Problem Set #6

Danny Edgel
Econ 709: Economic Statistics and Econometrics I
Fall 2020

October 18, 2020

Collaborated with Sarah Bass, Emily Case, Michael Nattinger, and Alex Von Hafften

Question 1

Say $P(X = 1) = p$ and $P(X = 0) = 1 - p$, where $0 < p < 1$.

(a) Say $f(x) = p^x(1-p)^{1-x}$. Then,

$$f(0) = p^0(1-p)^{1-0} = 1 - p = P(X = 0)$$
$$f(1) = p^1(1-p)^{1-1} = p = P(X = 1)$$

(b)

$$\ell_n = \sum_{i=1}^{n} \log(f(x_i)) = \sum_{i=1}^{n} x_i \log(p) + (1-x_i)\log(1-p) = n\log(p) + \log(1-p)\sum_{i=1}^{n} 1 - x_i$$

(c) To find $\hat{p}$, we simply maximize ℓ_n with repspect to p:

$$\begin{aligned}
\frac{\partial \ell_n}{\partial p} &= \frac{1}{p}\sum_{i=1}^{n} x_i - \frac{1}{1-p}\sum_{i=1}^{n} 1 - x_i = 0 \\
\frac{n}{p}\overline{X}_n &= \frac{n}{1-p} - \frac{n}{1-p}\overline{X}_n \\
\frac{p-1}{p}\overline{X}_n &= 1 - \overline{X}_n \\
\left(\frac{p-1}{p} + 1\right)\overline{X}_n &= 1 \\
\frac{1}{p}\overline{X}_n &= 1 \\
\hat{p}_n &= \overline{X}_n
\end{aligned}$$

Question 2

$X \sim f(x) = \frac{\alpha}{x^{1+\alpha}}$, $x \geq 1$

(a) The log-likelihood function is:

$$\ell_n = \sum_{i=1}^{n} \log\left(f(x_i)\right) = \sum_{i=1}^{n} \log(\alpha) - (1+\alpha)\log(x_i) = n\log(\alpha) - (1+\alpha)\sum_{i=1}^{n} \log(x_i)$$

(b) To find $\hat{\alpha}$, we simply maximize ℓ_n with repspect to α:

$$\frac{\partial \ell_n}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} \log(x_i) = 0$$

$$\frac{n}{\hat{\alpha}} = \sum_{i=1}^{n} \log(x_i)$$

$$\hat{\alpha}_n^{-1} = \frac{1}{n}\sum_{i=1}^{n} \log(x_i)$$

Question 3

$X \sim f(x) = \left[\pi(1+(x-\theta)^2)\right]^{-1}$, $x \in \mathbb{R}$

(a) The log-likelihood function is:

$$\ell_n = \sum_{i=1}^{n} \log\left(f(x_i)\right) = \sum_{i=1}^{n} \log(\pi) + \log\left(1+(x_i-\theta)^2\right) = -n\log(\pi) - \sum_{i=1}^{n} \log\left(1+(x_i-\theta)^2\right)$$

(b) The first-order condition for the MLE $\hat{\theta}$ is:

$$\frac{\partial \ell_n}{\partial \theta} = \sum_{i=1}^{n} \frac{2(x_i - \hat{\theta}_n)}{1+(x_i-\hat{\theta}_n)} = 0$$

Question 4

$X \sim f(x) = \frac{1}{2}\exp(-|x-\theta|)$, $x \in \mathbb{R}$

(a) The log-likelihood function is:

$$\ell_n = \sum_{i=1}^{n} \log\left(f(x_i)\right) = \sum_{i=1}^{n} \log\left(\frac{1}{2}\right) - |x_i - \theta| = n\log\left(\frac{1}{2}\right) - \sum_{i=1}^{n} -|x_i-\theta|$$

(b) The MLE will be $\hat{\theta}_n$ that minimizes $\sum_{i=1}^{n} |x_i - \hat{\theta}_n|$, so we want to choose theta that will minimize the sum of the absolute deviations from X_i. We already know that this value is $\frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X}_n$. Thus,

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

Question 5

$f(x) = \alpha x^{-1-\alpha}$, $x \geq 1$. I is defined as:

$$I_0 = -E\left[\frac{\partial^2 \log\left(f(x|\theta_0)\right)}{\partial\theta\partial\theta'}\right]$$

Thus, given $f(x|\alpha)$:

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

Therefore,

$$I = \frac{1}{\alpha^2}$$

Question 6

$f(x) = \theta\exp(-\theta x)$, $x \geq 0$, $\theta > 0$

(a) The Cramer-Rao Lower Bound (CRLB) is equal to $(nI_0)^{-1}$. then,

$$\begin{aligned}
\log\left(f(x)\right) &= \log\left(\theta\right) - \theta x \\
\frac{\partial \log\left(f(x)\right)}{\partial\theta} &= \frac{1}{\theta} - x \\
I = \frac{\partial^2 \log\left(f(x)\right)}{\partial\theta^2} &= -\frac{1}{\theta^2} \\
\text{CRLB } &= \frac{1}{n}\theta^2
\end{aligned}$$

(b) From a previous problem, $\hat{\theta}_n = (\overline{X}_n)^{-1}$. Then $\hat{\theta}_n = g(\overline{X}_n)$. By the central limit theorem,

$$\sqrt{n}(\overline{X}_n - \mu) \to_d \mathcal{N}(0, \sigma^2)$$

Where $Var(X) = \sigma^2$. Then, by the delta method,

$$\sqrt{n}(\hat{\theta}_n - \theta) \to_d g'(\overline{X}_n)\mathcal{N}(0, \sigma^2) = -\overline{X}_n^{-2}\mathcal{N}(0, \sigma^2)$$

Thus,

$$\hat{\theta}_n \to_d \mathcal{N}\left(\theta, \frac{\sigma^2}{n\overline{X}_n^4}\right)$$

Where $E(\overline{X}_n) = \frac{1}{\theta_0}$ and, because $X \sim \theta\exp(-\theta x)$, $\sigma^2 = \frac{1}{\theta_0^2}$. Therefore,

$$\hat{\theta}_n \to_d \mathcal{N}\left(\theta, \frac{1}{n}\theta_0^2\right)$$

(c) From section 6, $\sqrt{n}(\hat{\theta}_n - \theta_0) \to_d \mathcal{N}(0, I_0^{-1})$. Then,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \to_d \mathcal{N}(0, \theta^2)$$

Which is equal to the answer I got in (b).

Question 7

(a) In question 1, we found $\hat{p}_n = \overline{X}_n$. Then $\hat{p}_n = g(\overline{X}_n) = \overline{X}_n$, so by the delta method,

$$\sqrt{n}(\overline{X}_n - \mu) \to_d \mathcal{N}(0, \sigma^2) \Rightarrow \sqrt{n}(\hat{p}_n - p) \to_d \mathcal{N}(0, \sigma^2)$$

so $V = Var(X_i) = \sum_{i=1}^n (X_i - \overline{X}_n)^2$.

(b) By the weak law of large numbers, $\frac{1}{n}\sum_{i=1}^n X_i \to_p E(X)$ and $\sum_{i=1}^n (X_i - E(X))^2 = E(X - E(X))^2$, so V is a consistent esimator.

(c) Recall that $\hat{p}_n = \overline{X}_n$. We know that $Var(\overline{X}_n) = \frac{1}{n}\sigma^2$, and our estimator for σ^2 is given by $\hat{\sigma}_n = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n$. Thus,

$$s(\hat{p}_n) = \frac{1}{n^2}\sum_{i=1}^n (X_i - \hat{p}_n)^2$$

Question 8

(a) Let F_X be the CDF of Uniform$[0, \theta]$. The PDF of any uniform distribution over $[A, B]$ is $f(x) = \frac{1}{B-A}$, so we can derive:

$$F_X(c) = \int_0^c f(x)dx = \int_0^c \frac{1}{\theta}dx = \frac{x}{\theta}|_0^c = \begin{cases} 0, & c < 0 \\ \frac{c}{\theta}, & c \in [0, \theta] \\ 1, & c > \theta \end{cases}$$

(b) From the definition of $F_{n(th\hat{e}ta_n - \theta)}(x)$, we can solve:

$$\begin{aligned} F_{n(th\hat{e}ta_n - \theta)}(x) &= \Pr\left(\max_{i=1,\ldots,n}(n(X_i - \theta)) \le x\right) \\ &= \prod_{i=1}^n \Pr\left(n(X_i - \theta) \le x\right) \\ &= \prod_{i=1}^n \Pr\left(X_i \le \frac{x}{n} + \theta\right) \\ &= \prod_{i=1}^n F_X(\theta + \frac{x}{n}) \\ F_{n(th\hat{e}ta_n - \theta)}(x) &= \left(F_X(\theta + \frac{x}{n})\right)^n \end{aligned}$$

(c) Knowing that $\lim\limits_{x\to\infty}(1+\frac{y}{n}) = e^y\ \forall y \in \mathbb{R}$,

$$\lim_{x\to\infty} F_{n(\hat{\theta}_n-\theta)}(x) = \lim_{x\to\infty}(F_X(\theta+\frac{x}{n}))^n = \begin{cases} \lim\limits_{x\to\infty} 0^n, & \theta+\frac{x}{n} < 0 \\ \lim\limits_{x\to\infty} \left(\frac{\theta+\frac{x}{n}}{\theta}\right)^n, & \theta+\frac{x}{n} \in [0,\theta] \\ \lim\limits_{x\to\infty} 1^n, & \theta+\frac{x}{n} > \theta \end{cases}$$

