ECMA 31000: Solutions to Problem Set 2

Due October 14 by 11:59PM

Question 1 a) Prove, using the definition of convergence in probability, that if X_n is a sequence of random variables such that $\mathrm{E}(X_n) = 0$ and $Var(X_n) = \frac{1}{n}$, then $X_n \xrightarrow{p} 0$.

ANS: By Chebyshev's inequality, for any $\epsilon > 0$:

$$\mathrm{P}(|X_n - 0| > \epsilon) \leq \frac{\mathrm{E}\left(|X_n - 0|^2\right)}{\epsilon^2} = \frac{Var(X_n)}{\epsilon^2} = \frac{1}{n\epsilon^2} \to 0$$

as $n \to \infty$.

b) Fix $(\Omega, \mathcal{F}, \mathrm{P})$. Show that if $A_1, \ldots, A_n$ is any sequence of events in $\mathcal{F}$, then

$$\mathrm{P}(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathrm{P}(A_n).$$

ANS: As in Question 2 of Problem Set 1, let $E_1 = A_1$, $E_2 = A_2 \backslash A_1$,... $E_k = A_k \backslash \left(\cup_{n=1}^{k-1} A_n\right)$. Then

$$\begin{aligned}
\mathrm{P}(\cup_{n=1}^{\infty} A_n) &= \mathrm{P}(\cup_{n=1}^{\infty} E_n) \\
&= \sum_{n=1}^{\infty} \mathrm{P}(E_n) \\
&\leq \sum_{n=1}^{\infty} \mathrm{P}(A_n)
\end{aligned}$$

where the final inequality follows because for each n, $E_n \subset A_n$.

c) If $Var(X_n) = \frac{1}{n^2}$ then $X_n \xrightarrow{a.s.} 0$. (Hint: Use part (b), and the 2nd definition of $\xrightarrow{a.s.}$ discussed in class).

ANS: We need to show that for all $\epsilon > 0$:

$$\mathrm{P}(\cup_{k \geq n} \{|X_k - 0| > \epsilon\}) \to 0.$$

Note that

$$\begin{aligned}
\mathrm{P}(\cup_{k \geq n} \{|X_k - 0| > \epsilon\}) &\leq \sum_{k=n}^{\infty} \mathrm{P}(|X_k - 0| > \epsilon) \\
\text{(Chebyshev's inequality)} &\leq \sum_{k=n}^{\infty} \frac{Var(X_k)}{\epsilon^2}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=n}^{\infty} \frac{1}{k^2 \epsilon^2} \\
&= \frac{1}{\epsilon^2} \left(\sum_{k=1}^{\infty} \frac{1}{k^2} - \sum_{k=1}^{n-1} \frac{1}{k^2} \right) \to 0.
\end{aligned}
$$

Question 2 Let $X_n : [0,1] \to \mathbb{R}$ be a sequence of random variables defined on $([0,1], \mathcal{B}([0,1]), \lambda)$ where $\mathcal{B}([0,1])$ is the Borel sigma algebra on $[0,1]$. The only property of this probability space you will need is that λ satisfies

$$\lambda([a,b]) = \lambda((a,b]) = \lambda((a,b)) = \lambda([a,b)) = b - a$$

for $0 \leq a \leq b \leq 1$. Consider

$$X_n(\omega) = 2^n \mathbf{1}\left(\omega \leq \frac{1}{n}\right)$$

a) Show that $X_n \overset{a.s.}{\to} 0$.

ANS: If $\omega \in (0,1]$, then for all $n > \frac{1}{\omega}$, $X_n = 0$. So $\mathrm{P}(\omega : X_n(\omega) \to X(\omega)) = 1$, and $X_n \overset{a.s.}{\to} 0$.

b) Is it true that $\mathrm{E}(X_n) \to 0$?

ANS: No.

$$\mathrm{E}(X_n) = \frac{2^n}{n} \to \infty.$$

c) Now consider another sequence Y_n of non-negative random variables such that $Y_n \leq K$ for some constant $K > 0$. If $Y_n \overset{a.s.}{\to} 0$, does $\mathrm{E}(Y_n) \to 0$? Prove it or provide a counterexample.

