UC Berkeley
Department of Electrical Engineering and Computer Sciences

ELECTRICAL ENGINEERING 126: PROBABILITY AND RANDOM PROCESSES

Problem Set 4
Fall 2017

Self-Graded Scores Due: 5 PM, Monday, October 2, 2017
Submit your self-graded scores via the Google form:
https://goo.gl/forms/kYqjyQfAhrIFPgaE2.
Make sure you use your **SORTABLE NAME** on CalCentral.

---

1. **Almost Sure Convergence**

   In this question, we will explore almost sure convergence and compare it to convergence in probability. Recall that a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ converges **almost surely** (abbreviated a.s.) to X if $\mathbb{P}(\lim_{n \to \infty} X_n = X) = 1$.

   (a) Suppose that, with probability 1, the sequence $(X_n)_{n \in \mathbb{N}}$ oscillates between two values $a \neq b$ infinitely often. Is this enough to prove that $(X_n)_{n \in \mathbb{N}}$ does *not* converge almost surely? Justify your answer.

   (b) Suppose that Y is uniform on $[-1, 1]$, and X_n has distribution

   $$\mathbb{P}\big(X_n = (y + n^{-1})^{-1} \mid Y = y\big) = 1.$$

   Does $(X_n)_{n=1}^{\infty}$ converge a.s.?

   (c) Define random variables $(X_n)_{n \in \mathbb{N}}$ in the following way: first, set each X_n to 0. Then, for each $k \in \mathbb{N}$, pick j uniformly randomly in $\{2^k, \ldots, 2^{k+1} - 1\}$ and set $X_j = 2^k$. Does the sequence $(X_n)_{n \in \mathbb{N}}$ converge a.s.?

   (d) Does the sequence $(X_n)_{n \in \mathbb{N}}$ from the previous part converge in probability to some X? If so, is it true that $\mathbb{E}[X_n] \to \mathbb{E}[X]$ as $n \to \infty$?

   Solution:

   (a) Yes. If a sequence oscillates between two values infinitely often, then it does not converge. Here, we have a sequence that oscillates between two values infinitely often (with probability 1), which means that the sequence does not converge (with probability 1). (Perhaps we could name this "almost surely not converging"!)

   The above paragraph was very cumbersome to read, which is why we often abbreviate "with probability 1" with a.s. With this abbreviation, here is how the above justification reads: $(X_n)_{n \in \mathbb{N}}$ oscillates between two values infinitely often a.s., so $(X_n)_{n \in \mathbb{N}}$ does not converge a.s.

(b) Yes. Observe that when $Y = y \neq 0$, $(X_n)_{n \in \mathbb{N}}$ will converge to y^{-1}. When $Y = 0$, $(X_n)_{n \in \mathbb{N}}$ does not converge; however, $\mathbb{P}(Y = 0) = 0$ since Y is a continuous random variable. In other words,

$$\mathbb{P}(X_n \text{ does not converge as } n \to \infty) = \mathbb{P}(Y = 0) = 0,$$
$$\mathbb{P}(X_n \text{ converges as } n \to \infty) = \mathbb{P}(Y \neq 0) = 1,$$

so $(X_n)_{n \in \mathbb{N}}$ converges a.s.

(c) No. The sequence $(X_n)_{n \in \mathbb{N}}$ oscillates between 0 and successively higher powers of two infinitely often a.s., so it does not converge a.s.

(d) Yes. Fix $\varepsilon > 0$. For $n \in \mathbb{Z}_+$, one has

$$\mathbb{P}(|X_n| > \varepsilon) = \frac{1}{2^k},$$

where $k = \lfloor \log n \rfloor$. As $n \to \infty$, the above probability goes to 0, so $X_n \to 0$ in probability. Intuitively, $(X_n)_{n \in \mathbb{N}}$ has infinitely many oscillations, so it cannot converge a.s. However, the probability of each oscillation shrinks to 0, so $(X_n)_{n \in \mathbb{N}}$ converges in probability.

The expectations do not converge. For all n, one has $\mathbb{E}[X_n] = 1$, so it is not the case that $\mathbb{E}[X_n] \to 0$ as $n \to \infty$. Hence, convergence in probability is not sufficient to imply that the expectations converge (in fact, almost sure convergence is not sufficient either).

2. **Convergence in Probability**

Let $(X_n)_{n=1}^{\infty}$, be a sequence of i.i.d. random variables distributed uniformly in $[-1, 1]$. Show that the following sequences $(Y_n)_{n=1}^{\infty}$ converge in probability to some limit.