Simplifying and recognizing that $\lim\limits_{x\to\infty}\left(\frac{\theta+\frac{x}{n}}{\theta}\right)^n = \lim\limits_{x\to\infty}\left(1+\frac{x/\theta}{n}\right)^n$, we get:

$$\frac{\partial}{\partial x}\left(F_{n(th\hat{e}ta_n-\theta)}(x)\right) = \begin{cases} 0, & x < -n\theta \\ \frac{1}{\theta}e^{\frac{x}{\theta}}, & x \in [-n\theta, 0] \\ 0, x > 0 \end{cases}$$

Thus, $n(\hat{\theta}_n - \theta) \to_d f(-x|\theta)$, where $f(-x|\theta)$ is an exponential distribution with parameter $\frac{1}{\theta}$.

Question 9

$X \sim \mathcal{N}(\mu, \sigma^2)$, H_0: $\mu = 1$, H_1: $\mu \neq 1$. To test this hypothesis, collect an i.i.d. sample, $\{X_1, ..., X_n\}$ and calculate:

$$T = \frac{\sqrt{n}(\overline{X}_n - 1)|}{s_x}, \text{ where } \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \text{ and } s_x = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$$

$$t_{\frac{\alpha}{2},n-1}, \text{ where } t_{\frac{\alpha}{2},n-1} \text{ is the } \left(1-\frac{\alpha}{2}\right)^{\text{th}} \text{ percentile of } t_{n-1}$$

Then, choose $\alpha = 0.05$. If $T > t_{\frac{\alpha}{2},n-1}$, then reject H_0. Otherwise, do not reject H_0.

Question 10

$X \sim \mathcal{N}(\mu, 1)$, H_0: $\mu \in \{0, 1\}$, H_1: $\mu \notin \{0, 1\}$, where:

$$T = \min\left\{|\sqrt{n}\overline{X}_n|, |\sqrt{n}(\overline{X}_n - 1)|\right\}$$

And the critical value, c, is the $(1-\alpha)^{\text{th}}$ quantile of $\min\{|Z|, |Z - \sqrt{n}|\}$, where $Z \sim \mathcal{N}(0, 1)$.

If $\mu = 0$, then, by the central limit theorem (CLT),

$$\sqrt{n}\overline{X}_n = \sqrt{n}(\overline{X}_n - E(\overline{X}_n)) \to_d \mathcal{N}(0,1)$$
$$\sqrt{n}(\overline{X}_n - 1) = \sqrt{n}\overline{X}_n - \sqrt{n} \to_d \mathcal{N}(-\sqrt{n}, 1)$$

And if $\mu = 1$, then,

$$\begin{aligned}\sqrt{n}(\overline{X}_n - 1) = \sqrt{n}(\overline{X}_n - E(\overline{X}_n)) \to_d \mathcal{N}(0,1) \\ \sqrt{n}\overline{X}_n = \sqrt{n}(\overline{X}_n - 1) + \sqrt{n} \to_d \mathcal{N}(\sqrt{n},1)\end{aligned}$$

Note that, since the normal distribution is symmetric and Z is mean zero, Z and $-Z$ have the same distribution. Further, we can define $Z = \sqrt{n}\overline{X}_n$. Then, we can solve:

$$\begin{aligned}\Pr\left(T > c|\mu = 0\right) &= \Pr\left(\min\{|\sqrt{n}\overline{X}_n|, |\sqrt{n}(\overline{X}_n - 1)|\} > c|\mu = 0\}\right) \\ &= \Pr\left(\min\{|Z|, |Z - \sqrt{n}| > c\}\right) \\ &= \alpha \text{ (by construction)} \\ \Pr\left(T > c|\mu = 1\right) &= \Pr\left(\min\{|\sqrt{n}\overline{X}_n|, |\sqrt{n}(\overline{X}_n - 1)|\} > c|\mu = 0\}\right) \\ &= \Pr\left(\min\{|Z + \sqrt{n}|, |Z|\} > c\right) \\ &= \Pr\left(\min\{|(-1)(-Z - \sqrt{n})|, |Z|\} > c\right) \\ &= \Pr\left(\min\{|Z - \sqrt{n}|, |Z|\} > c\right) \\ &= \alpha\end{aligned}$$

$\therefore \Pr\left(T > c|\mu = 0\right) = \Pr\left(T > c|\mu = 1\right) = \alpha$