UNIVERSITY OF OSLO

Faculty of mathematics and natural sciences

Exam in: STK4011/9011 — Statistical Inference Theory

Day of examination: Friday, 4 December 2020

Examination hours: 9.00 – 13.00

This solution proposal consists of 5 pages.

Appendices: None

Permitted aids: All examination aids are allowed. However, it is not allowed to cooperate or communicate with others regarding the exam.

Solution Proposal

Problem 1

(a) Calculate the expected value (or first moment):

$$E(X) = \int_1^\infty x\theta x^{-(\theta+1)}dx = \theta\left[\frac{x^{-\theta+1}}{-\theta+1}\right]_1^\infty = \frac{\theta}{1-\theta}(0-1) = \frac{\theta}{\theta-1}.$$

Set the first population moment equal to the first sample moment (sample mean) and solve for θ:

$$E(X) = \overline{X} \Leftrightarrow \theta = (\theta-1)\overline{X} \Leftrightarrow \theta(1-\overline{X}) = -\overline{X} \Leftrightarrow \theta = \frac{\overline{X}}{\overline{X}-1}.$$

Thus, the method of moments estimator is $\hat{\theta}_{MM} = \dfrac{\overline{X}}{\overline{X}-1}$.

(b) For every $\theta > 1$, by Thm 5.5.2, $\overline{X}_n$ converges in probability to

$$E(X) = \frac{\theta}{\theta-1} \quad \text{(See (a))},$$

and by Thm 5.5.4, we have that

$$\frac{\overline{X}}{\overline{X}-1} \to \frac{\frac{\theta}{\theta-1}}{\frac{\theta}{\theta-1}-1} = \theta \quad \text{in probability.}$$

By Def 10.1.1, $\hat{\theta}_{MM}$ is thus a consistent estimator.

(c) The likelihood function is

$$L(\theta\,|\,\mathbf{x}) = \prod_{i=1}^n \theta x_i^{-(\theta+1)} = \theta^n\left(\prod_{i=1}^n x_i\right)^{-(\theta+1)},$$

(Continued on page 2.)

and the log-likelihood is

$$l(\theta \mid \mathbf{x}) = \log\left(L(\theta \mid \mathbf{x})\right) = n\log(\theta) - (\theta+1)\log\left(\prod_{i=1}^{n} x_i\right).$$

Find the extreme point(s):

$$\frac{d}{d\theta} l(\theta \mid \mathbf{x}) = \frac{n}{\theta} - \log\left(\prod_{i=1}^{n} x_i\right) \overset{\text{set}}{=} 0 \Leftrightarrow \theta = \frac{n}{\log\left(\prod_{i=1}^{n} x_i\right)}.$$

Show that the found solution is a global maximum and conclude the we have found the MLE:

$$\frac{d^2}{d\theta^2} l(\theta \mid \mathbf{x}) = -\frac{n}{\theta^2} < 0 \Rightarrow \hat{\theta}_{ML} = \frac{n}{\log\left(\prod_{i=1}^{n} X_i\right)} = \frac{n}{\sum_{i=1}^{n} \log(X_i)}.$$

(d)

$$\begin{aligned} E(\hat{\theta}_{ML}) &= E(n/T) = nE(1/T) \\ &= n\int_0^\infty \frac{1}{t}\frac{\theta^n}{\Gamma(n)} t^{n-1} e^{-t\theta}\,dt \\ &= n\frac{\theta^n}{\Gamma(n)}\frac{\Gamma(n-1)}{\theta^{n-1}} \underbrace{\int_0^\infty \frac{\theta^{n-1}}{\Gamma(n-1)} t^{n-2} e^{-t\theta}\,dt}_{=1} \\ &= \frac{n}{n-1}\theta \end{aligned}$$

The above calculations show that $\hat{\theta}_{ML}$ is biased (the bias vanishes as $n \to \infty$).

(e) The pdf belongs to an exponential family since it can be expressed as

$$f(x \mid \theta) = \frac{\theta}{x^{\theta+1}} \mathbb{I}_{(1,\infty)}(x) = \underbrace{\mathbb{I}_{(1,\infty)}(x)}_{=h(x)} \underbrace{\theta}_{=c(\theta)} \exp\left(\underbrace{-(\theta+1)}_{=w(\theta)} \underbrace{\log(x)}_{=t(x)}\right)$$

By Thm 6.2.10 and Thm 6.2.25, $T = \sum_{i=1}^{n} t(X_i) = \sum_{i=1}^{n} \log(X_i)$ is a complete sufficient statistic for θ.

(f) By Thm 7.3.23, $\hat{\theta}_{ML}$ is the unique best unbiased estimator for $\frac{n}{n-1}\theta$, since it is based on T, which is a complete sufficient statistic for θ. Hence, $\phi(T) = \frac{n-1}{n}\hat{\theta}_{ML}$ is the unique best unbiased estimator of θ.