ANS: Yes: In fact $Y_n \overset{p}{\to} 0$ is enough. Fix $\epsilon > 0$ and $\delta > 0$. Choose n large such that $\mathrm{P}(\omega : |Y_n| > \epsilon) < \delta$. Thus

$$Y_n \leq K\mathbf{1}(|Y_n| > \epsilon) + \epsilon\mathbf{1}(|Y_n| \leq \epsilon)$$

Taking expectations yields

$$\mathrm{E}(Y_n) \leq K\delta + \epsilon$$

Since ϵ, δ were arbitrary, the conclusion follows.

Question 3 Suppose that X_n is a sequence of random variables such that $\mathrm{E}(X_n) \to \mu$ and $Var(X_n) \to 0$. Show that $X_n \overset{p}{\to} \mu$.

ANS: Note that

$$
\begin{aligned}
\mathrm{E}\left[(X_n - \mu)^2\right] &= \mathrm{E}\left[(X_n - \mathrm{E}(X_n) + \mathrm{E}(X_n) - \mu)^2\right] \\
&= \mathrm{E}\left[(X_n - \mathrm{E}(X_n))^2\right] + \mathrm{E}\left[(\mathrm{E}(X_n) - \mu)^2\right] \\
&+ 2\mathrm{E}\left[(X_n - \mathrm{E}(X_n))(\mathrm{E}(X_n) - \mu)\right] \\
&= Var(X_n) + (\mathrm{E}(X_n) - \mu)^2 + 0,
\end{aligned}
$$

because

$$2\mathrm{E}\left[\left(X_n-\mathrm{E}\left(X_n\right)\right)\left(\mathrm{E}\left(X_n\right)-\mu\right)\right]=2\left(\mathrm{E}\left(X_n\right)-\mu\right)\mathrm{E}\left[\left(X_n-\mathrm{E}\left(X_n\right)\right)\right]=0.$$

Therefore, if $Var\left(X_n\right)\to 0$ and $\mathrm{E}\left(X_n\right)\to\mu$, then $\mathrm{E}\left[\left(X_n-\mu\right)^2\right]\to 0$, which means $X_n\overset{\text{r-th}}{\to}\mu$ for $r=2$. Since convergence in r-th mean implies convergence in probability, the result follows.

Question 4 Suppose $\{X_i\}_{\geq 1}$ is a sequence of independent random variables with $\mathrm{E}\left(X_i\right)=\mu$ for all $i\geq 1$, and $\max_i\mathrm{E}\left(X_i^4\right)=K<\infty$. Show that $\bar{X}_n\overset{a.s.}{\to}\mu$.

Hint: Assume $\mu=0$ for brevity.

ANS: For brevity assume $\mu=0$. Now note that $\mathrm{E}\left(X_i^2\right)\leq\left[\mathrm{E}\left(X_i^4\right)\right]^{1/2}$, so by the Cauchy-Schwarz inequality, $\mathrm{E}\left(X_i^2X_j^2\right)\leq\max_i\mathrm{E}\left(X_i^4\right)$. Further, by independence,

$$\begin{aligned}\mathrm{E}\left(X_iX_jX_kX_l\right)&=0\text{ when }i\neq j\neq k\neq l\\ \mathrm{E}\left(X_i^2X_jX_k\right)&=0\text{ when }i\neq j\neq k\\ \mathrm{E}\left(X_i^3X_j\right)&=0\text{ when }i\neq j\end{aligned}$$

It follows that

$$\begin{aligned}\mathrm{P}\left(\left|\bar{X}_n-\mu\right|>\epsilon\right)&\leq\frac{\mathrm{E}\left(\left|\bar{X}_n-\mu\right|^4\right)}{\epsilon^4}\\ (\mu=0\text{ WLOG})&=\frac{\mathrm{E}\left(\sum_i\sum_j\sum_k\sum_l X_iX_jX_jX_l\right)}{n^4\epsilon^4}\\ &=\frac{\mathrm{E}\left(\sum_i X_i^4\right)}{n^4\epsilon^4}+\frac{3\mathrm{E}\left(\sum_i\sum_{j\neq i}X_i^2X_j^2\right)}{n^4\epsilon^4}\\ &\leq\frac{nK}{n^4\epsilon^4}+\frac{3n\left(n-1\right)K}{n^4\epsilon^4}\\ &\leq\frac{3K}{n^2\epsilon^4}.\end{aligned}$$

