Problem Set 2

ECON 21020 Spring, 2021

Jake Underland   Collaborated with Jack Surgeoner, Desmond Hui

2021-07-06

Contents

Question 1

a) Find an example of a sequence of random variables $\{X_n\}_{n\geq 1}$ such that $X_n \xrightarrow{p} 0$ but $E(X_n) \nrightarrow 0$ as $n \to \infty$. Why does the expectation fail to converge to 0 in your example?

Solution. Define X_n as below:

$$X_n = \begin{cases} 0 \ldots 1 - \frac{1}{n} \\ 2n \ldots \frac{1}{n} \end{cases}$$

Then, as $n \to \infty$, $P(X_n = 0) \to 1$ and $X_n \to 0$. However,

$$\begin{aligned} \lim_{n\to\infty} E(X) &= \lim_{n\to\infty} 0 \cdot (1 - \frac{1}{n}) + 2n \cdot \frac{1}{n} \\ &= \lim_{n\to\infty} 2 \\ &= 2 \end{aligned}$$

and the expected value fails to converge to 0. This is because as n grows larger, the value of $2n$ grows larger as well, balancing the effect of a lower $P(X_n = 2n)$.

b) Show that the method of moments estimator $\hat{\theta}_{2n} = \sqrt{\frac{3}{n}\Sigma_{i=1}^n X_i^2}$ derived in class is a consistent estimator of θ.

Solution. $X \sim Uni[0, \theta]$. We know that $E(X^2) = \frac{\theta^2}{3}$. Then, by the Law of Large Numbers,

$$\frac{1}{n}\Sigma_{i=1}^n X_i^2 \xrightarrow{p} E(X^2) = \frac{\theta^2}{3}$$

Let $g(x) = \sqrt{3x}$. g is continuous everywhere above 0, and thus continuous at $\frac{\theta^2}{3}$. Then, from the Continuous Mapping Theorem,

$$\begin{aligned}\overline{X^2} \xrightarrow{p} \frac{\theta^2}{3} &\implies g(\overline{X^2}) \xrightarrow{p} g(\frac{\theta^2}{3}) \\ &\implies \sqrt{\frac{3}{n}\Sigma_{i=1}^{n} X_i^2} \xrightarrow{p} \sqrt{\frac{\theta^2}{3} \cdot 3} \\ &\implies \hat{\theta}_{2n} \xrightarrow{p} \theta\end{aligned}$$

Thus, $\hat{\theta}_{2n}$ is consistent.

Question 2

3 Let Y denote the sample average from a random sample with mean μ and variance σ^2. Consider two alternative estimators of μ: $W_1 = [(n-1)/n]\overline{Y}$ and $W_2 = \overline{Y}/2$.

(i) Show that W_1 and W_2 are both biased estimators of μ and find the biases. What happens to the biases as $n \to \infty$? Comment on any important differences in bias for the two estimators as the sample size gets large.

Solution.

$$\begin{aligned}Bias(W_1) &= E[\frac{n-1}{n}\overline{Y}] - \mu \\ &= \frac{n-1}{n}E[\overline{Y}] - \mu \\ &= \frac{n-1}{n}\mu - \mu = -\frac{1}{n}\mu\end{aligned}$$

$$\begin{aligned}Bias(W_2) &= E[\frac{\overline{Y}}{2}] - \mu \\ &= \frac{1}{2}E[\overline{Y}] - \mu \\ &= \frac{\mu}{2} - \mu = -\frac{\mu}{2}\end{aligned}$$

As $n \to \infty$,

$$\begin{aligned}Bias(W_1) &\to \lim_{n\to\infty} -\frac{1}{n}\mu = 0 \\ Bias(W_2) &\to \lim_{n\to\infty} -\frac{\mu}{2} = -\frac{\mu}{2}\end{aligned}$$

Thus, W_1's bias grows smaller as the sample size increases while W_2 remains constant.

