $n \to \infty$, the right hand side $\dfrac{\sigma^2(n-1)}{n^2\varepsilon} + \dfrac{1}{n} \to 0$, so

$$\lim_{n\to\infty} P\left(|T_n - \mu| > \varepsilon\right) \to 0,$$

showing T_n is consistent.

(b) We need to compute $E(T_n)$. Conditioned on $s \in \{2, \ldots, n\}$, we have $E(T_n \mid S = s) = E(\bar{X}) = \mu$, and conditioned on $S = 1$, we have $E(T_n \mid S = 1) = n^2$. Therefore, using the law of iterated expectation, we obtain

$$\begin{aligned}
E(T_n) &= E\left(E(T_n) \mid S\right) \\
&= \sum_{s=1}^{n} E(T_n \mid S = s)P(S = s) \\
&= E(T_n \mid S = 1)P(S = 1) + \sum_{s=2}^{n} E(T_n \mid S = s)P(S = s) \\
&= n^2 \times \frac{1}{n} + (n-1) \times \mu \times \frac{1}{n} = n + \frac{(n-1)\mu}{n}.
\end{aligned}$$

We can see $E(T_n) \neq \mu$, so T_n is not unbiased. Even if $n \to \infty$, the expectation $E(T_n) \to \infty$, so T_n is not asymptotically unbiased either.

4. (a) The likelihood based on observations $\mathbf{X} = (X_1, \ldots, X_n)^T$ is

$$L(\theta; \mathbf{X}) = \prod_{i=1}^{n} \left[\frac{1}{\theta} e^{-X_i/\theta}\right] = \frac{1}{\theta^n} e^{-T/\theta}$$

where $T = \sum_{i=1}^n X_i$ is the sample total. The log-likelihood is

$$\log \ell(\theta; \mathbf{X}) = -\frac{T}{\theta} - n \log \theta$$

and the score function is

$$\frac{\partial \log \ell(\theta; \mathbf{X})}{\partial \theta} = \frac{t}{\theta^2} - \frac{n}{\theta} = \frac{n}{\theta^2}\left(\frac{T}{n} - \theta\right) = \frac{n}{\theta^2}\left(\bar{X} - \theta\right).$$

We know that $E(\bar{X}) = E(X_1) = \theta$, and the score function is in the form

$$C_\theta\left(\bar{X} - \theta\right),$$

with $C_\theta = n/\theta^2$, so $\bar{X}$ is the MVU estimator for θ. To verify that $\bar{X}$ achieves the CRLB for estimating θ, first note that

$$\mathrm{Var}_\theta\left[\frac{\partial \log \ell(\theta; \mathbf{X})}{\partial \theta}\right] = \mathrm{Var}_\theta\left(\frac{n}{\theta^2}\bar{X}\right) = \frac{n^2}{\theta^4}\mathrm{Var}_\theta\left(\bar{X}\right) = \frac{n^2}{\theta^4}\frac{\mathrm{Var}_\theta\left(X_1\right)}{n} = \frac{n^2}{\theta^4}\frac{\theta^2}{n} = \frac{n}{\theta^2}.$$

So for any unbiased estimator $\hat{\theta}(\boldsymbol{X})$, the CRLB is

$$\mathrm{Var}_\theta\left[\hat{\theta}(\mathbf{X})\right] \geq \frac{1}{\mathrm{Var}_\theta\left[\frac{\partial \ell(\theta;\mathbf{X})}{\partial \theta}\right]} = \frac{\theta^2}{n} = \mathrm{Var}_\theta\left(\bar{X}\right).$$

(b) (i) We have

$$E(X_1^2) = \int_1^\infty x^2\left(\frac{\theta}{\theta-1}\right)x^{-\left(\frac{2\theta-1}{\theta-1}\right)}\,dx = \left(\frac{\theta}{\theta-1}\right)\int_1^\infty x^{-\left(\frac{1}{\theta-1}\right)}\,dx$$
$$= \begin{cases} \left(\frac{\theta}{\theta-2}\right)\left[x^{\frac{\theta-2}{\theta-1}}\right]_1^\infty = \left(\frac{\theta}{2-\theta}\right) & \text{for } 1 < \theta < 2, \\ \infty & \text{for } \theta \geq 2. \end{cases}$$

