Problem Set #5

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

October 8, 2020

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

Question 1

(a) For any $\varepsilon > 0$, let $N = 1/\varepsilon$. Then, $\frac{1}{n} < \varepsilon$. Therefore,

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ s.t } n \geq N \Rightarrow |a_n - 0| < \varepsilon$$

Thus, $a_n \to 0$ as $n \to \infty$

(b) We have already proven that $\frac{1}{n} \to \infty$. $\forall n$, $sin(\frac{\pi n}{2}) \in [-1, 1]$, so if you let $N = 1/\varepsilon$, it will still be the case that, for any $\varepsilon > 0$, $n \geq N \Rightarrow |a_n - 0| < \varepsilon$

Question 2

Consider:

$$X_n = \begin{cases} -n, & \text{with probability } 1/n \\ 0, & \text{with probability } 1 + 2/n \\ n, & \text{with probability } 1/n \end{cases}$$

(a) X_n converges to zero in probability if, $\forall \varepsilon > 0$:

$$\lim_{n \to 0} P(|X_n - 0| < \varepsilon) = 0$$

From question 1 we know that $\frac{1}{n} \to 0$ as $n \to \infty$. This is true also of $\frac{2}{n}$, so as $n \to \infty$, $1 + \frac{2}{n} \to 1$. Thus, $P(X_n = 0) \to 1$ as $n \to \infty$. So X_n converges to zero in probability.

(b)

$$E(X_n) = P(X_n = -n)(-n) + P(X_n = 0)(0) + P(X_n = n)(n)$$
$$E(X_n) = \left(\frac{1}{n}\right)(-n) + 0 + \left(\frac{1}{n}\right)(n) = -1 + 1 = 0$$

(c)

$$Var(X_n) = E(X_n^2) - [E(X_n)]^2 = P(X_n = -n)(-n)^2 + P(X_n = 0)(0)^2 + P(X_n = n)(n)^2 - 0^2$$
$$Var(X_n) = \left(\frac{1}{n}\right) n^2 + 0 + \left(\frac{1}{n}\right) n^2 = n + n = 2n$$

(d)

$$E(X_n) = P(X_n = 0)(0) + P(X_n = n)(n) = 0 + \left(\frac{1}{n}\right)(n) = 1$$

(e) With the modified distribution from (d), X_n will still converge in probability to 0. Clearly, $X_n \to_p 0$ as $n \to 0$ is not sufficient for $E(X_n) \to_p 0$.

Question 3

Let $\overline{Y}^* = \frac{1}{n}\sum_{i=1}^n w_i Y_i$.

(a)

$$E(\overline{Y}^*) = E(\frac{1}{n}\sum_{i=1}^n w_i Y_i) = \frac{1}{n}\sum_{i=1}^n w_i E(Y_i) = \mu \frac{1}{n}\sum_{i=1}^n w_i = \mu$$

(b) Letting σ^2 be the variance of Y_i,

$$Var(\overline{Y}^*) = Var\left(\sum_{i=1}^n w_i Y_i\right) = \frac{1}{n^2}\sum_{i=1}^n w_i^2 Var(Y_i) = \left(\frac{1}{n^2}\sum_{i=1}^n w_i^2\right)\sigma^2$$

(c) Let $\sum_{i=1}^n w_i^2$ be represented by the constant k. Then, by Chebyshev's inequality, $\forall \varepsilon > 0$,

$$P(|\overline{Y}^* - \mu| \geq \varepsilon) \leq \frac{Var(\overline{Y}^*)}{\varepsilon} = \frac{k\sigma^2}{\varepsilon n^2}$$

where $\frac{k\sigma^2}{\varepsilon n^2}$ is equal to some constant times $\frac{1}{n^2}$, which converges to zero as $n \to \infty$. Thus, $\overline{Y}^* \to_p \mu$.

(d) Another way of writing the probability from (c) is $\frac{\sigma^2}{\varepsilon}\sum_{i=1}^n \left(\frac{w_i}{n}\right)^2$. Thus, if we let $w^* = \max_{i \leq n} w_i$,

$$P(|\overline{Y}^* - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon}\sum_{i=1}^n \left(\frac{w^*}{n}\right)^2$$

So if $\left(\frac{w^*}{n}\right)^2 \to 0$, then this probability also converges to 0.

Question 4

Each of the following answers assumes that the moment in question exists.

(a) $g(x) = x^2$ is continuous, so, under the continous mapping theorem, $E(\frac{1}{n}\sum_{i=1}^n X_i^2) \to_p E(X^2)$ because $E(X_i) \to_p E(X)$

(b) $g(x) = x^3$ is continuous, so, under the continous mapping theorem, $E(\frac{1}{n}\sum_{i=1}^n X_i^3) \to_p E(X^3)$ because $E(X_i) \to_p E(X)$

(c) $g(X_i) = max_{i \leq n}(X_i)$ is not a continuous function, so the continuous mapping theorem and weak law of large numbers cannot tell us anything about this statisic's convergence in probability

(d) This statistic is the same as (a), but minus a constant. Thus, $g(\cdot)$ is continuous and the statistic converges, by the continuous mapping theorem

(e) We know that $\sum_{i=1}^n X_i$ converges to $n\mu$ and that $\sum_{i=1}^n X_i$ converges to $nE(X^2)$. Further, the function applied to X_i, $g(x) = x$, where $x > 0$, is continuous. So, by the continuous mapping theorem and weak law of large numbers, this statistic converges.

