Statistics 852 Fall 2011 Final Exam – Solutions

1. We will begin by showing that $W(X_1,\ldots,X_n)$ is an unbiased estimator of θ^2. This follows since

$$\begin{aligned}
E_\theta(W(X_1,\ldots,X_n)) = \frac{1}{n(n-1)}E_\theta\left(T(T-1)\right) &= \frac{1}{n(n-1)}\big(E_\theta(T^2)-E_\theta(T)\big)\\
&= \frac{1}{n(n-1)}\big(\mathrm{Var}_\theta(T)+[E_\theta(T)]^2-E_\theta(T)\big)\\
&= \frac{n\theta(1-\theta)+n^2\theta^2-n\theta}{n(n-1)}\\
&= \frac{n^2\theta^2-n\theta^2}{n(n-1)}\\
&= \theta^2.
\end{aligned}$$

In order to apply the Rao-Blackwell theorem, we need to show that T is a complete and sufficient statistic for θ. One might try to argue that since $T \sim \mathrm{Bin}(n,\theta)$, the density for T follows an exponential family so that T is therefore complete and sufficient. This would be correct if the parameter space were $0<\theta<1$. However, since we are considering $0\le\theta\le1$, the density for T does *not* follow an exponential family for all $\theta\in[0,1]$. If $\theta\in\{0,1\}$, then the support of the distribution *does* depend on θ. Instead, one can use the factorization theorem to conclude that T is a sufficient statistic for θ. Completeness follows from the fact that

$$\sum_{i=0}^{n}\binom{n}{i}\theta^i(1-\theta)^{n-1}g(i)=0$$

for all $0\le\theta\le1$ if and only if $g(i)=0$ for all $i=0,1,\ldots,n$. Hence, we can now apply the Rao-Blackwell theorem (Theorem 7.3.23) to conclude that $W(X_1,\ldots,X_n)$ is the MVUE of θ^2 since $W(X_1,\ldots,X_n)$ is a function of the sufficient and complete statistic T.

2. **(a)** The joint density of $X_1,\ldots,X_n$ is

$$f(x_1,\ldots,x_n|\theta)=\frac{1}{(2\pi)^{n/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-i\theta)^2\right\}$$

and so the log-likelihood function is

$$\ell(\theta)=\log L(\theta)=\log f(x_1,\ldots,x_n|\theta)=-\frac{n}{2}\log(2\pi)-\frac{1}{2}\sum_{i=1}^{n}(x_i-i\theta)^2.$$

Since

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\ell(\theta)=\sum_{i=1}^{n}i(x_i-i\theta)=\sum_{i=1}^{n}ix_i-\theta\sum_{i=1}^{n}i^2=0$$

if and only if

$$\theta = \frac{\sum_{i=1}^{n} i x_i}{\sum_{i=1}^{n} i^2},$$

and since

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}\ell(\theta) = \sum_{i=1}^{n} i^2 < 0$$

for all θ, we conclude that the MLE of θ is

$$\hat{\theta}(X_1, \ldots, X_n) = \frac{\sum_{i=1}^{n} i X_i}{\sum_{i=1}^{n} i^2}$$

as required.

(b) Since $X_1, \ldots, X_n$ are independent, the variance of $\hat{\theta}(X_1, \ldots, X_n)$ is

$$\mathrm{Var}(\hat{\theta}(X_1, \ldots, X_n)) = \frac{\sum_{i=1}^{n} i^2 \,\mathrm{Var}(X_i)}{\left(\sum_{i=1}^{n} i^2\right)^2} = \frac{\sum_{i=1}^{n} i^2}{\left(\sum_{i=1}^{n} i^2\right)^2} = \frac{1}{\sum_{i=1}^{n} i^2}.$$

(c) The Cramér-Rao lower bound for unbiased estimation of θ is

$$\left[\mathbb{E}_\theta\left(\frac{\partial}{\partial\theta}\log f(X_1, \ldots, X_n|\theta)\right)^2\right]^{-1} = \left[\sum_{i=1}^{n} \mathbb{E}_\theta\left(\frac{\partial}{\partial\theta}\log f(X_i|\theta)\right)^2\right]^{-1}$$

where

$$f(x_i|\theta) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(x_i - i\theta)^2\right\}$$

