Problem Set 1

Lionel Claubon - 11.9.24

Exercise 1

1) $$\frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} \left(Y_i^2 - 2 Y_i \bar{Y} + \bar{Y}^2 \right)$$

$$= \frac{1}{n-1} \left(\sum_{i=1}^{n} Y_i^2 - 2 \sum_{i=1}^{n} Y_i \bar{Y} + n \bar{Y}^2 \right)$$, since $n \bar{Y} = \sum_{i=1}^{n} Y_i$

$$= \frac{1}{n-1} \left(\sum_{i=1}^{n} Y_i^2 - 2n \bar{Y}^2 + n \bar{Y}^2 \right)$$

$$= \frac{1}{n-1} \left(\sum_{i=1}^{n} Y_i^2 - n \bar{Y}^2 \right)$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} Y_i^2 - \frac{n}{n-1} (\bar{Y})^2$$

2) We know: $V(Y_i) = E(Y_i^2) - E(Y_i)^2$

Hence: $E(Y_i^2) = V(Y_i) + E(Y_i)^2$

$$E\left(\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y}^2)\right)$$

$$= E\left(\frac{1}{n-1}\sum_{i=1}^{n} Y_i^2 - \frac{n}{n-1}(\bar{Y})^2\right) \quad \text{(from 1.1)}$$

$$= \frac{1}{n-1} n E(Y_i^2) - \frac{n}{n-1} E(\bar{Y}^2))$$

$$= \frac{1}{n-1}(n E(Y_i^2) - n(E\bar{Y}^2))$$

$$= \frac{1}{n-1}(n(V(Y_i) + (E Y_i)^2) - n(E\bar{Y}^2))$$

We also know that: $E(\bar{Y}^2) = V(\bar{Y}) - E(\bar{Y})^2$

and $E(\bar{Y}) = E(Y)$. Also, $V(\bar{Y}) = \frac{1}{n}V(Y)$.

Hence,

$$= \frac{1}{n-1}\left(n(V(Y) + (E Y)^2) - n\left(\left(\frac{V(Y)}{n} - E(Y)^2\right)\right)\right.$$

$$= \frac{1}{n-1}(nV(Y) + n(EY)^2 - V(Y) - n(EY)^2)$$

$$= \frac{1}{n-1}(n-1)V(Y)$$

$$= V(Y).$$

Exercise 2

We are given $Y \sim U(0, \theta)$.

a) $\mathbb{E}(Y) = \frac{\theta}{2}$, thus $\theta = 2\mathbb{E}(Y)$.

b) Since $\mathbb{E}(Y) = \bar{Y}$, let $\hat{\theta}_{MM}$ be s.t. $\bar{Y} = \frac{\hat{\theta}}{2}$ or

$\hat{\theta} = 2\bar{Y}$.

c) Since Y_i's are iid, we invoke the CLT

since $\mathbb{E}(\hat{\theta}) = \mathbb{E}(2\bar{Y}) = 2\mathbb{E}(Y) = \mathbb{E}(\theta)$:

$$\sqrt{n}\,(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}\left(\theta, \mathrm{Var}(\sqrt{n}\,\hat{\theta})\right)^{*}$$

where

$$\begin{aligned}\mathrm{Var}(\sqrt{n}\,2\bar{Y}) &= 4n\,\mathrm{Var}(\bar{Y}) \\ &= 4n\,\frac{V(Y)}{n} \\ &= 4\,V(Y)\end{aligned}$$

I remember this formulation from Moshe's class - I believe he told me this was the expression for finite samples? But I am not sure.

d) Because for any $U \sim [0, \theta]$, θ is the upper bound of the distribution, so it might seem sensible to set the highest observed value (i.e., $\hat{\theta}_{ML}$) as such.

e) We assume now that θ is the upper bound of the distribution. Then,

$$Pr\{\hat{\theta}_{ML} \leq x\} = 0 \quad \text{for} \quad x < 0 \quad \text{and}$$
$$Pr\{\hat{\theta}_{ML} \leq x\} = 1 \quad \text{for} \quad x > 0$$

follow from the properties of a uniform distribution.