(ii) Find the probability limits of W_1 and W_2. Which estimator is consistent?

Solution. From the Law of Large Numbers, we know $\bar{Y} \xrightarrow{p} \mu$. Note also that $\frac{n-1}{n} \xrightarrow{p} 1$. Thus,

$$\text{plim}\frac{n-1}{n}\bar{Y} = \text{plim}\frac{n-1}{n}\text{plim}\bar{Y} = 1 \cdot \mu = \mu$$

$$\text{plim}\frac{\bar{Y}}{2} = \frac{1}{2}\text{plim}\bar{Y} = \frac{1}{2}\mu$$

Thus, W_1 is consistent.

(iii) Find $Var(W_1), Var(W_2)$.

Solution.

$$\begin{aligned}
Var(W_1) &= Var(\frac{n-1}{n}\bar{Y}) \\
&= (\frac{n-1}{n})^2 Var(\frac{1}{n}\Sigma_i Y_i) \\
&= (\frac{n-1}{n})^2 \frac{1}{n^2}\Sigma_i \sigma^2 \ldots (iid) \\
&= \frac{(n-1)^2}{n^3}\sigma^2 \\
Var(W_2) &= Var(\frac{1}{2}\bar{Y}) \\
&= \frac{1}{4}Var(\bar{Y}) \\
&= \frac{\sigma^2}{4n}
\end{aligned}$$

(iv) Argue that W_1 is a better estimator than $\bar{Y}$ if μ is "close" to zero.

Solution.

$$\begin{aligned}
MSE(W_1) &= \frac{(n-1)^2}{n^3}\sigma^2 + \frac{\mu^2}{n^2} \\
MSE[\bar{Y}] &= \frac{\sigma^2}{n}
\end{aligned}$$

When $\mu \approx 0$,

$$MSE[W_1|\mu = 0] = \frac{(n-1)^2}{n^3}\sigma^2 = (\frac{n-1}{n})^2\frac{\sigma^2}{n} < \frac{\sigma^2}{n} = MSE[\bar{Y}|\mu = 0]$$

Thus, W_1 becomes a better predictor.

4 For positive random variables X and Y, suppose the expected value of Y given X is $E[Y|X] = \theta X$. The unknown parameter θ shows how the expected value of Y changes with X.

(i) Define the random variable $Z = Y/X$. Show that $E(Z) = \theta$.

Solution.

$$\begin{aligned}
E[Z|X] &= E[Y/X|X] \\
&= \frac{1}{X}E[Y|X] \\
&= \frac{1}{X}\theta X \\
&= \theta
\end{aligned}$$

Then, by the tower law,

$$\begin{aligned}
& E[E[Z|X]] = E[Z] \\
\implies & E[\theta] = \theta \\
\implies & E[Z] = \theta
\end{aligned}$$

(ii) Use part (i) to prove that the estimator $W_1 = n^{-1}\Sigma_{i=1}^n (Y_i/X_i)$ is unbiased for θ, where $\{(X_i, Y_i) : i = 1, 2, \ldots, n\}$ is a random sample.

Solution.

$$\begin{aligned} E[W_1] &= E[n^{-1}\Sigma_{i=1}^{n}(Y_i/X_i)] \\ &= \frac{1}{n}\Sigma_i E[Y_i/X_i] \\ &= \frac{1}{n}\Sigma_i E[Z_i] \\ &= \frac{1}{n}n\theta \\ &= \theta \end{aligned}$$

(iii) Explain why the estimator $W_2 = \bar{Y}/\bar{X}$ where the overbars denote sample averages, is not the same as W_1. Nevertheless, show that W_2 is also unbiased for θ

Solution. In general, the average of ratios, which is W_1, does not equal the ratio of averages, W_2, thus the two are not completely equivalent expressions.

$$\begin{aligned} E[W_2] &= E[\frac{\bar{Y}}{\bar{X}}] \\ &= E[E[\frac{\bar{Y}}{\bar{X}}|\bar{X}]] \\ &= E[E[\bar{Y}|\bar{X}] \cdot E[\frac{1}{\bar{X}}|\bar{X}]] \\ &= E[\frac{1}{\bar{X}}E[\bar{Y}|\bar{X}]] \\ &= E[\frac{1}{\bar{X}}\theta\bar{X}] \\ &= E[\theta] \\ &= \theta \end{aligned}$$

Thus, W_2 is unbiased.