It follows that

$$\begin{aligned}\mathrm{P}\left(\cup_{k\geq n}\left|\bar{X}_k-\mu\right|>\epsilon\right)&\leq\sum_{k=n}^{\infty}\mathrm{P}\left(\left|\bar{X}_k-\mu\right|>\epsilon\right)\\ &\leq\frac{3K}{\epsilon^4}\sum_{k=n}^{\infty}\frac{1}{k^2}\\ &=\frac{3K}{\epsilon^4}\left(\sum_{k=1}^{\infty}\frac{1}{k^2}-\sum_{k=1}^{n-1}\frac{1}{k^2}\right)\to 0\end{aligned}$$

as $n\to\infty$.

Question 5 Show that if $X_n\overset{a.s.}{\to}X$ and $X_n\overset{p}{\to}Y$ then $X_n\overset{a.s.}{\to}Y$.

Hint: If $X_n\overset{a.s.}{\to}X$ but $X_n\overset{a.s.}{\nrightarrow}Y$, then $\mathrm{P}\left(X\neq Y\right)>0$. This means there are constants $c,\delta>0$ such that $\mathrm{P}\left(\left|X-Y\right|>c\right)>\delta$.

ANS: Suppose $X_n \overset{a.s.}{\to} X$ and $X_n \overset{p}{\to} Y$ but $X_n \overset{a.s.}{\nrightarrow} Y$. Note that

$$|X - Y| \leq |X - X_n| + |X_n - Y|,$$

so if $|X - Y| > \epsilon$, then $|X - X_n| + |X_n - Y| > \epsilon$, so either $|X - X_n| > \epsilon/2$ or $|X_n - Y| > \epsilon/2$. Thus

$$\begin{aligned} \mathrm{P}(|X - Y| > \epsilon) &\leq \mathrm{P}(\{|X - X_n| > \epsilon/2\} \cup \{|X_n - Y| > \epsilon/2\}) \\ &\leq \mathrm{P}(|X - X_n| > \epsilon/2) + \mathrm{P}(|X_n - Y| > \epsilon/2) \end{aligned}$$

However, both terms on the RHS converge to 0 because $X_n \overset{a.s.}{\to} X$ and $X_n \overset{p}{\to} Y$. However, choosing $\epsilon \leq c$ gives $\mathrm{P}(|X - Y| > \epsilon) < \delta$, a contradiction.

(Optional: Proof of Hint) Let $A = \{\omega : X_n(\omega) \to X(\omega)\}$ and $B = \{\omega : X(\omega) = Y(\omega)\}$. Then $\mathrm{P}(A) = 1$ since $X_n \overset{a.s.}{\to} X$. Since for any events A, B, $\mathrm{P}(A \cap B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cup B)$, we have

$$\begin{aligned} \mathrm{P}(\{\omega : X_n(\omega) \to X(\omega) \text{ and } X(\omega) = Y(\omega)\}) &= \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cup B) \\ &= 1 + \mathrm{P}(B) - 1 \\ &= \mathrm{P}(\{\omega : X(\omega) = Y(\omega)\}). \end{aligned}$$

Since $X_n \overset{a.s.}{\nrightarrow} Y$, the LHS probability is less than 1, which means $\mathrm{P}(X = Y) < 1$. For the second part, note that $\mathrm{P}(X \neq Y) = \mathrm{P}(|X - Y| > 0) > 0$ and define

$$A_n = \left\{|X - Y| > \frac{1}{n}\right\}.$$

Note that

$$\{|X - Y| > 0\} = \cup_{n=1}^{\infty} A_n$$

and suppose that $\mathrm{P}(A_n) = 0$ for all n. Then

$$\mathrm{P}(\{|X - Y| > 0\}) = \mathrm{P}(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathrm{P}(A_n) = 0,$$

which is a contradiction. So $\mathrm{P}\left(|X - Y| > \frac{1}{n}\right) > 0$ for some n sufficiently large. Choose $0 < c \leq \frac{1}{n}$ and set $\delta = \frac{1}{2} \cdot \mathrm{P}\left(|X - Y| > \frac{1}{n}\right)$. This concludes the proof.

