ECMA31000: Introduction to Empirical Analysis
Solutions to Exam 1

Tuesday, October 26, 2021

INSTRUCTIONS: This is an 80 minute exam. There are 4 questions and a total of 80 points. There is no choice, so answer all questions. You may use any results stated/proven in class without re-proving them, but you must justify all your answers. Write your answers to Questions 1+2 in one blue book and to Questions 3+4 in the other. Good Luck!

Question 1 (16 points)

a) (5 points) Find a sequence of random variables $\{X_n\}_{n\geq 1}$ and a random variable X such that $\mathrm{E}\left([X_n - X]^2\right) \to 0$ as $n \to \infty$ but $\mathrm{E}\left([X_n - X]^4\right) \not\to 0$ as $n \to \infty$.

ANS: Let $X = 0$ and let

$$X_n = \begin{cases} n & \text{with probability } \frac{1}{n^3}, \\ 0 & \text{with probability } 1 - \frac{1}{n^3}. \end{cases}$$

Then $\mathrm{E}\left([X_n - X]^2\right) = \frac{1}{n} \to 0$ but $\mathrm{E}\left([X_n - X]^4\right) = n \to \infty$.

b) (6 points) Suppose $X_n \xrightarrow{p} X$ and $Y_n - X_n \xrightarrow{d} 0$. Prove that $Y_n \xrightarrow{p} X$.

ANS: $Y_n - X_n \xrightarrow{d} 0$ implies $Y_n - X_n \xrightarrow{p} 0$, and since joint convergence in probability is equivalent to convergence in probability of the marginals, we obtain

$$\begin{pmatrix} X_n \\ Y_n - X_n \end{pmatrix} \xrightarrow{p} \begin{pmatrix} X \\ 0 \end{pmatrix}.$$

By the continuous mapping theorem with $g(x, y) = x + y$ we see that:

$$Y_n = X_n + (Y_n - X_n) \xrightarrow{p} X + 0 = X.$$

c) (5 points) Suppose $X_n \xrightarrow{p} X$ and $X_n \xrightarrow{d} Y$. Prove that $\mathrm{P}(X = Y) = 1$ or find a counterexample.

ANS: This is false. Let $X_n = X \sim \mathcal{N}(0, 1)$ for all n and let $Y = -X$. Then $Y \sim \mathcal{N}(0, 1)$ also, so

$X_n \xrightarrow{d} Y$ but

$$\mathrm{P}(X = Y) = \mathrm{P}(2X = 0) = 0.$$

Question 2 (24 points) Our friend from PSET 1 is throwing darts at a dartboard again. The area of the dartboard is 1, and the location of each of their throws is uniformly distributed across the board. Each throw is independent of all others. For each n, let A_n be a subset of the board and define

$$X_n = \begin{cases} 1 & \text{if dart } n \text{ lands in } A_n, \\ 0 & \text{if dart } n \text{ lands outside } A_n. \end{cases}$$

a) (6 points) Explain in non-technical terms what it means for $X_n \xrightarrow{p} 0$ and what it means for $X_n \xrightarrow{a.s.} 0$ in the context of this game. It may help you to write the mathematical descriptions first.

ANS: We say the dart misses if it lands outside A_n. $X_n \xrightarrow{p} 0$ means that as our friend plays more rounds, the probability that the next dart misses approaches 1. $X_n \xrightarrow{a.s.} 0$ means that as our friend plays more rounds, the probability that all future darts miss approaches 1.

In parts **b)** and **c)**, assume A_n has area $\frac{1}{n}$ for each n.

b) (4 points) Prove that $X_n \xrightarrow{p} 0$.

ANS: For any $\epsilon > 0$ we have

$$\mathrm{P}(|X_n - 0| > \epsilon) \leq \mathrm{P}(X_n = 1) = \frac{1}{n} \to 0,$$

where the last equality follows because $\mathrm{P}(X_n = 1)$ is the probability that the dart lands in A_n. Since the throws are uniformly distributed across the board, this equals the area of A_n, which is $\frac{1}{n}$.

c) (6 points) Prove that

$$\mathrm{P}\left(\cap_{k=n}^{\infty} \{X_k = 0\}\right) = 0$$

for all n. Use your answer to part a) to conclude that $X_n \overset{a.s.}{\not\to} 0$.

