Econ 241A Probability, Statistics and Econometrics
Fall 2017

Solutions Problem Set 7

1. (Consistency of the sample slope)

   Recall that we introduced earlier on what was the best linear predictor for Y given X, when Y and X were jointly distributed. We defined the best linear predictor as

$$\mathbb{E}^*(Y|X) = \alpha + \beta X,$$

   where

$$\beta = \frac{\sigma_{XY}}{\sigma_X^2}$$

$$\alpha = \mu_Y - \beta\mu_X$$

   Generally σ_{XY} and σ_X^2 are unknown (if the joint distribution of X and Y is unknown). However, we can estimate σ_{XY} and σ_X^2 using a random sample.

   Use Theorem 5.5.4 and Slutsky Theorem to prove that

$$\hat{\beta} = \frac{S_{XY}}{S_X^2} \to_p \beta$$

   Remember that

$$\begin{aligned} S_X^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \\ &= \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right) \\ &= \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2\right) \end{aligned}$$

   By the WLLN (provided that $E[x_i^4] < \infty$, why?),

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 \xrightarrow{p} E[x_i^2]$$

$$\bar{x} \xrightarrow{p} \mu_X$$

   and

$$\frac{n}{n-1} \to 1$$

   Sum and multiplication are continuous functions so by the CMT

$$S_X^2 \xrightarrow{p} E[x_i^2] - \mu_X^2 = \sigma_X^2$$

Similarly,

$$
\begin{aligned}
S^2_{XY} &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) \\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}\right) \\
&= \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}\right)
\end{aligned}
$$

$(x_i y_i)$ is independent of $(x_j y_j)$ for $i \neq j$, then by the WLLN applied on $\{x_i y_i\}_{i=1}^n$ (provided that $E[x_i^2 y_i^2] < \infty$)

$$
\frac{1}{n}\sum_{i=1}^{n} x_i y_i \xrightarrow{p} E[x_i y_i]
$$

and

$$
\begin{aligned}
\bar{x} &\xrightarrow{p} \mu_X \\
\bar{y} &\xrightarrow{p} \mu_Y
\end{aligned}
$$

Multiplication is a continuous function so by the CMT

$$
\bar{x}\bar{y} \xrightarrow{p} \mu_X \mu_Y
$$

Putting together all the pieces (sum is again continuous

$$
S^2_{XY} \xrightarrow{p} E[x_i y_i] - \mu_X \mu_Y = \sigma_{XY}
$$

Again division by a non-zero real number is a continuous function (and $\sigma_X^2 > 0$) then

$$
\hat{\beta} = \frac{S_{XY}}{S_X^2} \rightarrow_p \frac{\sigma_{XY}}{\sigma_X^2} = \beta
$$

2. (Uniform MLE) Assume you are given the following random sample from a uniform(a, b) distribution

   $\{0.6849204, 3.216103, 2.789009, 3.023975, 3.42088, 0.5433397, 3.092291, 0.3053189, 2.776194, 4.357245\}$

   a) Recall that the uniform pdf is

$$
f_X(x) = \begin{cases} \frac{1}{b-a} \text{ if } x \in [a,b] \\ 0 \text{ if } x \notin [a,b] \end{cases}
$$

   with $b > a$.

   Therefore, the likelihood function for the uniform distribution is

$$
L(a, b|x_1, .., x_n) = \prod_{i=1}^{n} \frac{1}{b-a} I_{[a,b]}(x_i)
$$

   Recall that the $I_{[a,b]}(x_i)$ function takes the value of one if $x_i \in [a, b]$ and 0 otherwise.

(a) Using the random sample values provided, evaluate the likelihood function above at the following parameter values:

   (i) $a = 2.7, b = 4.3$
   $L(2.7, 4.3|\boldsymbol{x}) = 0$

   (ii) $a = 0.30,\ b = 3.02.$
   $L(0.3, 3.02|\boldsymbol{x}) = 0$

   (iii) $a = 0.68,\ b = 4.5$
   $L(0.68, 4.5|\boldsymbol{x}) = 0$

   (iv) $a = 0.2,\ b = 4.5$
   $L(0.2, 4.5|\boldsymbol{x}) = \frac{1}{(4.5-0.2)^{10}} = 4.6 \times 10^{-7}$

   (v) $a = -2, b = 6$
   $L(-2, 6|\boldsymbol{x}) = \frac{1}{(6-(-2))^{10}} = 9.3 \times 10^{-10}$

(b) If you had to choose a and b out of the five choices in (a), which values would maximize the likelihood function?
$a = 0.2,\ b = 4.5$