Question 6 Show that for $(K \times 1)$ random vectors $\{X_n\}_{n \geq 1}, X$:

$$\begin{gathered} X_n \overset{a.s.}{\to} X \iff X_{n,i} \overset{a.s.}{\to} X_i \text{ for all } i = 1, \ldots, K; \\ \mathrm{E}(\|X_n - X\|^r) \to 0 \iff \mathrm{E}(|X_{n,i} - X_i|^r) \to 0 \text{ for all } i = 1, \ldots, K. \end{gathered}$$

ANS: Use the inequalities

$$|X_{n,i} - X_i| \leq \|X_n - X\| \leq \sqrt{K} \max_i |X_{n,i} - X_i| \,.$$

From the first inequality,

$$\begin{aligned} \mathrm{E}\left(\|X_n - X\|^r\right) \to 0 &\implies \mathrm{E}\left(|X_{n,i} - X_i|^r\right) \to 0 \text{ for all } i = 1, \dots, K; \\ \|X_n(\omega) - X(\omega)\| \to 0 &\implies |X_{n,i}(\omega) - X_i(\omega)| \to 0 \text{ for all } i = 1, \dots, K. \end{aligned}$$

This proves the forward implications. Now note for convergence in r-th mean:

$$\begin{aligned} \mathrm{E}\left(\|X_n - X\|^r\right) &\leq \sqrt{K}\mathrm{E}\left(\max_i |X_{n,i} - X_i|^r\right) \\ &\leq \sqrt{K}\mathrm{E}\left(\sum_{i=1}^{K} |X_{n,i} - X_i|^r\right) \\ &= \sqrt{K}\sum_{i=1}^{K} \mathrm{E}\left(|X_{n,i} - X_i|^r\right) \to 0 \end{aligned}$$

For convergence a.s. note that if $X_{n,i} \overset{a.s.}{\to} X_i$ for all $i = 1, \dots, K$, then

$$\mathrm{P}\left(\{\omega : X_{n,i}(\omega) \to X_i(\omega) \text{ for all } i\}\right) = 1.$$

To see this note that

$$\begin{aligned} \mathrm{P}\left(\{\omega : X_{n,i}(\omega) \to X_i(\omega) \text{ for all } i\}\right) &= \mathrm{P}\left(\cap_{i=1}^{K} \{\omega : X_{n,i}(\omega) \to X_i(\omega)\}\right) \\ &= 1 - \mathrm{P}\left(\cup_{i=1}^{K} \{\omega : X_{n,i}(\omega) \nrightarrow X_i(\omega)\}\right) \\ &\geq 1 - \sum_{i=1}^{K} \mathrm{P}\left(\{\omega : X_{n,i}(\omega) \nrightarrow X_i(\omega)\}\right) \\ &= 1. \end{aligned}$$

To conclude, let $\omega \in \{\omega : X_{n,i}(\omega) \to X_i(\omega) \text{ for all } i\}$, and consider

$$\|X_n(\omega) - X(\omega)\| \leq \sqrt{K} \max_i |X_{n,i}(\omega) - X_i(\omega)| \to 0.$$

Question 7 Show that if $\mathrm{E}\left(\|X_n - X\|^r\right) \to 0$, then $\mathrm{E}\left(\|X_n - X\|^s\right) \to 0$ for $0 < s < r$.

(Hint: Jensen!)

ANS: By Jensen's inequality, with $g(x) = x^{r/s}$ a convex function, we have:

$$0 \leq \mathrm{E}\left(\|X_n - X\|^s\right)^{r/s} \leq \mathrm{E}\left(\|X_n - X\|^{s \cdot r/s}\right) = \mathrm{E}\left(\|X_n - X\|^r\right) \to 0.$$

Question 8 a) Let $\{X_i\}_{i \geq 1}$ be an iid sequence of $U[0, \theta]$ random variables. For each n, derive the distribution of $\max_{i\leq n} X_i$,and show that $\max_{i\leq n} X_i \xrightarrow{p} \theta$.