Thus

$$Var_\theta(X_1) = E(X_1^2) - \left\{\bar{E}(X_1)\right\}^2 = \frac{\theta}{2-\theta} - \theta^2 = \frac{\theta - 2\theta^2 + \theta^3}{2-\theta} = \frac{\theta(\theta-1)^2}{2-\theta}$$

for $1 < \theta < 2$, otherwise $Var_\theta(X_1) = \infty$; dividing by n gives the variance of $\bar{X}$, i.e

$$\mathrm{Var}(\bar{X}) = Var_\theta(X_1) = \frac{\theta(\theta-1)^2}{n(2-\theta)}.$$

(ii) The log likelihood is

$$\ell_\theta(\theta; \mathbf{X}) = \sum_{i=1}^n \log f(X_i; \theta) = n\log\left(\frac{\theta}{\theta-1}\right) - \left(\frac{2\theta-1}{\theta-1}\right)\sum_{i=1}^n \log X_i$$
$$= n\log\theta - n\log(\theta-1) - \left(2 + \frac{1}{\theta-1}\right)\sum_{i=1}^n \log X_i.$$

(iii) The score function is

$$\begin{aligned}\frac{\partial \ell(\theta;\mathbf{X})}{\partial\theta} &= \frac{n}{\theta} - \frac{n}{\theta-1} + \frac{1}{(\theta-1)^2}\sum_{i=1}^{n}\log(X_i)\\ &= \frac{-n}{\theta(\theta-1)} + \frac{1}{(\theta-1)^2}\sum_{i=1}^{n}\log(X_i)\\ &= \frac{n}{(\theta-1)^2}\left\{\frac{1}{n}\sum_{i=1}^{n}\log(X_i) - \frac{\theta-1}{\theta}\right\}.\end{aligned}$$

This is *not* in the special form, but it suggests that we should change parameters to $\eta = \eta(\theta) = \frac{\theta-1}{\theta} = 1 - \frac{1}{\theta}$.

(iv) Let $Y = \log(X_1) = g(X_1)$, so $X_1 = e^Y$. This transformation is one-to-one, with $g^{-1}(y) = e^y$. Hence,

$$f_Y(y) = f_{X_1}(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right| = \frac{\theta}{\theta-1}\left(e^y\right)^{-\left(\frac{2\theta-1}{\theta-1}\right)}e^y = \frac{\theta}{\theta-1}e^{-\left(\frac{\theta}{\theta-1}\right)y},\ y>0$$

which is the pdf of an exponential random variable with *mean* $\eta = \frac{\theta-1}{\theta} = 1 - \frac{1}{\theta}$. Thus $Var_\theta\left[\log(X_1)\right] = \eta^2 = \left(\frac{\theta-1}{\theta}\right)^2$, so the variance of the score is

$$\frac{1}{(\theta-1)^4}\,n\left(\frac{\theta-1}{\theta}\right)^2 = \frac{n}{\theta^2(\theta-1)^2}.$$

Thus the CRLB for estimating θ is

$$\frac{\theta^2(\theta-1)^2}{n}.$$

(v) If we change to $\eta = \eta(\theta) = \frac{\theta-1}{\theta} = 1 - \frac{1}{\theta}$, $\theta = \frac{1}{1-\eta}$, $\theta - 1 = \frac{\eta}{1-\eta}$, then

$$\frac{2\theta-1}{\theta-1} = 2 + \frac{1}{\theta-1} = 2 + \frac{1-\eta}{\eta} = 1 + \frac{1}{\eta} = \frac{\eta+1}{\eta}.$$

