THE UNIVERSITY OF NEW SOUTH WALES

DEPARTMENT OF STATISTICS

MID SESSION TEST - 2020 - Thursday, 2nd April (Week 7)
Solutions

MATH5905

Time allowed: 135 minutes

1. Let $\mathbf{X} = (X_1, X_2, \dots, X_n)$ be a sample of i.i.d. Bernoulli(θ) random variables with density function

$$f(x, \theta) = \theta^x (1-\theta)^{1-x}, x = \{0, 1\}; \theta \in (0, 1)).$$

a) The statistic $T(X) = \sum_{i=1}^n X_i$ is complete and sufficient for θ. Provide justification for why this statement is true.

Using the property of the one-parameter exponential family, we observe that

$$f(x, \theta) = \theta^x (1-\theta)^{1-x} = (1-\theta) \exp\left(x \log \frac{\theta}{1-\theta} \right).$$

Thus, the Bernoulli distribution belongs to the one-parameter exponential family. Then this implies that $T(X) = \sum_{i=1}^n X_i$ is (minimal) sufficient and complete for θ.

b) Derive the UMVUE of $h(\theta) = \theta^k$ where $k = 1, 2, \dots, n$ is a known integer. You must justify each step in your answer.
Hint: Use the interpretation that $P(X_1 = 1) = \theta$ and therefore we have that $P(X_1 = 1, \dots, X_k = 1) = P(X_1 = 1)^k = \theta^k$.

1) **T is sufficient and complete for θ.**

2) **Let $W = I_{\{X_1=1,\dots,X_k=1\}}(X)$ and then using the fact that $P(X_i = 1) = \theta$ we see that**

$$E(W) = P(X_1 = 1, \dots, X_k = 1) = P(X_1 = 1)^k = \theta^k,$$

which is unbiased for $h(\theta)$.

3) **Now, apply the Theorem of Lehmann-Scheffe, and obtain the following**

$$
\begin{aligned}
\hat{\tau}(T) = E(W|T=t) &= E\left(I_{\{X_1=1,\ldots,X_k=1\}} | \sum_{i=1}^{n} X_i = t \right) \\
&= P\left(X_1 = 1, \ldots, X_k = 1 | \sum_{i=1}^{n} X_i = t \right) \\
&= \frac{P\left(X_1 = 1, \ldots, X_k = 1, \sum_{i=1}^{n} X_i = t \right)}{P\left(\sum_{i=1}^{n} X_i = t \right)}
\end{aligned}
$$

The events in the numerator must be satisfied simultaneously to have a non-zero probability and hence this reduces to

$$
\begin{aligned}
\hat{\tau}(T) &= \frac{P\left(X_1 = 1, \ldots, X_k = 1, \sum_{i=k+1}^{n} X_i = t-k \right)}{P\left(\sum_{i=1}^{n} X_i = t \right)} \\
&= \theta^k \frac{P\left(\sum_{i=k=1}^{n} X_i = t-k \right)}{P\left(\sum_{i=1}^{n} X_i = t \right)} \\
&= \theta^k \frac{\binom{n-k}{t-k} \theta^{t-k} (1-\theta)^{n-k-(t-k)}}{\binom{n}{t} \theta^t (1-\theta)^{n-t}} \\
&= \frac{\binom{n-k}{t-k}}{\binom{n}{t}}.
\end{aligned}
$$

Therefore the UMVUE for $h(\theta) = \theta^k$ is $\widehat{h_{umvue}}(\theta) = \frac{\binom{n-k}{t-k}}{\binom{n}{t}}$.

c) Calculate the Cramer-Rao lower bound for the minimal variance of an unbiased estimator of $h(\theta) = \theta^k$.

First we calculate the fisher information as follows. We have that

$$
\begin{aligned}
\log f(x) &= x \log \theta + (1-x) \log(1-\theta) \\
\frac{\partial}{\partial \theta} \log f(x) &= \frac{x}{\theta} - \frac{1-x}{1-\theta} \\
\frac{\partial^2}{\partial \theta^2} \log f(x) &= -\frac{x}{\theta^2} - \frac{1-x}{(1-\theta)^2}
\end{aligned}
$$

and so the Fisher information in a single sample is

$$
I_{X_1}(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} \log f(x) \right] = \frac{\theta}{\theta^2} + \frac{1-\theta}{(1-\theta)^2} = \frac{1}{\theta} + \frac{1}{1-\theta} = \frac{1}{\theta(1-\theta)}
$$