ANS: We have

$$\begin{aligned} \mathrm{P}\left(\max_{i\leq n} X_i \leq x\right) &= \mathrm{P}\left(\cap_{i=1}^{n}\left\{X_i \leq x\right\}\right) \\ &= \Pi_{i=1}^{n} \mathrm{P}\left(X_i \leq x\right) \\ &= \begin{cases} 0 & \text{if } x < 0, \\ \left(\frac{x}{\theta}\right)^n & \text{if } 0 \leq x \leq \theta, \\ 1 & \text{if } x > \theta. \end{cases} \end{aligned}$$

It follows that

$$F_{X_n}(x) \to \begin{cases} 0 & x < \theta \\ 1 & x \geq \theta \end{cases},$$

which is the distribution of the constant θ. Since convergence in distribution to a constant implies convergence in probability, the result follows.

b) Show that $n\left(\theta - \max_{i\leq n} X_i\right) \xrightarrow{d} X$ where X has an exponential distribution with CDF

$$F_X(x) = \begin{cases} 0 & x < 0, \\ 1 - \exp\left(-\frac{x}{\theta}\right) & x \geq 0. \end{cases}$$

ANS: If $x < 0$, $\mathrm{P}\left(n\left(\theta - \max_{i\leq n} X_i\right) \leq x\right) = 0$ for all n. If $x \geq 0$:

$$\begin{aligned} \mathrm{P}\left(n\left(\theta - \max_{i\leq n} X_i\right) \leq x\right) &= \mathrm{P}\left(\theta - \frac{x}{n} \leq \max_{i\leq n} X_i\right) \\ &= 1 - \mathrm{P}\left(\max_{i\leq n} X_i \leq \theta - \frac{x}{n}\right) \\ &= 1 - \left(1 - \frac{x/\theta}{n}\right)^n \\ &\to 1 - \exp\left(-\frac{x}{\theta}\right). \end{aligned}$$

Question 9 (Computational Question) Let $\{X_i\}_{i\geq 1}$ be an iid sequence such that $X_i \sim Bernoulli(p)$. That is:

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

a) Show that $\mathrm{E}(X_i) = p$ and $Var(X_i) = p(1-p)$, and that $\mathrm{P}\left(\bar{X}_n - \epsilon < p < \bar{X}_n + \epsilon\right) \geq 1 - \frac{1}{4n\epsilon^2}$. (Hint: Chebyshev!)

ANS:

$$\mathrm{E}(X_i) = 1 \times p + 0 \times (1-p) = p$$

$$
\begin{aligned}
Var\left(X_i\right) &= \mathrm{E}\left(X_i^2\right) - \left[\mathrm{E}\left(X_i\right)\right]^2 \\
&= 1^2 \times p - p^2 \\
&= p\left(1-p\right).
\end{aligned}
$$

Note that

$$
\begin{aligned}
p > \bar{X}_n - \epsilon &\iff \bar{X}_n - p < \epsilon \\
p < \bar{X}_n + \epsilon &\iff \bar{X}_n - p > -\epsilon
\end{aligned}
$$

Therefore,

$$
\bar{X}_n - \epsilon < p < \bar{X}_n + \epsilon \iff \left|\bar{X}_n - p\right| < \epsilon.
$$

Chebyshev's inequality gives

$$
\mathrm{P}\left(\left|\bar{X}_n - p\right| \geq \epsilon\right) \leq \frac{Var\left(X_i\right)}{n\epsilon^2} = \frac{p\left(1-p\right)}{n\epsilon^2}.
$$

The largest value of $p\left(1-p\right)$ is $\frac{1}{4}$, which occurs when $p = \frac{1}{2}$. So, for any p,

$$
\mathrm{P}\left(\left|\bar{X}_n - p\right| < \epsilon\right) = 1 - \mathrm{P}\left(\left|\bar{X}_n - p\right| \geq \epsilon\right) \geq 1 - \frac{1}{4n\epsilon^2}.
$$

b) We call $\left[\bar{X}_n - \epsilon, \bar{X}_n + \epsilon\right]$ a confidence interval for p. Suppose you want to use your bound to ensure that the probability the true parameter p lies inside your confidence interval is at least 0.95. How large a sample must you take if $\epsilon = 0.1$?