Note: You may use without proof the fact that if $B_n, B_{n+1}, B_{n+2}, \ldots$ are independent events, then

$$\mathrm{P}\left(\cap_{k=n}^{\infty} B_k\right) = \prod_{k=n}^{\infty} \mathrm{P}(B_k).$$

ANS: Since the throws are independent, the X_n are independent, so

$$\begin{aligned} \mathrm{P}\left(\cap_{k=n}^{\infty} \{X_k = 0\}\right) &= \Pi_{k=n}^{\infty} \mathrm{P}(X_k = 0) \\ &= \Pi_{k=n}^{\infty} \left(\frac{k-1}{k}\right) \\ &= \lim_{N\to\infty} \Pi_{k=n}^{N} \left(\frac{k-1}{k}\right) \end{aligned}$$

$$= \lim_{N\to\infty} \left(\frac{n-1}{N}\right) = 0.$$

Furthermore, $\mathrm{P}\left(\cap_{k=n}^{\infty} \{X_k = 0\}\right)$ is the probability that all future darts miss after the $(n-1)$–st throw. This probability would have to converge to 1 as n increases for $X_n \overset{a.s.}{\to} 0$ to hold, but we see that it is zero for every n, so it cannot.

Now suppose only that A_n has been chosen to satisfy $X_n \overset{p}{\to} 0$ (but not necessarily $X_n \overset{a.s.}{\to} 0$).

d) (8 points) Define

$$Y_n = \begin{cases} 1 & \text{if } X_k = 1 \text{ for all } k \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

Prove that $Y_n \overset{a.s.}{\to} 0$. Hint: Argue that $\{Y_n = 1\} = \cup_{k=n}^{\infty} \{Y_k = 1\}$.

ANS: Note that $\{\{Y_n = 1\}\}_{n\geq 1}$ is a decreasing family of sets, so $\{Y_n = 1\} = \cup_{k\geq n} \{Y_k = 1\}$. To see this, note that if $Y_n = 1$ then $Y_k = 1$ for some $k \geq n$ since $Y_n = 1$. If, on the other hand, $Y_k = 1$ for some $k \geq n$, then it must be that $X_j = 1$ for all $j \leq k$. But $Y_n = 1$ must also hold because $n \leq k$. Therefore $\{Y_n = 1\} = \cup_{k\geq n} \{Y_k = 1\}$, so for any $\epsilon > 0$:

$$\begin{aligned} \mathrm{P}\left(\cup_{k\geq n} \{|Y_k - 0| > \epsilon\}\right) &\leq \mathrm{P}\left(\cup_{k\geq n} \{Y_k = 1\}\right) \\ &= \mathrm{P}\left(Y_n = 1\right) \\ &\leq \mathrm{P}\left(X_n = 1\right) \to 0, \end{aligned}$$

where the last line follows because $Y_n = 1 \implies X_n = 1$ and because $X_n \overset{p}{\to} 0$. An argument more like that we saw in class notes that by definition of Y_n:

$$\{Y_n = 1\} = \cap_{j=1}^{n} \{X_j = 1\}.$$

Therefore, for any $\epsilon > 0$:

$$\begin{aligned} \mathrm{P}\left(\cup_{k\geq n} \{|Y_k - 0| > \epsilon\}\right) &\leq \mathrm{P}\left(\cup_{k\geq n} \{Y_k = 1\}\right) \\ &\leq \sum_{k=n}^{\infty} \mathrm{P}\left(Y_k = 1\right) \\ &= \sum_{k=n}^{\infty} \prod_{j=1}^{k} \mathrm{P}\left(X_j = 1\right) \\ &\leq \sum_{k=n}^{\infty} \prod_{j=n}^{k} \mathrm{P}\left(X_j = 1\right). \end{aligned}$$

where the equality follows by independence of the X_n. Now choose n large such that for all $j \geq n$,

$\mathrm{P}(X_j = 1) < \frac{1}{2}$, which can be done because $X_n \xrightarrow{p} 0$. It follows that for such n

$$\sum_{k=n}^{\infty} \prod_{j=n}^{k} \mathrm{P}(X_j = 1) \leq \sum_{k=n}^{\infty} \frac{1}{2^k} = \frac{1}{2^{n-1}} \to 0,$$

so $Y_n \xrightarrow{a.s.} 0$.

Question 3 (15 points)

a) (5 points) Suppose $X_n = o_p(n)$ and $Y_n = O_p(n^{-1})$. Prove that $X_n Y_n = o_p(1)$ is true or provide a counterexample.

ANS: The statement is true. $X_n = o_p(n)$ means $\frac{X_n}{n} \xrightarrow{p} 0$, so $\frac{X_n}{n} = o_p(1)$. $Y_n = O_p(n^{-1})$ means for all $\epsilon > 0$ there exists B_ϵ such that for all n:

$$\mathrm{P}(|nY_n| > B_\epsilon) \leq \epsilon.$$

Therefore $nY_n = O_p(1)$. So $X_n Y_n = \frac{X_n}{n} \cdot nY_n = o_p(1) \cdot O_p(1) = o_p(1)$.

b) (10 points) Suppose that $X_n = O_p(a_n)$ and $Y_n = O_p(b_n)$ for some sequences of strictly positive numbers $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n\geq 1}$. Show that $X_n + Y_n = O_p(\max\{a_n, b_n\})$.