Hence the Fisher information for the whole sample is

$$I_{\boldsymbol{X}} = nI_{X_1}(\theta) = \frac{n}{\theta(1-\theta)}.$$

Then notice that

$$\frac{\partial}{\partial\theta}h(\theta) = k\theta^{k-1}.$$

Hence the Cramer-Rao lower bound is

$$\frac{\left(\frac{\partial}{\partial\theta}h(\theta)\right)^2}{I_{\boldsymbol{X}}(\theta)} = \frac{\theta(1-\theta)}{n}k^2\theta^{2(k-1)} = \frac{k^2\theta^{2k-1}(1-\theta)}{n}$$

d) Find the value of k for which the variance of the UMVUE of $h(\theta)$ attains the Cramer-Rao lower bound found in part (c).
To show that the Cramer-Rao bound is NOT attainable we will look at the score function:

$$\begin{aligned}
V(\boldsymbol{X},\theta) &= \frac{1}{\theta}\sum_{i=1}^{n} X_i - \frac{n-\sum_{i=1}^{n} X_i}{1-\theta} \\
&= \frac{n\bar{X}}{\theta} - \frac{n(1-\bar{X})}{1-\theta} \\
&= \frac{n\bar{X}}{\theta} - \frac{n}{1-\theta} + \frac{n\bar{X}}{1-\theta} \\
&= \frac{n\bar{X}}{\theta(1-\theta)} - \frac{n}{1-\theta} \\
&= \frac{n}{\theta^k(1-\theta)}\left(\bar{X}\theta^{k-1} - \theta^k\right)
\end{aligned}$$

Therefore $\bar{X}\theta^{k-1}$ is only a statistic (it does not depend on θ) when $k=1$ and the CRLB is attainable when $k=1$.

e) Determine the MLE $\hat{h}$ of $h(\theta)$.
The MLE for θ is simply $\hat{\theta} = \bar{X}$ and hence the

$$\widehat{h(\theta)}_{\textbf{mle}} = h(\hat{\theta}_{\textbf{mle}}) = \bar{X}^k.$$

f) Suppose that $n=6$, $T=3$ and $k=2$ compute the numerical values of the UMVUE in part (b) and the MLE in part (e). What would happen to these values as $n\to\infty$ but the ratio $T/n = 1/2$ remained the same.
The UMVUE is

$$\widehat{h(\theta)}_{\textbf{umvue}} = \frac{\binom{6-2}{3-2}}{\binom{6}{3}} = \frac{\binom{4}{1}}{\binom{6}{3}} = \frac{4}{20} = 0.2$$

The MLE is

$$\widehat{h(\theta)}_{\mathbf{mle}} = \left(\frac{3}{6}\right)^2 = 0.25.$$

As $n \to \infty$ the UMVUE would converge to the MLE at the value $(T/n)^2 = 0.25$.

g) Consider testing $H_0 : \theta \leq 0.5$ versus $H_1 : \theta > 0.5$ with a 0-1 loss in Bayesian setting with the prior $\tau(\theta) = 60\theta^3(1-\theta)^2$. What is your decision when $n = 6$ and $T = 3$. You may use:

$$\int_0^{0.5} x^6(1-x)^5 \mathrm{d}x = 0.0000698.$$

Note: The continuous random variable X has a beta density f with parameters $\alpha > 0$ and $\beta > 0$ if

$$f(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, x \in (0, 1)$$

where

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)},$$

and

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha) = \alpha!$$

First we need to compute the posterior by observing that

$$h(\theta|x) \propto 60\theta^3(1-\theta)^2 \times \binom{6}{3}\theta^3(1-\theta)^3 \propto \theta^6(1-\theta)^5$$

which implies that

$$\theta|X \sim \mathbf{Beta}(7, 6).$$

Hence we are interested in computing the posterior probability

$$\begin{aligned} P(\theta < 0.5|X) &= \int_0^{0.5} \frac{1}{B(7,6)} \theta^6(1-\theta)^5 \mathrm{d}\theta \\ &= \frac{\Gamma(7+6)}{\Gamma(7)\Gamma(6)} \times \int_0^{0.5} \theta^6(1-\theta)^5 \mathrm{d}\theta \\ &= \frac{12!}{6! \times 5!} \times 0.0000698 \\ &= 5544 \times 0.0000698 \\ &= 0.387 \end{aligned}$$

We compare this posterior probability with 0.5 since we are dealing with a 0-1 loss. Since this probability is smaller than 0.5 we must reject H_0.