(f) $\mathbb{1}(\cdot)$ is not a continuous function and we don't know anything about the distribution of X_i, so the weak law of large numbers and continuous mapping theorem cannot tell us anything about the convergence of this statistic.

Question 5

Since $\{X_1, ..., X_n\}$ is a random sample, we know that $\frac{1}{n}\sum_{i=1}^n X_i \to_p E[X]$. Then, by the continuous mapping theorem, for some continuous function $g(\cdot)$, $\frac{1}{n}\sum_{i=1}^n g(X_i) \to_p g(E[X])$. Now, let $g(x) = log(x)$. Then,

$$\log(\hat{\mu}) = \log\left(\left(\prod_{i=1}^n X_i\right)^{1/n}\right) = \frac{1}{n}\sum_{i=1}^n \log(X_i) \to_p E(\log(X))$$

And by the contraction mapping theorem, $g(X_i) \to_p g(X) \Rightarrow g^{-1}(X_i) \to g^{-1}(X)$. So we can conclude:

$$\hat{\mu} \to_p e^{E(\log(X))} = \mu$$

Question 6

(a) $\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^n X_i^k$ is a consistent estimator of $\mu_k = E(X^k)$, according to the weak law of large numbers.

(b) Let σ be the variance of X^k. Then, $g(x) = \sigma x$ is a continuous function, so $X_i^k \to_d X^k \Rightarrow \sigma X_i^k \to_d \sigma X^k$. By the central limit theorem, since $\{X_i\}$ is an i.i.d. sample, then if $\sigma^2_{X^k} < \infty$, $\frac{\sqrt{n}(\hat{\mu}_k - \mu_k)}{\sigma_{X^k}} \to_d \mathcal{N}(0,1)$. In the last assignment, we fount that $Var(X_k) = \mu_{2k} - \mu_k$. This value is clearly finite. Thus,

$$\sigma_{X^k}\sqrt{n}((\hat{\mu}_k - \mu_k)) \to_d \sigma_{X^k}\mathcal{N}(0,1) \sim \mathcal{N}(0, \mu_{2k} - \mu_k)$$

Question 7

(a) Since $g(x) = 1/k$ is a continuous function and $\hat{\mu}_k$ converges to μ_k, so $\hat{m}_k = \left(\frac{1}{n}\sum_{i=1}^n X_i^k\right)^{1/k}$ is a consistent estimator of $m_k = \left(E(X^k)\right)^{1/k}$ by the continuous mapping theorem.

(b) Given the answer to 6(b), we know that $\sigma_{X^k}\sqrt{n}((\hat{\mu}_k - \mu_k)) \to_d \mathcal{N}(0, \mu_{2k} - \mu_k)$, where $m_k = h(\mu_k)$. Then, since $h(x) = x^{1/k}$ is continuously differentiable, we can use the delta method to derive:

$$\sqrt{n}\left(\hat{\mu}_k^{1/k} - \mu_k^{1/k}\right) \to_d \frac{1}{k}\mu_k^{\frac{k-1}{k}}\mathcal{N}(0, \mu_{2k} - \mu_k)) \sim \mathcal{N}\left(0, \frac{1}{k^2}\mu_k^{\frac{2k-2}{k}}(\mu_{2k} - \mu_k)\right)$$

Question 8

(a)

$$\sqrt{n}(\hat{\beta} - \beta) \to_d 2\mu\mathcal{N}(0, v^2) \sim \mathcal{N}(0, 4(\mu v)^2)$$

(b) If $\mu = 0$, then the variance of $\hat{\beta}'s$ distribution is zero, meaning that $\hat{\beta}$'s distribution is just a mass at 0, equal to 1.

(c) If $\mu = 0$, then we know that $\sqrt{n}\hat{\mu} \to_d \mathcal{N}(0, v^2)$. Thus, we can use the continuous mapping theorem to calculate:

$$\begin{aligned} \frac{\sqrt{n}\hat{\mu}}{v} &\to_d \mathcal{N}(0,1) \\ \left(\frac{\sqrt{n}\hat{\mu}}{v}\right)^2 &\to_d \mathcal{N}(0,1) \\ n\beta &\to_d \mathcal{N}(0, v^4) \end{aligned}$$

(d) In part (a), $\hat{\beta}$ converged to its asymptotic distribution at a rate of $\sqrt{n}$, which had a variance that's proportionate to μ. In (c), we find that, when $\mu = 0$, $\hat{\beta}$ converges to its asymptotic distribution at a rate of n, and its variance is, intuitively, the square of $\hat{\mu}$'s variance. Unlike in (a), the variance does not depend on the value of μ itself.