(iv) (omitted)

Solution.

$$W_1 = \frac{1}{17}\Sigma\frac{Y_i}{X_i} = 0.4179674$$

$$W_2 = \bar{Y}/\bar{X} = 0.4180967$$

The estimates are similar.

Question 3

Let $\{X_i\}$ be an *iid* sample drawn from a distribution F_X which has mean $E(X)$. Consider the estimator

$$\hat{\theta}_n = \Sigma_{i=1}^{n} a_i X_i$$

for some constants $a_1, a_2, \dots, a_n$.

a) Show that if $\hat{\theta}_n$ is an unbiased estimator of $E(X)$, then $\Sigma_{i=1}^{n} a_i = 1$.

Solution.

$$\begin{aligned} E[\hat{\theta}_n] &= E[\Sigma_{i=1}^{n} a_i X_i] \\ &= \Sigma_{i=1}^{n} a_i E[X_i] \\ &= \Sigma_{i=1}^{n} a_i E(X) = E(X) \\ &\implies \Sigma_{i=1}^{n} a_i = 1 \end{aligned}$$

b) Show that $Var(\hat{\theta}_n) = Var(X)\Sigma_{i=1}^n a_i^2$.

Solution.

$$\begin{aligned} Var(\hat{\theta}_n) &= Var(\Sigma_{i=1}^n a_i X_i) \\ &= \Sigma_{i=1}^n Var(a_i X_i) \dots (Independence) \\ &= \Sigma_{i=1}^n a_i^2 Var(X_i) \\ &= Var(X)\Sigma_{i=1}^n a_i^2 \dots (Identical) \end{aligned}$$

c) Find $a_1, \dots, a_n$ that minimize $Var(\hat{\theta}_n)$ subject to the condition that $\hat{\theta}_n$ is an unbiased estimator.

Solution. Since $Var(\hat{\theta}_n) = Var(X)\Sigma_{i=1}^n a_i^2$ and $Var(X)$ is a constant, this problem is equivalent to the one below:

$$\begin{aligned} &\min_{a_i}\{\Sigma_i a_i^2 \;\; s.t. \;\; \Sigma_i a_i = 1\} \\ \implies &\min_{a_i} \mathcal{L} = \Sigma_i a_i^2 + \lambda[1 - \Sigma_i a_i] \\ \text{FOC} \;\; &[a_i]: \;\; 2a_i = \lambda \\ &[\lambda]: \;\; 1 - \Sigma_i a_i = 0 \\ \implies &a_i = \frac{\lambda}{2} \implies \lambda = \frac{2}{n} \implies a_i = \frac{1}{n} \end{aligned}$$

Question 4

Let $\{X_i\}$ be an *iid* sample drawn from a distribution F_X which has mean $E(X)$ and variance $Var(X)$. Define the following estimator of $Var(X)$:

$$\hat{\sigma}_X^2 = \frac{1}{n}\Sigma_{i=1}^n (X_i - \bar{X}_n)^2$$

Show that $\hat{\sigma}_X^2$ is consistent.

