Econ 241A Probability, Statistics and Econometrics
Fall 2017

Problem Set 8

1. Let $X_1, X_2, ..., X_n$ be iid with pdf $f(x|\theta) = \frac{1}{\theta}$.

   (a) What is the MLE of θ?
   This exercise is similar to the MLE in PS7.

$$L(\theta|\mathbf{x}) = \begin{cases} \prod_{i=1}^{n} \frac{1}{\theta} = \left(\frac{1}{\theta}\right)^n \text{ if } x_{(1)} \geq 0 \text{ and } x_{(n)} < \theta \\ 0 \text{ otherwise} \end{cases}$$

   Alternatively

$$L(\theta|x) = \left(\frac{1}{\theta}\right)^n I_{[-\infty,\theta]}(x_{(n)}) I_{[0,\infty]}(x_{(1)})$$

   Since the likelihood function is decreasing in θ, as long as $\theta > x_{(n)}$, the MLE of θ, θ_{MLE}, should be $x_{(n)}$.

   (b) What is the Method of Moments estimator of θ?
   We need to equate the first sample moment to the first population moment.

$$\mathbb{E}(X) = \frac{\theta}{2} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

   Hence

$$\theta_{MM} = 2\bar{X}$$

   (c) Denote the MLE of θ, $\hat{\theta}$. Also, denote the Method of Moments estimator of θ as $\tilde{\theta}$. Finally, define $\hat{\hat{\theta}} = \frac{n+1}{n}\hat{\theta}$. Fill the following table:

Estimator	Expected Value	Variance	Mean Squared Error
$\tilde{\theta}$	θ	$\frac{\theta^2}{3n}$	$\frac{\theta^2}{3n}$
$\hat{\theta}$	$\frac{n}{n+1}\theta$	$\frac{n}{(n+2)(n+1)^2}\theta$	$\frac{2}{(n+1)(n+2)}\theta^2$
$\hat{\hat{\theta}}$	θ	$\frac{\theta^2}{n(n+2)}$	$\frac{\theta^2}{n(n+2)}$

   Hint: For the expected value of $\hat{\theta}$, derive first the exact cdf, $F_{\hat{\theta}_n}(x|n)$. Then, calculate $f_{\hat{\theta}_n}(x|n)$.

   <u>Expected values</u>

   The expected value of $\bar{\theta}$ is easy to obtain because it is a linear function of the sample observations. Hence,

$$\mathbb{E}\left(\bar{\theta}\right) = \frac{2}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \frac{2}{n}n\frac{\theta}{2} = \theta$$

   The expected value of the MLE is a bit trickier, since $\max(X_1, .., X_n)$ is not a linear function of X_is. Hence, we need to use the definition of expectation:

$$\mathbb{E}(\hat{\theta}) = \int_0^{\theta} x f_{\hat{\theta}}(x)dx$$

where $f_{\hat{\theta}}(x)$ is the pdf of $\hat{\theta}$ evaluated at x.
Using the hint, we start by calculating the cdf.

$$
\begin{aligned}
F_{\hat{\theta}}(x) &= \Pr\left(\hat{\theta} < x\right) = \Pr\left(X_{(n)} < x\right) \\
&= \Pr(X_1 < x) \times \ldots \times \Pr(X_n < x) \\
&= \prod_{i=1}^{n} \frac{1}{\theta} x = \left(\frac{1}{\theta}\right)^n x^n, \ x \in [0, \theta]
\end{aligned}
$$

Then, following the hint, we obtain $f_{\hat{\theta}}(x)$ by taking the derivative of the cdf wrt x.

$$
f_{\hat{\theta}}(x) = \frac{1}{\theta^n} n x^{n-1}, \ x \in [0, \theta]
$$

Now we can take the expectation:

$$
\begin{aligned}
\mathbb{E}(\hat{\theta}) &= \int_0^{\theta} \frac{1}{\theta^n} n x^{n-1} x dx \\
&= \frac{1}{\theta^n} n \int_0^{\theta} x^n dx \\
&= \frac{1}{\theta^n} n \left[\frac{1}{n+1} x^{n+1} \Big|_0^{\theta} \right] dx \\
\mathbb{E}(\hat{\theta}) &= \frac{n}{n+1} \theta
\end{aligned}
$$