Problem 2

(a)

$$\begin{aligned} M_X(t) &= E\left(e^{tX}\right) \\ &= \sum_{x=1}^{\infty} e^{tx}\theta(1-\theta)^{x-1} \\ &= \theta e^t \sum_{x=1}^{\infty} \left(e^t(1-\theta)\right)^{x-1} \\ &\overset{*}{=} \frac{\theta e^t}{1 - e^t(1-\theta)}, \quad \text{for } t < -\log(1-\theta). \end{aligned}$$

(Continued on page 3.)

*A converging geometric series if $e^t(1-\theta) < 1 \Leftrightarrow t < -\log(1-\theta)$.

(b) By Thm 2.3.7, we have that

$$E(X^n) = \frac{d^n}{dt^n} M_X(t)\Big|_{t=0}.$$

Calculate the first and second derivative:

$$\frac{d}{dt} M_X(t) = \ldots = \frac{\theta e^t}{\big(1 - e^t(1-\theta)\big)^2},$$

$$\frac{d^2}{dt^2} M_X(t) = \ldots = \frac{\theta e^t\big(1 + e^t(1-\theta)\big)}{\big(1 - e^t(1-\theta)\big)^3},$$

and use the above formula to show that

$$E(X) = \frac{d}{dt} M_X(t)\Big|_{t=0} = \frac{\theta}{\theta^2} = \frac{1}{\theta},$$

$$E(X^2) = \frac{d^2}{dt^2} M_X(t)\Big|_{t=0} = \frac{\theta(2-\theta)}{\theta^3} = \frac{2-\theta}{\theta^2},$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{2-\theta}{\theta^2} - \frac{1}{\theta^2} = \frac{1-\theta}{\theta^2}.$$

(c) Since the ratio

$$\frac{f(\mathbf{x}\,|\,\theta)}{f(\mathbf{y}\,|\,\theta)} = \frac{\prod_{i=1}^n \theta(1-\theta)^{x_i - 1}}{\prod_{i=1}^n \theta(1-\theta)^{y_i - 1}} = \frac{\theta^n(1-\theta)^{\sum x_i - n}}{\theta^n(1-\theta)^{\sum y_i - n}} = (1-\theta)^{\sum x_i - \sum y_i}$$

is constant w.r.t. θ if and only if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$, by Thm 6.2.13, $T = \sum_{i=1}^n X_i$ is a minimal sufficient statistic for θ.

(d) By Thm 10.1.12,

$$\sqrt{n}(\hat{\theta}_{ML} - \theta) \to N\big(0, v(\theta)\big)$$

in distribution, where

$$v(\theta) = \frac{1}{E\Big(\big(\frac{\partial}{\partial\theta}\log(f(X\,|\,\theta))\big)^2\Big)}$$

is the Cramér-Rao Lower Bound (and asymptotic variance). The denominator in the above expression is

$$\begin{aligned}
E\Big(\big(\frac{\partial}{\partial\theta}\log(f(X\,|\,\theta))\big)^2\Big) &= E\Big(\big(\frac{\partial}{\partial\theta}\big[\log(\theta) + (x-1)\log(1-\theta)\big]\big)^2\Big) \\
&= E\Big(\big(\frac{1-\theta x}{\theta(1-\theta)}\big)^2\Big) \\
&= E\left(\frac{1 - 2\theta x + \theta^2 x}{\theta^2(1-\theta)^2}\right) \\
&= \frac{1 - 2\theta E(X) + \theta^2 E(X^2)}{\theta^2(1-\theta)^2} \\
&\overset{(b)}{=} \frac{1 - 2 + 2 - \theta}{\theta^2(1-\theta)^2} \\
&= \frac{1}{\theta^2(1-\theta)},
\end{aligned}$$

and, hence, the asymptotic variance is $\theta^2(1-\theta)$.

(Continued on page 4.)

(e) By the invariance property (Thm 7.2.10), the MLE of $\mu(\theta) = \theta^{-1}$ is $\hat{\mu}_{ML} = \hat{\theta}_{ML}^{-1} = \overline{X}$. To show that the estimator attains the Cramér-Rao Lower bound, show that the derivative of the log-likelihood

$$l(\theta \,|\, \mathbf{x}) = n\log(\theta) - (n - \sum_{i=1}^{n} x_i)\log(1-\theta)$$

w.r.t. θ can be written as

$$\begin{aligned}\frac{\partial}{\partial\theta} l(\theta \,|\, \mathbf{x}) &= \frac{n}{\theta} + \frac{n - \sum_{i=1}^{n} x_i}{1-\theta} \\ &= \frac{n - n\theta + n\theta - n\theta\bar{x}}{\theta(1-\theta)} \\ &= \underbrace{-\frac{n}{1-\theta}}_{=a(\theta)}(\underbrace{\bar{x}}_{=W(\mathbf{x})} - \underbrace{\frac{1}{\theta}}_{=\tau(\theta)}).\end{aligned}$$

Then, by Thm 7.3.15, we have that $\hat{\mu}_{ML} = W(\mathbf{X}) = \overline{X}$ attains the bound.