Solution.

$$\begin{aligned} \frac{1}{n}\Sigma(X_i - \bar{X}_n)^2 &= \frac{1}{n}\Sigma(X_i^2 - 2X_i\bar{X} + \bar{X}^2) \\ &= \frac{1}{n}\Sigma X_i^2 - 2\bar{X}\underbrace{\frac{1}{n}\Sigma X_i}_{\bar{X}} + \frac{1}{n}\underbrace{\Sigma\bar{X}^2}_{n\bar{X}^2} \\ &= \frac{1}{n}\Sigma X_i^2 - 2\bar{X}^2 + \bar{X}^2 \\ &= \frac{1}{n}\Sigma X_i^2 - \bar{X}^2 \end{aligned}$$

Now, by the Law of Large Numbers,

$$\begin{aligned} \frac{1}{n}\Sigma X_i^2 &\xrightarrow{p} E(X^2) \\ \bar{X} &\xrightarrow{p} E(X) \end{aligned}$$

Let $g(a, b) = a - b^2$. Then, g is continuous at $(E[X^2], E[X])$. Thus, from the Continuous Mapping Theorem, we have that

$$\begin{aligned} \frac{1}{n}\Sigma X_i^2 \xrightarrow{p} E(X^2), \;\; \bar{X} \xrightarrow{p} E(X) &\implies g(\frac{1}{n}\Sigma X_i^2, \bar{X}) \xrightarrow{p} g(E[X^2], E[X]^2) \\ &\implies \frac{1}{n}\Sigma X_i^2 - \bar{X}^2 \xrightarrow{p} E[X^2] - E[X]^2 = Var(X) \end{aligned}$$

Thus, $\hat{\sigma}_X^2$ is consistent.

Question 5

Suppose you observe an *iid* sample of n observations of the (2×1) random vectors $\{(Yi, Xi)\}_{i=1}^n$ (e.g. an individual's height and age), drawn from the bivariate distribution F_{YX}. Suppose that $Var(X) = \sigma_X^2 < \infty$ and $Var(Y) = \sigma_Y^2 < \infty$. Consider estimating $E(X)E(Y)$ using the estimator $\bar{X}_n \cdot \bar{Y}_n$.

a) Is $\bar{X}_n \cdot \bar{Y}_n$ an unbiased estimator of $E(X)E(Y)$ in general?

   Solution. No. In general, $E(AB) = E(A)E(B)$ only strictly holds true when $A \perp\!\!\!\perp B$, and otherwise the equality does not hold necessarily. That's why, when $\bar{X}_n$ and $\bar{Y}_n$ are dependent, other than a few exceptions, the following holds:

$$E(\bar{X}_n \cdot \bar{Y}_n) \neq E(\bar{X}_n)E(\bar{Y}_n) = E(X)E(Y)$$

   An example of this is when $X_i = Y_i$ and X_i is nondegenerate. Then,

$$E(\bar{X}_n \cdot \bar{Y}_n) = E(\bar{X}_n^2) > E(\bar{X}_n)^2 = E(X)^2 = E(X)E(Y)$$

   We know this from Jensen's inequality, and since $g(x) = x^2$ is strictly convex above zero the inequality holds strictly.

b) Is $\bar{X}_n \cdot \bar{Y}_n$ a consistent estimator of $E(X)E(Y)$ in general? If yes, prove it. If no, find a counterexample.

   Solution. Yes, it is. Observe from the Law of Large Numbers the following:

$$\bar{X}_n \xrightarrow{p} E(X)$$
$$\bar{Y}_n \xrightarrow{p} E(Y)$$

   Now, consider $g(x, y) = x \cdot y$. This is continuous in $\mathbb{R}^2$, and we can apply the continuous mapping theorem.

$$g(\bar{X}_n, \bar{Y}_n) \xrightarrow{p} g(E[X], E[Y])$$
$$\implies \bar{X}_n \cdot \bar{Y}_n \xrightarrow{p} E[X] \cdot E[Y]$$

Question 6

a) Suppose n individuals are surveyed about their employment status. Let X_i denote individual i's employment status:

$$X_i = \begin{cases} 1 & \text{if individual } i \text{ is employed,} \\ 0 & \text{otherwise} \end{cases}$$

   Suppose the true rate of employment in the population is p. Let $\{X_i\}_{i=1}^n$ be an *iid* sample and suppose $X_i \sim Bernoulli(p)$. That is:

$$X_i = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1 - p \end{cases}$$

   Show that $E(X_i) = p$ and $Var(X_i) = p(1 - p)$.