The expected value of $\hat{\hat{\theta}}$ is very easy after this since

$$
\mathbb{E}\left(\hat{\hat{\theta}}\right) = \frac{n+1}{n} \mathbb{E}(\hat{\theta}) = \theta
$$

<u>Variances</u>
The variance of the MM estimator is the variance of X_i divided by $n/4$. Hence, we get that

$$
Var(\bar{\theta}) = \frac{\theta^2}{3n}
$$

The variance of the MLE can be computed once we get the expected value of $\hat{\theta}^2$.

$$
\begin{aligned}
\mathbb{E}\left[\hat{\theta}^2\right] &= \frac{1}{\theta^n} n \int_0^\theta x^{n+1} dx \\
&= \frac{1}{\theta^n} n \left[\frac{1}{n+2} x^{n+2} \Big|_0^\theta \right] dx \\
&= \frac{n}{n+2} \theta^2 \\
Var\left[\hat{\theta}^2\right] &= \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\right)^2 \theta^2 \\
&= \left[\frac{n}{n+2} - \left(\frac{n}{n+1}\right)^2\right] \theta^2 \\
&= \frac{n}{(n+2)(n+1)^2} \theta^2
\end{aligned}
$$

<u>MSE</u>

The MSE of unbiased estimators is equal to the variance of each estimator which is true for $\tilde{\theta}$ and $\hat{\hat{\theta}}$.

For biased estimators $MSE(\hat{\theta}) = Var(\hat{\theta}) + Bias(\hat{\theta})^2$

(d) What is the Cramer-Rao lower bound for the variance of θ?

The Cramer-Rao lower bound (CRLB) for iid data is defined as

$$
CRLB = \frac{\left(\frac{dE(W(X))}{d\theta}\right)^2}{nE\left[\frac{\partial}{\partial \theta} \log f(x_i|\theta)^2\right]}
$$

For an unbiased estimator $W(X)$ such as $E[W(X)] = \theta$ it reduces to

$$
CRLB = \frac{1}{n\frac{1}{\theta^2}} = \frac{\theta^2}{n}
$$

(e) Does any estimate among $\tilde{\theta}$, $\hat{\theta}$, and $\hat{\hat{\theta}}$ 'attain' the Cramer-Rao lower bound for unbiased estimates of θ (note that not all of them are unbiased)? Discuss (Hint: notice that the uniform distribution does not belong to the exponential family).

The variance of all 3 estimators is below the CRLB. This is a consequence that the uniform does not belong to the uniform distribution family

In addition, solve the following problems from Casella and Berger: 7.19, 7.20, 7.21, 7.48, 8.2 and 8.18.

7.19 (a) The joint pdf for the sample is

$$
\begin{aligned}
f(\mathbf{y}|\beta,\sigma^2) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(\frac{1}{2\sigma^2}\sum_i (y_i-\beta x_i)^2\right) \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(\frac{1}{2\sigma^2}\sum_i (y_i-\beta x_i)^2\right) \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(\frac{1}{2\sigma^2}\sum_i (y_i^2-2\beta x_i y_i+\beta^2 x_i^2)\right) \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(\frac{1}{2\sigma^2}(\sum_i y_i^2-2\beta\sum_i x_i y_i+\beta^2\sum_i x_i^2)\right)
\end{aligned}
$$

We can factor the sample pdf into two functions: the first one is a function of the statistics $\sum_i Y_i^2$ and $\sum_i x_i Y_i$ given the parameters of interest β and σ^2.

$$
\begin{aligned}
g\left(\sum_i y_i, \sum_i x_i y_i|\beta,\sigma^2\right) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(\frac{1}{2\sigma^2}(\sum_i y_i^2-2\beta\sum_i x_i y_i+\beta^2\sum_i x_i^2)\right) \\
h(\mathbf{y}) &= 1
\end{aligned}
$$

Hence the two dimenssional sufficient statistic for $\left(\beta,\sigma^2\right)$ is $\mathbf{W}(\mathbf{Y}) = \left(\sum_i x_i Y_i, \sum_i Y_i^2\right)$.