Problem 3

(a) The log-likelihood function of θ is

$$l(\theta \,|\, \mathbf{x}) = \log\big(\prod_{i=1}^{n} \theta e^{-\theta x_i}\big) = \log\big(\theta^n e^{-n\bar{x}\theta}\big) = n\log(\theta) - n\bar{x}\theta.$$

Find the extreme point(s):

$$\frac{d}{d\theta} l(\theta \,|\, \mathbf{x}) = \frac{n}{\theta} - n\bar{x} \overset{\text{set}}{=} 0 \Leftrightarrow \theta = \frac{1}{\bar{x}}$$

and show that the found point is a global maximum, that is, the MLE:

$$\frac{d^2}{d\theta^2} l(\theta \,|\, \mathbf{x}) = -\frac{n}{\theta^2} < 0 \Rightarrow \hat{\theta}_{ML} = \frac{1}{\bar{x}}$$

The likelihood ratio (LR) statistic is thus

$$\lambda(\mathbf{x}) = \frac{\theta_0^n e^{-n\bar{x}\theta_0}}{\frac{1}{\bar{x}^n} e^{-n}} = (\theta_0\bar{x})^n e^{-n(\bar{x}\theta_0 - 1)}$$

By Thm 10.3.1, $-2\log\big(\lambda(\mathbf{x})\big) \to \chi_1^2$ in distribution as $n \to \infty$. Hence, we obtain a large-sample approximate test of size α by rejecting H_0 if

$$2n(\bar{x}\theta_0 - 1 - \log(\bar{x}\theta_0)) > \chi_{1,\alpha}^2.$$

(b) Let $g(t) = \theta_0^n t^n e^{-n(t\theta_0 - 1)}$, that is, $g(\bar{x}) = \lambda(\mathbf{x})$. We then have that

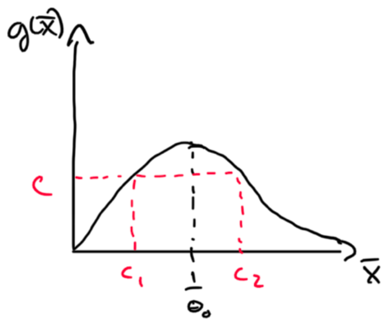
$$\begin{aligned}\frac{d}{dt} g(t) &= \theta_0^n n t^{n-1} e^{-n(t\theta_0-1)} + \theta_0^n t^n(-n\theta_0) e^{-n(t\theta_0-1)} \\ &= \theta_0^n n t^{n-1} e^{-n(t\theta_0-1)}(1 - t\theta_0) \overset{\text{set}}{=} 0 \Leftrightarrow t = \frac{1}{\theta_0},\end{aligned}$$

(Continued on page 5.)

and, furthermore,

$$g'(t) > 0 \text{ for } 0 < t < \frac{1}{\theta_0} \quad \text{and} \quad g'(t) < 0 \text{ for } t > \frac{1}{\theta_0}.$$

That is, $g(t)$ is increasing for $0 < t < \frac{1}{\theta_0}$ and decreasing for $t > \frac{1}{\theta_0}$. This means that the LRT, which rejects H_0 if $\lambda(\mathbf{x}) = g(\bar{x}) \leq c$, is equivalent to rejecting H_0 if $\overline{X} \leq c_1$ or $\overline{X} \geq c_2$ for some constants c_1 and c_2. This is illustrated in the figure below.



(c) Under H_0 we have that

$$X_i \sim \text{Exp}(\frac{1}{\theta_0}) \quad \text{and} \quad M_{X_i}(t) = \frac{1}{1 - \frac{t}{\theta_0}}, \quad t < \theta_0.$$

By Thm 5.2.7,

$$M_{\overline{X}}(t) = \left(\frac{1}{1 - \frac{t}{n\theta_0}} \right)^n, \quad t < \theta_0,$$

which is the mgf of a Gamma$(n, \frac{1}{n\theta_0})$ variable. Since we know the distribution of $\overline{X}$ under H_0, we can select c_1 and c_2 such that

$$\alpha = P(H_0 \text{ is rejected} \,|\, H_0 \text{ is true}) = 1 - \int_{c_1}^{c_2} f_{\overline{X}}(\bar{x} \,|\, \theta_0) d\bar{x},$$

and, possibly, such that $g(c_1) = g(c_2)$, where g is the function in (b).