2. Let $X_1, X_2, \ldots, X_n$ be independent random variables, with a density

$$f(x;\theta) = \begin{cases} \dfrac{\theta}{x^2}, & x > \theta, \\ 0 & \text{otherwise} \end{cases}$$

where $\theta > 0$ is an unknown parameter. Let $T = \min\{X_1, \ldots, X_n\} = X_{(1)}$ be the minimal of the n observations.

a) Show that T is a sufficient statistic for the parameter θ.
First calculate the likelihood as follows

$$L(\boldsymbol{X}, \theta) = \frac{\theta^n}{\prod_{i=1}^n X_i^2} \prod_{i=1}^n I_{(0,X_i)}(\theta) = \frac{\theta^n}{\prod_{i=1}^n X_i^2} I_{(0,X_{(1)})}(\theta).$$

Thus T is sufficient by the Neyman Fisher Factorization Criterion.

b) Show that the density of T is

$$f_T(t) = \begin{cases} \dfrac{n\theta^n}{t^{n+1}}, & t > \theta, \\ 0 & \text{otherwise} \end{cases}$$

Hint: You may compute the CDF first by using

$$P(X_{(1)} < x) = 1 - P(X_1 > x \cap X_2 > x \cdots \cap X_n > x).$$

First note that for $x < \theta$ we have $P(X_1 \geq x) = 1$ and for $x \geq \theta$ we have,

$$(X_1 \geq x) = \int_x^\infty \frac{\theta}{y^2} \mathrm{d}y = \left[-\frac{\theta}{y} \right]_{y=x}^{y=\infty} = \left(\frac{\theta}{x} \right).$$

Hence,

$$\begin{aligned} F_T(x,\theta) &= P(T \leq t) \\ &= 1 - P(X_1 \geq t, X_2 \geq t, \ldots, X_n \geq t) \\ &= 1 - P(X_1 \geq t)^n \\ &= \begin{cases} 1 - \left(\dfrac{\theta}{t} \right)^n & \textbf{if} \quad x \geq \theta \\ 0 & \textbf{if} \quad x < \theta. \end{cases} \end{aligned}$$

Then by differentiation

$$f_T(t,\theta) = \frac{n\theta^n}{t^{n+1}}, \quad t \geq 0,$$

otherwise zero.

c) Find the maximum likelihood estimator of θ and provide justification.
The MLE for θ is calculated by maximizing over all θ values

$$L(\boldsymbol{X}, \theta) = \frac{\theta^n}{\prod_{i=1}^n X_i^2} I_{(0,X_{(1)})}(\theta).$$

The graph of this function starts at zero and increases until the value $X_{(1)}$ and from here stays at zero. Hence $X_{(1)}$ must be the MLE. You should plot this.

d) Show that the MLE is a biased estimator.
By calculating

$$\begin{aligned}
E(X_{(1)}) &= \int_\theta^\infty t \frac{n\theta^n}{t^{n+1}} \mathrm{d}t \\
&= \int_\theta^\infty \frac{n\theta^n}{t^n} \mathrm{d}t \\
&= \left[-\frac{n}{n-1} \theta^n t^{-n+1} \right]_{t=0}^{t=\infty} \\
&= \frac{n}{n-1} \theta \neq \theta
\end{aligned}$$

and hence the MLE is a biased estimator.

e) Show that $T = X_{(1)}$ is complete for θ.
We know that the density of T is

$$f_T(t) = nt^{n-1}/\theta^n, 0 < t < \theta \textbf{ (and 0 else).}$$

Let $E_\theta g(T) = 0$ for all $\theta > 0$. This implies:

$$\int_0^\theta g(t) \frac{nt^{n-1}}{\theta^n} dt = 0 = \frac{1}{\theta^n} \int_0^\theta g(t) nt^{n-1} dt$$

for all $\theta > 0$ must hold. Since $\frac{1}{\theta^n} \neq 0$ we get $\int_0^\theta g(t) nt^{n-1} dt = 0$ for all $\theta > 0$. Differentiating both sides with respect to θ we get

$$ng(\theta)\theta^{n-1} = 0$$

for all $\theta > 0$. This implies $g(\theta) = 0$ for all $\theta > 0$. This also means $P_\theta(g(T) = 0) = 1$. Hence $T = X_{(n)}$ is complete.

f) Hence determine the UMVUE of θ.
From part d) we know that $W = \frac{n-1}{n} X_{(1)}$ is an unbiased estimator for θ. Thus by Lehmann-Scheffe theorem,

$$\hat{\theta} = E(W|X_{(1)}) = \frac{n-1}{n} X_{(1)}$$

is the unique UMVUE of θ.