(c) Notice that we can write the likelihood function as

$$L(a, b|x_1, .., x_n) = \left[\frac{1}{b-a}\right]^n I_{[a,\infty]}(x_{(1)}) I_{[-\infty,b]}(x_{(n)}),$$

where $x_{(1)} = \min(x_1, ..., x_n)$ and $x_{(n)} = \max(x_1, ..., x_n)$. Explain why.

$$\begin{aligned} L(a, b|x_1, .., x_n) &= \prod_{i=1}^{n} \frac{1}{b-a} I_{[a,b]}(x_i) \\ &= \frac{1}{(b-a)^n} \prod_{i=1}^{n} I_{[a,b]}(x_i) \end{aligned}$$

Note that $\prod_{i=1}^{n} I_{[a,b]}(x_i) = 1 \iff a \le x_{(1)} \le x_{(n)} \le b \iff I_{[a,\infty]}(x_{(1)}) I_{[-\infty,b]}(x_{(n)}) = 1$

$$L(a, b|x_1, .., x_n) = \left[\frac{1}{b-a}\right]^n I_{[a,\infty]}(x_{(1)}) I_{[-\infty,b]}(x_{(n)})$$

(d) What is the MLE for a and b in the general case with a random sample realization given by $x_1, ..., x_n$ that is distributed uniform$[a, b]$.
$a_{MLE} = x_{(1)}$ and $b_{MLE} = x_{(n)}$.

In addition, solve the following problem from Casella and Berger: 7.12.

7.12 (a) $E[X] = (1-\theta)(0) + (\theta)(1) = \theta$ so $\theta_{MME} = \bar{x}$

$$L(\theta|\boldsymbol{x}) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}$$

$$l(\theta|\boldsymbol{x}) = \log\left(\prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}\right)$$

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

The F.O.C. is

$$\begin{aligned}
\frac{dl(\theta|\boldsymbol{x})}{d\theta} = \frac{\sum_{i=1}^n x_i}{\theta} - \frac{\sum_{i=1}^n (1-x_i)}{1-\theta} &= 0 \\
\frac{(1-\theta)\sum_{i=1}^n x_i - \theta\sum_{i=1}^n (1-x_i)}{\theta(1-\theta)} &= 0 \\
\sum_{i=1}^n x_i - n\theta &= 0 \\
\hat{\theta} &= \bar{x}
\end{aligned}$$

So $\theta_{MLE} = \bar{x}$ if $\bar{x} \leq \frac{1}{2}$ given by the restriction on θ. If $\bar{x} > \frac{1}{2}$, then we know that

$$\begin{aligned}
\frac{dl(\theta = \hat{\theta}|\boldsymbol{x})}{d\theta} &= \frac{\sum_{i=1}^n x_i - n\hat{\theta}}{\hat{\theta}(1-\hat{\theta})} \\
&= \frac{n(\bar{x} - \hat{\theta})}{\hat{\theta}(1-\hat{\theta})} > 0
\end{aligned}$$

The inequality comes from $\bar{x} > \frac{1}{2}$ and $\hat{\theta} \leq \frac{1}{2}$. In other words, $l(\theta|\boldsymbol{x})$ is increasing in θ when $\bar{x} > \frac{1}{2}$ and $\hat{\theta} \leq \frac{1}{2}$. Thus, $\hat{\theta} = \frac{1}{2}$.
To sum up,

$$\theta_{MLE} = \begin{cases} \bar{x} & \text{if } \bar{x} \leq \frac{1}{2} \\ \frac{1}{2} & \text{if } \bar{x} > \frac{1}{2} \end{cases}$$

Or alternatively, $\theta_{MLE} = \min\{\frac{1}{2}, \bar{x}\}$.

(b)

$$MSE(\theta_{MME}) = E[(\theta_{MME} - \theta)^2] = Var(\theta_{MME}) + Bias(\theta_{MME})^2 = \theta(1-\theta)/n + 0^2 = \theta(1-\theta)/n$$

$$MSE(\theta_{MLE}) = E[(\theta_{MLE} - \theta)^2] = \sum_{y=0}^{n} (\theta_{MLE} - \theta)^2 \binom{n}{y} \theta^y (1-\theta)^{n-y}$$

with $Y = \sum_{i=1}^n X_i \sim Binomial(n, \theta)$.

(c) $MSE(\theta_{MLE}) < MSE(\theta_{MME})$. Even though the MME is unbiased, it has a higher mean square error (MSE) relative to the MLE.