   Solution.

$$\begin{aligned} E[X_i] &= 1 \cdot p + (1 - p) \cdot 0 = p \\ Var(X_i) &= E[X_i^2] - E[X_i]^2 \\ &= 1^2 \cdot p + 0^2 \cdot (1 - p) - p^2 \\ &= p - p^2 \\ &= p(1 - p) \end{aligned}$$

b) Let $\bar{X}_n = \frac{1}{n}\Sigma_{i=1}^n X_i$ and $\hat{\sigma}_X^2 = \frac{1}{n}\Sigma_{i=1}^n (X_i - \bar{X}_n)^2$. Show that in this problem $\hat{\sigma}_X^2 = \bar{X}_n(1-\bar{X}_n)$. Is $\hat{\sigma}_X^2$ a biased estimator of $Var(X_i)$? How would you alter $\hat{\sigma}_X^2$ to get an unbiased estimator?

Solution.

$$\begin{aligned}
\hat{\sigma}_X^2 &= \frac{1}{n}\Sigma_{i=1}^n (X_i - \bar{X}_n)^2 \\
&= \frac{1}{n}\Sigma X_i^2 - 2\bar{X}\frac{1}{n}\Sigma X_i + \bar{X}^2 \\
&= \frac{1}{n}\Sigma X_i^2 - \bar{X}^2
\end{aligned}$$

Now, since $X_i \sim Bernoulli(p)$, X_i only takes the values 1 and 0. Thus, for any value of X_i, $X_i^2 = X_i$. Plugging this relationship in to the equation above, we get

$$\begin{aligned}
\frac{1}{n}\Sigma X_i^2 - \bar{X}^2 &= \frac{1}{n}\Sigma X_i - \bar{X}^2 \\
&= \bar{X} - \bar{X}^2 \\
&= \bar{X}(1-\bar{X})
\end{aligned}$$

$\hat{\sigma}_X^2$ is a biased estimator, as we show below:

$$\begin{aligned}
E[\hat{\sigma}_X^2] &= E[\bar{X}(1-\bar{X})] \\
&= E[\bar{X}] - E[\bar{X}^2] \\
&= E[\frac{1}{n}\Sigma X_i] - E[(\frac{1}{n}\Sigma_i X_i)^2] \\
&= \frac{1}{n}\Sigma_i E[X_i] - E[\frac{1}{n^2}\Sigma_i\Sigma_j X_i X_j] \\
&= E[X_i] - \frac{1}{n^2}E[\Sigma_{i=j}X_i^2 + \Sigma_i\Sigma_{j\neq i}X_i X_j] \\
&= E[X_i] - \frac{1}{n}E[X_i^2] - \frac{n^2-n}{n^2}E[X_i X_j] \\
&= E[X_i] - \frac{1}{n}E[X_i] - \frac{n^2-n}{n^2}E[X_i]E[X_j] \ldots (Independence) \\
&= \frac{n-1}{n}(E[X_i] - E[X_i]^2) \ldots (Identical) \\
&= \frac{n-1}{n}(E[X_i^2] - E[X_i]^2) \ldots (\textit{Bernoulli, as expained earlier}) \\
&= \frac{n-1}{n}Var(X_i)
\end{aligned}$$

In order to fix this bias, consider the below estimator:

$$\begin{aligned}
\tilde{\sigma}_X^2 &= \frac{n}{n-1}\hat{\sigma}_X^2 \\
&= \frac{n}{n-1}\cdot\frac{1}{n}\Sigma_{i=1}^n (X_i - \bar{X}_n)^2 \\
&= \frac{1}{n-1}\Sigma_{i=1}^n (X_i - \bar{X}_n)^2
\end{aligned}$$