ANS: For any $\epsilon > 0$ there exist constants B_ϵ and C_ϵ such that for all n:

$$\mathrm{P}(|X_n| \geq B_\epsilon a_n) \leq \frac{\epsilon}{2}; \quad \mathrm{P}(|Y_n| \geq C_\epsilon b_n) \leq \frac{\epsilon}{2}.$$

We need to find D_ϵ such that

$$\mathrm{P}(|X_n + Y_n| \geq D_\epsilon \max\{a_n, b_n\}) \leq \epsilon.$$

Since $|X_n + Y_n| \leq |X_n| + |Y_n|$ by the triangle inequality, and

$$\{|X_n| + |Y_n| \geq D_\epsilon \max\{a_n, b_n\}\} \subset \left\{|X_n| \geq \frac{D_\epsilon}{2} \max\{a_n, b_n\}\right\} \cup \left\{|Y_n| \geq \frac{D_\epsilon}{2} \max\{a_n, b_n\}\right\}$$

It follows that $D_\epsilon \geq 2\max\{B_\epsilon, C_\epsilon\}$ gives:

$$\begin{aligned}
\mathrm{P}(|X_n + Y_n| \geq D_\epsilon \max\{a_n, b_n\}) &\leq \mathrm{P}(|X_n| + |Y_n| \geq D_\epsilon \max\{a_n, b_n\}) \\
&\leq \mathrm{P}(|X_n| \geq B_\epsilon a_n) + \mathrm{P}(|Y_n| \geq C_\epsilon b_n) \\
&\leq \epsilon.
\end{aligned}$$

Note that $D_\epsilon = B_\epsilon + C_\epsilon$ also works, since

$$\{|X_n| + |Y_n| \geq (B_\epsilon + C_\epsilon)\max\{a_n, b_n\}\} \subset \{|X_n| \geq B_\epsilon a_n\} \cup \{|Y_n| \geq C_\epsilon b_n\}.$$

Question 4 (25 points) Suppose we have an iid sample $\{X_i\}_{i=1}^n$ drawn from a distribution represented by the pdf

$$f_\lambda(x) = \begin{cases} (\lambda+1)\, x^\lambda & 0 < x < 1, \\ 0 & \text{otherwise,} \end{cases}$$

for some $\lambda > -1$. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.

a) (5 points) Find $\mathrm{E}(X)$ and show that

$$\hat{\lambda}_n = \frac{1-2\bar{X}_n}{\bar{X}_n - 1}$$

is a method of moments estimator of λ.

ANS: We have

$$\mathrm{E}(X) = \int_0^1 (\lambda+1)\, x^{\lambda+1} \mathrm{d}x = \left[\frac{\lambda+1}{\lambda+2} x^{\lambda+2}\right]_0^1 = \frac{\lambda+1}{\lambda+2}.$$

Now we use the sample analog principle to replace the population quantities with their sample counterparts and solve:

$$\bar{X}_n = \frac{\hat{\lambda}_n + 1}{\hat{\lambda}_n + 2},$$

so

$$\bar{X}_n \hat{\lambda}_n + 2\bar{X}_n = \hat{\lambda}_n + 1 \implies \hat{\lambda}_n = \frac{1-2\bar{X}_n}{\bar{X}_n - 1}$$

is the corresponding method of moments estimator.

b) (5 points) Show that $\hat{\lambda}_n$ is a strongly consistent estimator of λ.

ANS: By the SLLN:

$$\bar{X}_n \overset{a.s.}{\to} \mathrm{E}(X) = \frac{\lambda+1}{\lambda+2}.$$

Now let $g(x) = \frac{1-2x}{x-1}$. This function is continuous for all $x \neq 1$, and for any $\lambda > -1$, $\frac{\lambda+1}{\lambda+2} < 1$, so g is continuous at $\mathrm{E}(X)$. We conclude by the continuous mapping theorem that

$$\hat{\lambda}_n = g\left(\bar{X}_n\right) \overset{a.s.}{\to} g(\mathrm{E}(X)) = \frac{1 - 2\frac{\lambda+1}{\lambda+2}}{\frac{\lambda+1}{\lambda+2} - 1} = \lambda,$$

so $\hat{\lambda}_n \overset{a.s.}{\to} \lambda$, meaning it is strongly consistent.

c) (10 points) Find constants $r \geq 0$ and c such that $n^r\left(\hat{\lambda}_n - c\right) \overset{d}{\to} Y$ for some non-degenerate random variable Y.