(a) $Y_n = (X_n)^n$.

(b) $Y_n = \prod_{i=1}^{n} X_i$.

(c) $Y_n = \max\{X_1, X_2, \ldots, X_n\}$.

(d) $Y_n = (X_1^2 + \cdots + X_n^2)/n$.

Solution:

(a) For any $\epsilon > 0$, $\mathbb{P}(|Y_n| > \epsilon) = \mathbb{P}(|X_n| > \epsilon^{1/n}) = 1 - \epsilon^{1/n} \to 0$ as $n \to \infty$. Thus, the sequence converges to 0 in probability.

(b) By independence of the random variables,

$$\mathbb{E}[Y_n] = \mathbb{E}[X_1] \cdots \mathbb{E}[X_n] = 0,$$
$$\operatorname{var} Y_n = \mathbb{E}[Y_n^2] = (\operatorname{var} X_1)^n = \left(\frac{1}{3}\right)^n.$$

Now since $\operatorname{var} Y_n \to 0$ as $n \to \infty$, by Chebyshev's Inequality the sequence converges to its mean, that is, 0, in probability.

2

(c) Consider $\epsilon \in [0, 1]$. We see that:

$$\begin{aligned}
\mathbb{P}(|Y_n - 1| \geq \epsilon) &= \mathbb{P}(\max\{X_1, \ldots, X_n\} \leq 1 - \epsilon) \\
&= \mathbb{P}(X_1 \leq 1 - \epsilon, \ldots, X_n \leq 1 - \epsilon) \\
&= \mathbb{P}(X_1 \leq 1 - \epsilon)^n = \left(1 - \frac{\epsilon}{2}\right)^n
\end{aligned}$$

Thus, $\mathbb{P}(|Y_n - 1| \geq \epsilon) \to 0$ as $n \to \infty$ and we are done.

(d) The expectation is

$$\mathbb{E}[Y_n] = \frac{1}{n} \cdot n \, \mathbb{E}[X_1^2] = \frac{1}{3}.$$

Then, we bound the variance.

$$\text{var}\, Y_n = \frac{1}{n} \text{var}\, X_1^2 \leq \frac{1}{n} \to 0 \qquad \text{as } n \to \infty,$$

since $X_1^2 \leq 1$. Hence, we see that $Y_n \to 1/3$ in probability as $n \to \infty$.

Remark: We now provide an interpretation for the previous result. The sample space for Y_n is $\Omega_n = [-1, 1]^n$, which is an n-dimensional cube. The result we have just proved shows that, for any $\epsilon > 0$, the set

$$B_n = \left\{ \mathbf{x} \in \mathbb{R}^n : \frac{1}{3}(1 - \epsilon) \leq \frac{x_1^2 + \cdots + x_n^2}{n} \leq \frac{1}{3}(1 + \epsilon) \right\}$$

makes up "most" of the volume of Ω_n, in the sense that

$$\frac{\text{volume}(B_n \cap [-1, 1]^n)}{2^n} \to 1 \qquad \text{as } n \to \infty.$$

Since B_n is close to the boundary of a ball of radius $\sqrt{n/3}$, the result can be stated facetiously as "nearly all of the volume of a high-dimensional cube is contained in the boundary of a ball". Although this may seem like a meaningless comment, in fact various phenomena such as these contribute to the so-called "curse of dimensionality" in machine learning, which concerns the sparsity of data in high-dimensional statistics.

3. **Coupon Collector Convergence**

Recall the coupon collector problem: for a positive integer n, there are n different coupons, and you are trying to collect them all. Each time you purchase an item, you receive one of the n coupons uniformly at random. Let T_n denote the amount of time it takes to collect all n coupons.

Prove that $T_n/(n \ln n) \to 1$ in probability as $n \to \infty$.

Solution:

Recall that from the analysis of the coupon collector problem, we have

$$\mathbb{E}[T_n] = n \sum_{i=1}^{n} \frac{1}{i} = nH_n,$$

3

where $H_n = \sum_{i=1}^{n} i^{-1}$ is the harmonic sum. We also need to estimate the variance of T_n for this problem. Let X_i denote the amount of time required to collect the ith new coupon, so that $T_n = \sum_{i=1}^{n} X_i$. The X_i are independent, so $\operatorname{var} T_n = \sum_{i=1}^{n} \operatorname{var} X_i$, where X_i is a geometric random variable with probability $p = 1 - (i-1)/n$. Hence, $\operatorname{var} X_i = (1-p)/p^2 \leq p^{-2}$, which gives

$$\operatorname{var} T_n \leq \sum_{i=1}^{n} \left(1 - \frac{i-1}{n}\right)^{-2} = n^2 \sum_{i=1}^{n} \frac{1}{i^2} \leq n^2 \sum_{i=1}^{\infty} \frac{1}{i^2}.$$