We know the expectation of this is equal to $Var(X_i)$ through linearity of expectations.

c) Suppose $X_1, \ldots, X_n$ are *iid* with $X_i \sim Bernoulli(p)$ for each i. Let

$$Y_n = \Sigma_{i=1}^n X_i$$

Show that

$$\frac{Y_n - np}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, p(1-p)).$$

Solution. Observe that

$$Y_n = n\bar{X}$$

Then, from the central limit theorem,

$$\bar{X} \xrightarrow{d} \mathcal{N}(E[\bar{X}], Var(\bar{X}))$$

Here,

$$\begin{aligned}
E[\bar{X}] &= E[\frac{1}{n}\Sigma_i X_i] \\
&= \frac{1}{n}\Sigma_i E[X_i] \\
&= \frac{1}{n}np \\
&= p \\
Var(\bar{X}) &= Var(\frac{1}{n}\Sigma_i X_i) \\
&= \frac{1}{n^2}\Sigma_i Var(X_i) \ldots (Independence) \\
&= \frac{1}{n^2}np(1-p) \\
&= \frac{p(1-p)}{n}
\end{aligned}$$

Thus,

$$\begin{aligned}
&\bar{X} \xrightarrow{d} \mathcal{N}(E[\bar{X}], Var(\bar{X})) \\
\implies &\bar{X} \xrightarrow{d} \mathcal{N}(p, \frac{p(1-p)}{n}) \\
\implies &Y_n = n\bar{X} \xrightarrow{d} \mathcal{N}(np, np(1-p))
\end{aligned}$$

Standardizing the above, we get

$$\frac{Y_n - np}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, p(1-p))$$

d) What is the distribution of Y_n? Argue that, suitably centered and scaled, the distribution of a binomial random variable can be approximated by a normal distribution when n is large.

Solution. Y_n is by definition a binomial random variable, since it represents the sum of identical and independent Bernoulli trials. Thus, following from c), we can see that a binomial distribution can be approximated by a normal distribution via the Central Limit Theorem, which presupposes a large n.

e) How would you test the null hypothesis that $p = 0.9$ vs. the two sided alternative at a 5% significance level?

Solution. Following from c), we can compute the 95% confidence interval of p.

$$\begin{aligned}
&\frac{Y_n - np}{\sqrt{np(1-p)}} \sim \mathcal{N}(0,1) \\
\implies &Z_{0.025} \leq \frac{Y_n - np}{\sqrt{np(1-p)}} \leq Z_{0.975} \\
\implies &\frac{Y_n}{n} - 1.96\sqrt{\frac{p(1-p)}{n}} \leq p \leq \frac{Y_n}{n} + 1.96\sqrt{\frac{p(1-p)}{n}} \\
\implies &\bar{X} - 1.96\sqrt{\frac{p(1-p)}{n}} \leq p \leq \bar{X} + 1.96\sqrt{\frac{p(1-p)}{n}}
\end{aligned}$$

Since $E[\bar{X}] = p$, we can use $\bar{X} = \hat{p}$ as an estimator for p, and using Wald's method, construct the following 95% confidence interval:

$$\implies \hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Thus, if $p = 0.9$ is outside of this interval and in the rejection region, we reject the null hypothesis.

f) Your data reveal that $X_n = 0.93$. For what values of n would your test in part e) reject the null hypothesis? In other words, when is 0.93 statistically significantly far from 0.9?

Solution. We solve this by substituting $p = 0.9$ and $\hat{p} = 0.93$ in the above inequality and reversing the inequalities.

$$0.9 < 0.93 - 1.96\sqrt{\frac{0.93(1-0.93)}{n}} \text{ or } 0.9 > 0.93 + 1.96\sqrt{\frac{0.93(1-0.93)}{n}}$$

Rearranging this we get

$$\sqrt{n} > 0.5/0.03,\ \sqrt{n} < -0.5/0.03$$
$$\implies n > 277.777...$$
$$\implies n \geq 278$$