(b) The MLE for β, and σ^2 can be obtained by differentiating the log-likelihood wrt these parameters.

$$
\begin{aligned}
L\left(\beta,\sigma^2|\mathbf{y}\right) = f(\mathbf{y}|\beta,\sigma^2) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(\frac{1}{2\sigma^2}(\sum_i y_i^2-2\beta\sum_i x_i y_i+\beta^2\sum_i x_i^2)\right) \\
l\left(\beta,\sigma^2|\mathbf{y}\right) &= n\ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \left(\frac{1}{2\sigma^2}(\sum_i y_i^2-2\beta\sum_i x_i y_i+\beta^2\sum_i x_i^2)\right) \\
l_\beta(\beta,\sigma^2|\mathbf{y}) &= \frac{1}{2\sigma^2}\left(2\sum_i x_i y_i + 2\beta\sum_i x_i^2\right) = 0 \\
\hat{\beta} &= \frac{\sum_i x_i Y_i}{\sum_i x_i^2} \\
\mathbb{E}\left(\hat{\beta}\right) &= \mathbb{E}\left(\frac{\sum_i x_i Y_i}{\sum_i x_i^2}\right) \\
&= \frac{\sum_i x_i \mathbb{E}\left(Y_i\right)}{\sum_i x_i^2} \\
&= \frac{\sum_i x_i \beta x_i}{\sum_i x_i^2} = \frac{\beta\sum_i x_i^2}{\sum_i x_i^2} = \beta
\end{aligned}
$$

(c)

$$
\begin{aligned}
x_i Y_i &\sim \mathrm{n}(\beta x_i^2, x_i^2 \sigma^2) \\
\sum_i x_i Y_i &\sim \mathrm{n}\left(\beta \sum_i x_i^2, \sum_i x_i^2 \sigma^2\right) \\
\frac{\sum_i x_i Y_i}{\sum_i x_i^2} &\sim \mathrm{n}\left(\beta, \frac{\sum_i x_i^2}{(\sum_i x_i^2)^2} \sigma^2\right) \\
\frac{\sum_i x_i Y_i}{\sum_i x_i^2} &\sim \mathrm{n}\left(\beta, \frac{1}{\sum_i x_i^2} \sigma^2\right)
\end{aligned}
$$

7.20 (a)

$$
\begin{aligned}
E\left[\frac{\sum Y_i}{\sum x_i}\right] &= \frac{\sum E[Y_i]}{\sum x_i} \\
&= \frac{\sum \beta x_i}{\sum x_i} \\
&= \frac{\sum x_i}{\sum x_i} \beta \\
&= \beta
\end{aligned}
$$

(b)

$$
\begin{aligned}
\mathrm{Var}\,(\beta_{MLE}) &= \frac{\sigma^2}{\sum x_i^2} \\
\mathrm{Var}\left(\frac{\sum Y_i}{\sum x_i}\right) &= \frac{\mathrm{Var}(\sum Y_i)}{(\sum x_i)^2} \\
&= \frac{n\sigma^2}{(\sum x_i)^2} \\
&= \frac{n\sigma^2}{n^2 \bar{x}^2} \\
&= \frac{\sigma^2}{n\bar{x}^2}
\end{aligned}
$$

Since β_{MLE} is BLUE then it has lower variance than any other linear unbiased estimator.

7.20 (a)

$$
\begin{aligned}
E\left[\sum \frac{Y_i}{n x_i}\right] &= \sum \frac{E[Y_i]}{n x_i} \\
&= \sum \frac{\beta x_i}{n x_i} \\
&= \frac{\beta}{n} \sum \frac{x_i}{x_i} \\
&= \beta
\end{aligned}
$$

(b)

$$\begin{aligned}
\operatorname{Var}\left(\sum \frac{Y_i}{nx_i}\right) &= \sum \frac{\operatorname{Var}(Y_i)}{n^2x_i^2} \\
&= \sum \frac{\sigma}{n^2x_i^2} \\
&= \frac{\sigma}{n^2} \sum \frac{1}{x_i^2}
\end{aligned}$$

Since β_{MLE} is BLUE then it has lower variance than any other linear unbiased estimator.
Note that $g(u) = 1/x^2$ is convex in x. By Jensen Inequality

$$\begin{aligned}
\frac{1}{\bar{x}^2} &\leq \sum \frac{1}{n} \frac{1}{x_i^2} \\
\frac{\sigma}{n\bar{x}^2} &\leq \sum \frac{\sigma}{n^2} \frac{1}{x_i^2} \\
\operatorname{Var}\left(\frac{\sum Y_i}{\sum x_i}\right) &\leq \operatorname{Var}\left(\sum \frac{Y_i}{nx_i}\right)
\end{aligned}$$