Since, by the cdf of a uniform,
$Pr\{X \leq x\} = \frac{x-a}{b-a}$ for $x \in [a, b]$ and $Y_i \sim U[0, \theta]$,
we can write

$$Pr\{\hat{\theta}_{ML} \leq x\} = \frac{x}{\theta}$$

Since $Y_i \sim$ iid $\forall i$, $\prod_i \left[Pr\{\hat{\theta}_{ML} \leq x\}\right] = \left(\frac{x}{\theta}\right)^n$.

f) We want to show:

$$n\left(\frac{\theta - \hat{\theta}_{ML}}{\theta}\right) \xrightarrow{d} F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - e^{x} & \text{if } x > 0 \end{cases}$$

$$P\left(n\frac{\theta - \hat{\theta}_{ML}}{\theta} \leq x\right) = P\left(-\hat{\theta}_{ML} \leq \frac{x\theta}{n} - \theta\right)$$
$$= 1 - P\left(\hat{\theta}_{ML} \leq \theta - \frac{x}{n}\right)$$

From previous result:

$$1 - P\left(\hat{\theta}_{ML} \leq \frac{x\theta}{n} - \theta\right) = 1 - \frac{\theta - \frac{\theta x}{n}}{\theta}$$
$$= 1 - \left(1 - \frac{x}{n}\right)$$

Since we are dealing with n iid variables:

$$1 - \prod_i P\left\{ n \frac{n\theta - \hat{\theta}_{ML}}{\theta} \leq x \right\} = 1 - \left(1 - \frac{x}{n}\right)^n$$

Note: $\lim_{n \to \infty} (1-a)^n = e^{-na}$

Using our definition of convergence,

$$U_n \xrightarrow{d} V \quad \text{if} \quad \lim_{n\to\infty} P(U_n \leq x) = P(V \leq x)$$

Thus: $\lim_{n\to\infty} 1 - \left(1 - \frac{x}{n}\right)^n = 1 - e^{-\frac{nx}{n}} = 1 - e^{-x}$

which proves the desired result.

g) We know that the MLE is unbiased by the properties of the MLE, and have shown above that $E(\hat{\theta}_{MM}) = \theta_{MM}$.

To decide with estimator is best, compare variances. We also know that the variance of the MLE reaches the Cramer-Rao lower bound, the lowest possible asymptotic variance.

Hence, $\hat{\theta}_{MLE}$ is best.

h) See do-file.

$\hat{\theta}_{MM} = 0.965$

$\hat{\theta}_{ML} = 0.998 \longrightarrow$ is closer to 1.

i) Write

$$\hat{\theta}_{ML} \leq \theta \leq \hat{\theta}_{ML} + \hat{\theta}_{ML} \frac{t_{1-\alpha}}{n}$$

$$0 \leq \theta - \hat{\theta}_{ML} \leq \hat{\theta}_{ML} \frac{t_{1-\alpha}}{n}$$

$$0 \leq \frac{\theta - \hat{\theta}_{ML}}{\hat{\theta}_{ML}} \leq \frac{t_{1-\alpha}}{n}$$

$$0 \leq n \frac{\theta - \hat{\theta}_{ML}}{\hat{\theta}_{ML}} \leq t_{1-\alpha}$$

Equivalently,

$$n \frac{\theta - \hat{\theta}_{ML}}{\theta} \cdot \frac{\theta}{\hat{\theta}_{ML}}$$

From previously, we know that $n \frac{\theta - \hat{\theta}_{ML}}{\theta} \xrightarrow{d} U$, an exponential distribution, and $\frac{\hat{\theta}_{ML}}{\theta} \xrightarrow{p} 1$ by the CMT since $\hat{\theta}_{ML} \xrightarrow{p} \theta$.

So $n \frac{\theta - \hat{\theta}_{ML}}{\theta} \cdot \frac{\theta}{\hat{\theta}_{ML}} \xrightarrow{d} F(x) \cdot 1$ by the Slutsky lemma, where $F(x) = 1 - e^{-x}$.

It follows that

$$\lim_{n \to \infty} P(\hat{\theta}_{ML} \in IC(\alpha)) = \lim_{n \to \infty} P\left(0 \leq n \frac{\theta - \hat{\theta}_{ML}}{\theta} \leq t_{1-\alpha}\right)$$

$$\longleftrightarrow \quad F_{t_{1-\alpha}} - F(0) \quad \text{where} \quad F(x) = 1 - e^{-x}$$

$$= F_{t_{1-\alpha}} - 0$$

$$= 1 - e^{t_{1-\alpha}} - 0$$

$1 - e^{t_{1-\alpha}} \approx t_{1-\alpha}$ by Taylor approximation.

QED.

Exercise 3

1) If $U_n \xrightarrow{P} c$, $V_n \xrightarrow{P} c'$, then any function $g(U_n, V_n) \longrightarrow g(c, c')$ by the CMT.
Here, $g(U_n, V_n) = U_n \cdot V_n$, so $g \xrightarrow{P} c \cdot c'$ by the CMT.

2) We know that U_n, V_n converge in probability to c, c', which are real numbers. This implies convergence in distribution to those constants, so we can invoke the Slutsky lemma. The Slutsky lemma tells us that $U_n \cdot V_n$ will converge in probability towards the product of those two numbers, which is also a real number: $c\, c'$.