Now,

$$\frac{\partial}{\partial\theta}\log f(x_i|\theta) = -\frac{\partial}{\partial\theta}\frac{1}{2}(x_i - i\theta)^2 = i(x_i - \theta)$$

and so

$$\begin{aligned}\sum_{i=1}^{n} \mathbb{E}_\theta\left(\frac{\partial}{\partial\theta}\log f(X_i|\theta)\right)^2 = \sum_{i=1}^{n} \mathbb{E}_\theta\left(i(X_i - \theta)\right)^2 = \sum_{i=1}^{n} i^2 \mathbb{E}_\theta\left((X_i - \theta)\right)^2 &= \sum_{i=1}^{n} i^2 \,\mathrm{Var}_\theta(X_i) \\ &= \sum_{i=1}^{n} i^2.\end{aligned}$$

Hence, the Cramer-Rao lower bound is attained by the variance of $\hat{\theta}(X_1, \ldots, X_n)$.

3. (a) The joint density function of X_1 and X_2 is

$$\begin{aligned} f(x_1, x_2|\theta) = P_\theta(X_1 = x_1, X_2 = x_2) &= P_\theta(X_1 = x_1)P_\theta(X_2 = x_2) \\ &= \frac{1}{\theta^2} I(x_1 \in \{1, \dots, \theta\},\, x_2 \in \{1, \dots, \theta\}) \\ &= \frac{1}{\theta^2} I(x_1 \in \mathbb{N},\, x_2 \in \mathbb{N},\, x_1 \le \theta,\, x_2 \le \theta) \\ &= \frac{1}{\theta^2} I(x_1 \in \mathbb{N},\, x_2 \in \mathbb{N},\, \max\{x_1, x_2\} \le \theta). \end{aligned}$$

Thus, by the factorization theorem, T is a sufficient statistic for θ.

(b) For $t = 1, 2, \dots, \theta$, we find

$$\begin{aligned} P_\theta(T \le t) = P_\theta(\max\{X_1, X_2\} \le t) &= P_\theta(X_1 \le t, X_2 \le t) \\ &= P_\theta(X_1 \le t)P_\theta(X_2 \le t) \\ &= \left(\frac{t}{\theta}\right)^2. \end{aligned}$$

Thus,

$$P_\theta(T = t) = P_\theta(T \le t) - P_\theta(T \le t - 1) = \left(\frac{t}{\theta}\right)^2 - \left(\frac{t-1}{\theta}\right)^2 = \frac{t^2 - (t-1)^2}{\theta^2} = \frac{2t - 1}{\theta^2}$$

for $t = 1, 2, \dots, \theta$.

(c) In order to show that the family of distributions of T is complete, we need to show that $E_\theta[g(T)] = 0$ for all θ implies that $P_\theta(g(T) = 0) = 1$ for all θ. Now

$$E_\theta[g(T)] = \frac{1}{\theta^2} \sum_{t=1}^{\theta} g(t)(2t - 1)$$

so that $E_\theta[g(T)] = 0$ for all θ implies that

$$\sum_{t=1}^{\theta} g(t)(2t - 1) = 0$$

for all θ. If $\theta = 1$, then

$$0 = \sum_{t=1}^{\theta} g(t)(2t - 1) = \sum_{t=1}^{1} g(t)(2t - 1) = g(1)(2 - 1) = g(1).$$

If $\theta = 2$, then

$$0 = \sum_{t=1}^{\theta} g(t)(2t - 1) = \sum_{t=1}^{2} g(t)(2t - 1) = g(1)(2 - 1) + g(2)(4 - 1) = 3g(2)$$

since $g(1) = 0$. Continuing in this way shows that $g(t) = 0$ for all $t = 1, 2, \dots$ so that $P_\theta(g(T) = 0) = 1$ for all θ as required.

(d) By the Rao-Blackwell theorem (Theorem 7.3.23), the (unique) MVUE of θ is a function of T, say $\phi(T)$, satisfying

$$\theta = E_\theta(\phi(T)) = \sum_{t=1}^{\theta} \phi(t) P_\theta(T = t).$$

In other words, ϕ satisfies

$$\sum_{t=1}^{\theta} \phi(t) \frac{(2t-1)}{\theta^2} = \theta \quad \text{or, equivalently,} \quad \theta^3 = \sum_{t=1}^{\theta} (2t-1)\phi(t)$$

for all $\theta = 1, 2, \ldots$. Using the hint, we find

$$\sum_{t=1}^{\theta} (3t^2 - 3t + 1) = \sum_{t=1}^{\theta} (2t-1)\phi(t)$$

so that $(3t^2 - 3t + 1) = (2t-1)\phi(t)$. Thus,

$$\phi(T) = \frac{3T^2 - 3T + 1}{2T - 1}$$

is the MVUE of θ.