It is a well-known fact that $\sum_{i=1}^{\infty} i^{-2} = \pi^2/6$, but we don't need precise bounds for our purposes; it suffices to note that the summation converges. Notably,

$$\operatorname{var} \frac{T_n - nH_n}{n \ln n} \leq \frac{\not{n^2}}{\not{n^2}(\ln n)^2} \sum_{i=1}^{\infty} \frac{1}{i^2} \to 0 \qquad \text{as } n \to \infty,$$

and now Chebyshev's Inequality gives

$$\mathbb{P}\left(\left|\frac{T_n - nH_n}{n \ln n}\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \operatorname{var} \frac{T_n - nH_n}{n \ln n} \to 0 \qquad \text{as } n \to \infty$$

for every $\varepsilon > 0$. Hence, $(T_n - nH_n)/(n \ln n) \to 0$ in probability as $n \to \infty$. To conclude, we note that $H_n \sim \ln n$ asymptotically. Hence, $T_n/(n \ln n) \to 1$ in probability as $n \to \infty$.

Remark. From our previous analysis of the coupon collector problem, we know that $\mathbb{E}[T_n]$ is close to $n \ln n$, so we have shown a result which is similar in spirit to a "weak law of large numbers for the coupon collector problem": as $n \to \infty$, T_n is "close" to its expected value. However, since we are not dealing with i.i.d. random variables, we cannot use the variant of the WLLN proved in lecture to deal with this problem.

4. **Confidence Interval Comparisons**

In order to estimate the probability of a head in a coin flip, p, you flip a coin n times, where n is a positive integer, and count the number of heads, S_n. You use the estimator $\hat{p} = S_n/n$.

(a) You choose the sample size n to have a guarantee

$$\mathbb{P}(|\hat{p} - p| \geq \epsilon) \leq \delta.$$

Using Chebyshev Inequality, determine n with the following parameters:

(i) Compare the value of n when $\epsilon = 0.05$, $\delta = 0.1$ to when $\epsilon = 0.1, \delta = 0.1$.

(ii) Compare the value of n when $\epsilon = 0.1, \delta = 0.05$ to when $\epsilon = 0.1, \delta = 0.1$.

(b) Now, we change the scenario slightly. You know that $p \in (0.4, 0.6)$ and would now like to determine the smallest n such that

$$\mathbb{P}\left(\frac{|\hat{p} - p|}{p} \leq 0.05\right) \geq 0.95.$$

Use the CLT to find the value of n that you should use.

Solution:

(a) Chebyshev Inequality implies that:

$$\mathbb{P}\Big(\Big|\frac{S_n}{n} - p\Big| \geq \epsilon\Big) \leq \frac{\mathrm{var}(S_n/n - p)}{\epsilon^2} = \frac{p(1-p)}{n\epsilon^2}$$

Thus, we set $\delta = p(1-p)/(n\epsilon^2)$ or $n = p(1-p)/(\delta\epsilon^2)$. Thus, when ϵ is reduced to half of its original value, n is changed to 4 times its original value, and when δ is reduced to half of its original value, n will be twice its original value. In order to be more concrete, we may maximize $p(1-p)/(\delta\epsilon^2)$ by letting $p = 1/2$. Thus, when $\epsilon = 0.1, \delta = 0.1$, $n = 250$. Letting $\delta = 0.05$ results in $n = 500$, while letting $\epsilon = 0.05$ results in $n = 1000$.

(b) Note that by the CLT:

$$\sqrt{n}\frac{\hat{p} - p}{\sqrt{p(1-p)}} \sim \mathcal{N}(0, 1)$$

We are interested in the following:

$$\mathbb{P}\Big(\frac{|\hat{p} - p|}{p} \leq 0.05\Big) \equiv \mathbb{P}\Big(\Big|\sqrt{n}\frac{\hat{p} - p}{\sqrt{p(1-p)}}\Big| \leq 0.05\frac{\sqrt{np}}{\sqrt{1-p}}\Big)$$

$$\approx \mathbb{P}\Big(|\mathcal{N}(0,1)| \leq 0.05\frac{\sqrt{np}}{\sqrt{1-p}}\Big)$$