ANS: We need to ensure $1 - \frac{1}{4n\epsilon^2} \geq 0.95$, or $4n\epsilon^2 \geq 20$, which means $n \geq 500$.

c) Simulate n iid draws from this distribution with $p = 0.4$, for each of $n = 25, 50, 100$. Let $\epsilon = 0.1$ and compute the confidence intervals for each n based on your simulated data. Does the true value of p lie inside the confidence interval? Repeat this exercise 250 times for each value of n, (though you don't need to display the results of each replication). For each value n, report the proportion of your replications for which the true value of p lies in your confidence interval. Are these proportions generally greater or lower than the bound derived in a)? Why?

Question 9, Part C (Computational Question)

Part 1: Does the true p lie in the CI?

```
p <- 0.4
eps <- 0.1 # epsilon

boolean_array = array(NA,c(1,3)) # to fill
i <- 1
for (n in c(25, 50, 100)) {
  draws <- rbinom(n, 1, 0.4)
  ci <- c(mean(draws) - eps, mean(draws) + eps)
  boolean_array[, i] = (p  >= ci[1] & p <= ci[2])
  i <- i + 1
}
cat("Is p in the CI when n = 25?", boolean_array[, 1],
    "\nIs p in the CI when n = 50?", boolean_array[, 2],
    "\nIs p in the CI when n = 100?", boolean_array[, 3])
```

```
## Is p in the CI when n = 25? TRUE
## Is p in the CI when n = 50? TRUE
## Is p in the CI when n = 100? TRUE
```

Part 2: What is the proportion of replications where the true p lies in the CI?

```
set.seed(257) # for reproducibility
reps <- 20000 # number of repetitions

boolean_array = array(NA,c(reps,3)) # to fill
for (i in 1:reps){
  j <- 1
  for (n in c(25, 50, 100)) {
    draws <- rbinom(n, 1, 0.4)
    ci <- c(mean(draws) - eps, mean(draws) + eps)
    boolean_array[i, j] = (p  >= ci[1] & p <= ci[2])
    j <- j + 1
  }
}

proportions <- colMeans(boolean_array)
cat("Proportion of times p lies in the CI (n = 25):", proportions[1],
    "\nProportion of times p lies in the CI (n = 50)", proportions[2],
    "\nProportion of times p lies in the CI (n = 100):", proportions[3])
```

```
## Proportion of times p lies in the CI (n = 25): 0.69345
## Proportion of times p lies in the CI (n = 50) 0.88875
## Proportion of times p lies in the CI (n = 100): 0.96775
```

Part 3: Are these proportions generally greater or lower than the bound derived in (a)?

In part (a), we showed $P(\bar{X}_n - \epsilon < p < \bar{X}_n + \epsilon) \geq 1 - \frac{1}{4n\epsilon^2}$. For $n = 25, 50, 100$, we compute this bound to be:

```r
cat("Bound (n = 25):",1-1/(1),"\nBound (n = 50)",1-1/(2),"\nBound (n = 100):",1-1/(4))
```

```
## Bound (n = 25): 0
## Bound (n = 50) 0.5
## Bound (n = 100): 0.75
```

Observe that the proportions we obtained from sampling are generally (much) greater than the bound that we derived in (a). While what we derived in (a) gives us a lower bound, the true probability that p lies in the interval is significantly higher, and the proportions we obtain from our simulations will reflect this true probability. Moreover, note that the number of observations that we required in part (b) is greater than what is actually necessary. For example, the Hoeffding bound for Bernoulli random variables gives us a much tighter estimate for sample sizes in that region. In fact, we would need far fewer than 500 observations to satisfy our goal in part (b). The Hoeffding bound gives

$$\mathrm{P}\left(\left|\bar{X}_n - p\right| \leq \epsilon\right) \geq 1 - 2\exp\left(-2\epsilon^2 n\right),$$

which is greater than or equal to 0.95 iff $n \geq 185$.