4. (a) If $\theta = P_\lambda(X_1 \le 1)$, then

$$\theta = P_\lambda(X_1 = 0) + P_\lambda(X_1 = 1) = e^{-\lambda}(1 + \lambda).$$

Since the MLE of λ is

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i,$$

we conclude that the MLE of θ is

$$\hat{\theta}(X_1, \ldots, X_n) = e^{-\overline{X}}(1 + \overline{X}).$$

(b) Since

$$T = \sum_{i=1}^{n} X_i \sim \text{Poisson}(n\lambda)$$

and $\overline{X} = T/n$, we can analyze

$$E_\theta(\hat{\theta}(X_1, \ldots, X_n)) = \mathbb{E}_\theta(e^{-T/n}(1 + T/n))$$

directly. That is,

$$\begin{aligned} E_\theta(\hat{\theta}(X_1, \ldots, X_n)) = \mathbb{E}_\theta(e^{-T/n}(1 + T/n)) &= \sum_{t=0}^{\infty} e^{-t/n}(1 + t/n) \frac{e^{-n\lambda}(n\lambda)^t}{t!} \\ &= \sum_{t=0}^{\infty} e^{-t/n} \frac{e^{-n\lambda}(n\lambda)^t}{t!} + \frac{1}{n} \sum_{t=0}^{\infty} t e^{-t/n} \frac{e^{-n\lambda}(n\lambda)^t}{t!} \end{aligned}$$

Now

$$\sum_{t=0}^{\infty} e^{-t/n} \frac{e^{-n\lambda}(n\lambda)^t}{t!} = e^{-n\lambda} \exp\{n\lambda e^{-1/n}\}$$

and

$$\begin{aligned} \frac{1}{n}\sum_{t=0}^{\infty} t e^{-t/n} \frac{e^{-n\lambda}(n\lambda)^t}{t!} = \frac{1}{n}\sum_{t=1}^{\infty} t e^{-t/n} \frac{e^{-n\lambda}(n\lambda)^t}{t!} &= \frac{e^{-n\lambda}}{n} \sum_{t=1}^{\infty} \frac{(n\lambda e^{-1/n})^t}{(t-1)!} \\ &= \frac{e^{-n\lambda}}{n}(n\lambda e^{-1/n}) \sum_{t=1}^{\infty} \frac{(n\lambda e^{-1/n})^{t-1}}{(t-1)!} \\ &= \lambda e^{-n\lambda} e^{-1/n} \sum_{t=0}^{\infty} \frac{(n\lambda e^{-1/n})^t}{t!} \\ &= \lambda e^{-n\lambda} e^{-1/n} \exp\{n\lambda e^{-1/n}\} \\ &= \lambda e^{-n\lambda} \exp\{n\lambda e^{-1/n} - 1/n\} \end{aligned}$$

so that

$$\begin{aligned} E_\theta(\hat{\theta}(X_1,\ldots,X_n)) &= e^{-n\lambda}\exp\{n\lambda e^{-1/n}\} + \lambda e^{-n\lambda}\exp\{n\lambda e^{-1/n} - 1/n\} \\ &= e^{-n\lambda}\exp\{n\lambda e^{-1/n}\}\left[1 + \lambda e^{-1/n}\right] \\ &= \exp\{\lambda(n e^{-1/n} - n)\}\left[1 + \lambda e^{-1/n}\right]. \end{aligned}$$

Hence, $\hat{\theta}(X_1,\ldots,X_n)$ is not an unbiased estimator of θ. Note that

$$\lim_{n\to\infty}(n e^{-1/n} - n) = \lim_{n\to\infty} n(e^{-1/n} - 1) = \lim_{n\to\infty} n\left[1 - \frac{1}{n} + O\left(\frac{1}{n^2}\right) - 1\right] = -1$$

so that $E_\theta(\hat{\theta}(X_1,\ldots,X_n)) \to e^{-\lambda}(1+\lambda)$ implying that $\hat{\theta}(X_1,\ldots,X_n)$ is asymptotically an unbiased estimator of θ.