Now, we use the condition that we want:

$$\mathbb{P}\Big(|\mathcal{N}(0,1)| \leq 0.05\frac{\sqrt{np}}{\sqrt{1-p}}\Big) \geq 0.95$$

This implies that $0.05\sqrt{np/(1-p)} \geq 2$ (note we use 2 here for simplicity, if you used 1.96, this is completely correct), or equivalently, $n \geq 1600(1-p)/p$. We now use the fact that we know $p \in [0.4, 0.6]$. Since $p \in [0.4, 0.6]$, we can see that the value $(1-p)/p$ is maximized when $p = 0.4$. Thus, we note that $n \geq 1600(1-p)/p$ for all values of p, so the minimum value of n must be the maximum valid value of $1600(1-p)/p = 2400$.

5. **Poisson Bounds**

Let X be the sum of 20 i.i.d. Poisson random variables $X_1, \ldots, X_{20}$ with $\mathbb{E}[X_1] = 1$. Use Markov's Inequality, Chebyshev's Inequality, and the Chernoff Bound to find an upper bound of $\mathbb{P}(X \geq 26)$. Use the CLT to estimate $\mathbb{P}(X \geq 26)$.

Solution:

(a) Using Markov's Inequality:

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a},$$

for all $a > 0$. So,

$$\mathbb{P}(X \geq 26) \leq \frac{20}{26} \approx 0.769.$$

(b) Using Chebyshev's Inequality:

$$\mathbb{P}(|X - E[X]| \geq c) \leq \frac{\sigma_X^2}{c^2},$$

so we have

$$\mathbb{P}(|X - 20| \geq 6) \leq \frac{20}{36} \approx 0.5556.$$

(c) Using the Chernoff Bound we have

$$\mathbb{P}(X \geq 26) \leq \min_{s \geq 0} e^{-26s} e^{20(e^s - 1)} = \min_{s \geq 0} e^{-26s + 20e^s} e^{-20}.$$

To minimize this function we only have to minimize $-26s + 20e^s$. Taking the derivative and setting it equal to zero, we have

$$-26 + 20e^s = 0,$$
$$e^s = \frac{26}{20},$$
$$s = \ln \frac{26}{20} \approx 0.26236.$$

Plugging in s to the above equation, we have

$$\mathbb{P}(X \geq 26) \leq e^{-26s + 20e^s} e^{-20} \Big|_{s=\ln(13/10)} \approx 0.4398.$$

(d) The central limit approximation yields

$$\mathbb{P}(X \geq 26) \approx 1 - \Phi\Big(\frac{26 - \mathbb{E}[X]}{\sigma_X}\Big) = 1 - \Phi\Big(\frac{26 - 20}{\sqrt{20}}\Big) \approx 1 - \Phi(1.34)$$
$$\approx 0.0899.$$

6. **Breaking a Stick**

I break a stick n times, where n is a positive integer, in the following manner: the ith time I break the stick, I keep a fraction X_i of the remaining stick where X_i is uniform on the interval $[0, 1]$ and $X_1, X_2, \ldots, X_n$ are i.i.d. Let $P_n = \prod_{i=1}^{n} X_i$ be the fraction of the original stick that I end up with. Find P such that $P_n^{1/n} \to P$ a.s. as $n \to \infty$ and also compute $\mathbb{E}[P_n]^{1/n}$.

Solution:

We will first find $\lim_{n \to \infty} P_n^{1/n}$. We have:

$$\lim_{n \to \infty} P_n^{1/n} = \lim_{n \to \infty} \Big(\prod_{i=1}^{n} X_i\Big)^{1/n}$$

Noting the product and that exp is a continuous function, we consider

$$\lim_{n \to \infty} \ln P_n^{1/n} = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \ln X_i$$
$$= \mathbb{E}[\ln X_1]$$

6

where the second equality follows from the Strong Law of Large Numbers.
Note that $\mathbb{E}[\ln X_1] = \int_0^1 \ln x \, \mathrm{d}x = -1$. Thus, we have:

$$\lim_{n\to\infty} P_n^{1/n} = \lim_{n\to\infty} \mathrm{e}^{n^{-1}\ln P_n} = \frac{1}{\mathrm{e}}.$$

Now, we tackle the second part. We note that

$$\mathbb{E}[P_n] = \mathbb{E}\Big[\prod_{i=1}^{n} X_i\Big] = \mathbb{E}[X_1]^n = \Big(\frac{1}{2}\Big)^n.$$

Thus,

$$\mathbb{E}[P_n]^{1/n} = \frac{1}{2} \neq \lim_{n\to\infty} P_n^{1/n}.$$