7.48 (a)

$$CRLB = \frac{\left(\frac{dE(W(X))}{dp}\right)^2}{-nE\left[\frac{\partial^2}{\partial p^2}\log f(x_i|p)\right]}$$

for unbiased estimators

$$CRLB = \frac{1}{-nE\left[\frac{\partial^2}{\partial p^2}\log p^{x_i}(1-p)^{1-x_i}\right]}$$

$$CRLB = \frac{1}{-nE\left[-\frac{x_i}{p^2} - -\frac{(1-x_i)}{(1-p)^2}\right]} = \frac{p(1-p)}{n}$$

$p_{MLE} = \bar{x}$, and $\operatorname{var}(p_{MLE}) = \frac{\operatorname{var}(x)}{n} = \frac{p(1-p)}{n}$. Then p_{MLE} attains CRLB and is the best unbiased estimator.

(b) $E[X_1X_2X_3X_4] = E[X_1]E[X_2]E[X_3]E[X_4] = p^4$. The first equality comes from independence and the second from $X_i \sim Bernoulli(p)$. Then the estimator is unbiased.
By Theorem 7.3.17 (Rao-Balckwell) and Theorem 7.3.23, $E[X_1X_2X_3X_4|T(X)]$ is the best unbiased estimator for p^4 as long as $T(X)$ is a complete and sufficient statistic. Since $Bernoulli(p)$ belongs to the exponential family, we know that $T(X) = \sum X_i$ is a complete sufficient statistic (Theorem 6.2.25).
Then the best unbiased estimator is

$$\phi(T) = E[X_1X_2X_3X_4|\sum X] = \frac{C(n-4, \sum X_i - 4)}{C(n, \sum X_i)}$$

Where $C(a,b)$ is "a chooses b" or the number of possible b combinations out of a.

8.2 We can calculate the probability that the rate of accidents per year is 10 given that the mean is 15 and the distribution is Poisson. Assume that the CLT holds and the sample is large enough

$$P(Z \le \frac{10-15}{\sqrt{15}}) = P(Z \le \frac{-5}{\sqrt{15}}) = 0.0983$$

Using the exact pdf of a $Poisson(15)$ distribution

$$P(Y \le 10 | \lambda = 15) = 0.1185$$

8.18 (a) The power function is the probability of rejecting when the null is false.

$$\beta(\theta) = P_\theta\left(\frac{|\bar{X} - \theta_0|}{\sigma/\sqrt{n}} > c\right)$$

where P_θ is the probability given that the true parameter is $\theta \neq \theta_0$

$$\begin{aligned}
\beta(\theta) &= 1 - P_\theta\left(\frac{|\bar{X} - \theta_0|}{\sigma/\sqrt{n}} < c\right) \\
&= 1 - P_\theta\left(-\frac{c\sigma}{\sqrt{n}} < \bar{X} - \theta_0 < \frac{c\sigma}{\sqrt{n}}\right) \\
&= 1 - P_\theta\left(-\frac{c\sigma}{\sqrt{n}} < \bar{X} - \theta_0 < \frac{c\sigma}{\sqrt{n}}\right) \\
&= 1 - P_\theta\left(\frac{-c\sigma/\sqrt{n} + \theta_0 - \theta}{\sigma/\sqrt{n}} < \frac{\bar{X} - \theta_0}{\sigma/\sqrt{n}} < \frac{c\sigma/\sqrt{n} + \theta_0 - \theta}{\sigma/\sqrt{n}}\right) \\
&= 1 - P_\theta\left(\frac{-c\sigma/\sqrt{n} + \theta_0 - \theta}{\sigma/\sqrt{n}} < Z < \frac{c\sigma/\sqrt{n} + \theta_0 - \theta}{\sigma/\sqrt{n}}\right) \\
&= 1 + \Phi\left(-c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) - \Phi\left(c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right)
\end{aligned}$$

when $\theta = \theta_0$ the power reduces to the significance level.

(b) For a size of 5% we need that $\beta(\theta_0) = 1 + \Phi(-c) - \Phi(c) = 0.05$ which implies that $c = 1.96$. For a power of 0.25 at $\theta = \sigma + \theta_0$, it is required that $0.75 = \beta(\sigma + \theta) = 1 + \Phi(-1.96 - \sqrt{n}) - \Phi(1.96 - \sqrt{n}) \approx 1 - \Phi(1.96 - \sqrt{n})$