(c) From the central limit theorem, we know that

$$\frac{\sqrt{n}\,(\overline{X} - \lambda)}{\sqrt{\lambda}} \to \mathcal{N}(0,1)$$

in distribution as $n \to \infty$. Moreover, $\overline{X} \to \lambda$ in probability as $n \to \infty$ since

$$P_\lambda(|\overline{X} - \lambda| \geq \varepsilon) \leq \frac{\mathrm{Var}_\lambda(\overline{X})}{\varepsilon^2} = \frac{\lambda}{n\varepsilon^2}$$

so that $e^{-\overline{X}}(1+\overline{X}) \to e^{-\lambda}(1+\lambda)$ in probability as well. If we now let $g(y) = (1+y)e^{-y}$ for $y > 0$ so that $g'(y) = -ye^{-y}$, then

$$\frac{\sqrt{n}\left(e^{-\overline{X}}(1+\overline{X}) - e^{-\lambda}(1+\lambda)\right)}{\sqrt{\lambda}\cdot\lambda e^{-\lambda}} \to \mathcal{N}(0,1)$$

or, equivalently,

$$\sqrt{n}\,(\,\hat{\theta}(X_1,\ldots,X_n) - \theta\,) \to \mathcal{N}(0, \lambda^3 e^{-2\lambda})$$

in distribution as $n \to \infty$.

5. The distribution of

$$\frac{1}{n}\sum_{i=1}^{n} X_i$$

is normal with mean θ and variance

$$\frac{\sigma^2}{n} + \frac{n(n-1)}{n^2}\rho\sigma^2 = \frac{\sigma^2(1-\rho)}{n} + \rho\sigma^2.$$

This means that the asymptotic distribution of

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

is $\mathcal{N}(\theta, \rho\sigma^2)$ with $\rho\sigma > 0$. Hence,

$$\lim_{n\to\infty} P_\theta\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \theta\right| \le \epsilon\right) = \frac{1}{\sigma\sqrt{2\pi\rho}}\int_{-\epsilon}^{\epsilon} e^{-\frac{y^2}{2\rho\sigma^2}}\,\mathrm{d}y < 1$$

for all $\epsilon > 0$. Since this limit does not equal 1 for all $\epsilon > 0$, we conclude that $\{\overline{X}_n\}$ is not a consistent sequence of estimators of θ.

6. Note that $T = \max\{X_1, \ldots, X_n\}$ is sufficient and complete for θ. Moreover, if $0 \le t \le \theta$, then

$$P_\theta(T \le t) = \left(\frac{t}{\theta}\right)^n$$

so that

$$E_\theta(T) = \frac{1}{\theta^n}\int_0^\theta t \cdot nt^{n-1}\,\mathrm{d}t = \frac{n}{n+1}\theta$$

which implies that

$$\frac{n+1}{n}T$$

is an unbiased estimator of θ. Now observe that $E_\theta(X_i) = \theta/2$ so that $E_\theta(\,\overline{X}\,) = \theta/2$. Thus, $2\overline{X}$ is also an unbiased estimator of θ. We know from the Rao-Blackwell theorem (Theorem 7.3.17) that for any unbiased estimator $W(X_1, \ldots, X_n)$ of θ, the (unique) minimum variance unbiased estimator of θ is

$$\phi(T) = E(\,W(X_1, \ldots, X_n)\,|\,T\,).$$

If we set $W = \frac{n+1}{n}T$, then

$$\phi(T) = E(\,W\,|\,T\,) = E\left(\frac{n+1}{n}T\,\middle|\,T\right) = \frac{n+1}{n}T.$$

If we set $W = 2\overline{X}$, then

$$\phi(T) = E(\,W\,|\,T\,) = E\left(2\overline{X}\,|\,T\right) = 2E\left(\overline{X}\,|\,T\right).$$

Thus, equating these two expressions for $\phi(T)$ implies that

$$2E\left(\overline{X}\,|\,T\right) = \frac{n+1}{n}T \quad \text{and so} \quad E\left(\overline{X}\,|\,T = t\right) = \frac{n+1}{2n}t$$

as required.