Thus the density is

$$f_\eta(x) = \frac{1}{\eta}x^{-\frac{\eta+1}{\eta}},$$

the log-likelihood is

$$\ell(\eta;\mathbf{X}) = -n\log\eta - \left(1+\frac{1}{\eta}\right)\sum_{i=1}^{n}\log(X_i)$$

and the score function is

$$\frac{\partial \ell(\eta;\mathbf{X})}{\partial\eta} = -\frac{n}{\eta} + \frac{1}{\eta^2}\sum_{i=1}^{n}\log(X_i) = \frac{n}{\eta^2}\left\{\frac{1}{n}\sum_{i=1}^{n}\log X_i - \eta\right\}.$$

This is in "the magic form" which shows us that the logarithmic average $\hat{\eta}(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n}\log(X_i)$ is the MVU estimator of η, which is the expectation of the exponential $\log(X_i)$'s, which (unsurprisingly) is optimally estimated using their sample mean.

5. (a) Note $\bar{X} \sim N(\theta, 1/n)$. Hence, we have

$$E_\theta(T_n) = E_\theta\left(\bar{X}^2 - \frac{1}{n}\right) = E_\theta(\bar{X}^2) - \frac{1}{n} = \mathrm{Var}_\theta(\bar{X}) + \left\{E_\theta(\bar{X})\right\}^2 = \frac{1}{n} + \theta^2 - \frac{1}{n} = \theta^2,$$

so T_n is unbiased for θ^2. As shown in the lecture, $N(\theta, 1)$ is a full exponential family distribution and $\bar{X}$ is a sufficient statistic for θ. Hence, T_n, as a function of $\bar{X}$ is the best unbiased estimator for its expected value, i.e T_n is the MVU for θ^2.

(b) First, we can compute the variance for T_n to be

$$\mathrm{Var}_\theta(T_n) = \mathrm{Var}_\theta\left(\bar{X}^2 - \frac{1}{n}\right) = \mathrm{Var}_\theta(\bar{X}^2) = E_\theta(\bar{X}^4) - \left\{E_\theta(\bar{X}^2)\right\}^2.$$

Using the formula in the hint, we have

$$E_\theta(\bar{X}^4) = \theta^4 + \frac{6\theta^2}{n} + \frac{3}{n^2},$$

and we have

$$E_\theta(\bar{X}^2) = \mathrm{Var}_\theta(\bar{X}) + \left\{E_\theta(\bar{X})\right\}^2 = \frac{1}{n} + \theta^2.$$

Therefore,

$$\mathrm{Var}_\theta(T_n) = \theta^4 + \frac{6\theta^2}{n} + \frac{3}{n^2} - \left(\frac{1}{n} + \theta^2\right)^2 = \frac{4\theta^2}{n} + \frac{2}{n^2}.$$

To compute the CRLB for estimating θ^2, the log likelihood function is

$$\ell(\theta; \mathbf{X}) = \sum_{i=1}^n \log f(X_i; \theta) = \sum_{i=1}^n \log(\sqrt{2\pi}) - \frac{1}{2}\sum_{i=1}^n (X_i - \theta)^2.$$

The score function is

$$\frac{\partial}{\partial\theta}\ell(\theta; \mathbf{X}) = \sum_{i=1}^n (X_i - \theta) = \sum_{i=1}^n X_i - n\theta$$

and

$$\mathrm{Var}_\theta\left(\frac{\partial}{\partial\theta}\ell(\theta; \mathbf{X})\right) = \mathrm{Var}\left(\sum_{i=1}^n X_i\right) = n.$$

Therefore, the CRLB for estimating θ^2 is

$$\frac{\left\{\frac{\partial d}{d\theta}\theta^2\right\}^2}{\mathrm{Var}_\theta\left(\frac{\partial}{\partial\theta}\ell(\theta; \mathbf{X})\right)} = \frac{(2\theta)^2}{n} = \frac{4\theta^2}{n}.$$

Hence, $\mathrm{Var}_\theta(T_n) = \dfrac{4\theta^2}{n} + \dfrac{2}{n^2} > \dfrac{4\theta^2}{n}$, so T_n does not achieve the CRLB.