Note: If $f(x) = \frac{g(x)}{h(x)}$, then $f'(x) = \frac{h(x)g'(x) - h'(x)g(x)}{h(x)^2}$.

ANS: First check $r = 0$:

$$\left(\hat{\lambda}_n - c\right) \overset{a.s.}{\to} \lambda - c$$

which is degenerate, so this cannot be a possibility. Now choose $r > 0$. If

$$n^r \left(\hat{\lambda}_n - c\right) \xrightarrow{d} Y$$

then by Slutsky's theorem

$$n^{-r} n^r \left(\hat{\lambda}_n - c\right) \xrightarrow{d} 0,$$

meaning $\hat{\lambda}_n \xrightarrow{p} c$ because convergence in distribution to a constant is equivalent to convergence in probability to the same constant. We already know $c = \lambda$ from our previous work. Also recall that $\hat{\lambda}_n$ is a function of $\bar{X}_n$, so we start with the CLT for $\bar{X}_n$:

$$\sqrt{n}\left(\bar{X}_n - \mathrm{E}\left(X\right)\right) \xrightarrow{d} \mathcal{N}\left(0, Var\left(X_i\right)\right),$$

where

$$\begin{aligned} Var\left(X_i\right) &= \mathrm{E}\left(X_i^2\right) - \mathrm{E}\left(X_i\right)^2 \\ &= \int_0^1 \left(\lambda + 1\right) x^{\lambda+2} \mathrm{d}x - \left(\frac{\lambda+1}{\lambda+2}\right)^2 \\ &= \frac{\lambda+1}{\lambda+3} - \left(\frac{\lambda+1}{\lambda+2}\right)^2. \end{aligned}$$

Now choose $g\left(x\right) = \frac{1-2x}{x-1}$ and note that $g'\left(x\right) = \frac{1}{(x-1)^2}$. The delta method provides

$$\sqrt{n}\left(\hat{\lambda}_n - \lambda\right) = \sqrt{n}\left(g\left(\bar{X}_n\right) - g\left(\frac{\lambda+1}{\lambda+2}\right)\right) \xrightarrow{d} \mathcal{N}\left(0, g'\left(\frac{\lambda+1}{\lambda+2}\right)^2 Var\left(X_i\right)\right).$$

Finally, note that

$$g'\left(\frac{\lambda+1}{\lambda+2}\right) = \frac{1}{\left(\frac{\lambda+1}{\lambda+2} - 1\right)^2} = \left(\lambda+2\right)^2,$$

so

$$\sqrt{n}\left(\hat{\lambda}_n - \lambda\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\left(\lambda+1\right)\left(\lambda+2\right)^4}{\left(\lambda+3\right)} - \left(\lambda+1\right)^2\left(\lambda+2\right)^2\right).$$

d) (5 points) Use your answer to part c) to find a function f such that

$$\frac{n^r\left(\hat{\lambda}_n - c\right)}{f\left(\hat{\lambda}_n\right)} \xrightarrow{d} \mathcal{N}\left(0, 1\right),$$

and prove this convergence.

ANS: Since

$$\sqrt{n}\left(\hat{\lambda}_n - \lambda\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\left(\lambda+1\right)\left(\lambda+2\right)^4}{\left(\lambda+3\right)} - \left(\lambda+1\right)^2\left(\lambda+2\right)^2\right),$$

we let

$$f\left(\lambda\right)=\sqrt{\frac{\left(\lambda+1\right)\left(\lambda+2\right)^{4}}{\left(\lambda+3\right)}-\left(\lambda+1\right)^{2}\left(\lambda+2\right)^{2}}.$$

This is a continuous function of λ for all $\lambda>-1$ because

$$\frac{\left(\lambda+1\right)\left(\lambda+2\right)^{4}}{\left(\lambda+3\right)}-\left(\lambda+1\right)^{2}\left(\lambda+2\right)^{2}>0$$

for any $\lambda>-1$, so by the continuous mapping theorem $f\left(\hat{\lambda}_{n}\right)\xrightarrow{p}f\left(\lambda\right)$. By Slutsky's theorem, since $f\left(\lambda\right)>0$:

$$\frac{\sqrt{n}\left(\hat{\lambda}_{n}-\lambda\right)}{f\left(\hat{\lambda}_{n}\right)}\xrightarrow{d}\mathcal{N}\left(0,1\right).$$