UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: PROBABILITY AND RANDOM PROCESSES

Homework 04
Spring 2023

1. **Moment-Generating Functions Practice**

The **moment-generating function** (mgf) of a random variable X is the function

$$M_X(s) = \mathbb{E}(e^{sX}) = \mathbb{E}\left(\sum_{k=0}^{\infty} \frac{(sX)^k}{k!}\right) = \sum_{k=0}^{\infty} \frac{s^k}{k!}\,\mathbb{E}(X^k).$$

In this class, we will not worry about technical details about the convergence of Taylor series, so we will say that the mgf is equal to any of the expressions above.

The mgf gets its name because it is the function *generating* the moments $\mathbb{E}(X^p)$, $p \geq 1$, of X. More specifically, by evaluating the pth derivative of the mgf at $s = 0$, we have a method to explicitly find the pth moment of X from its mgf:

$$\left[\frac{\mathrm{d}^p}{\mathrm{d}s^p} M_X(s)\right]_{s=0} = \left[\sum_{k=p}^{\infty} \frac{s^{k-p}}{p!}\,\mathbb{E}(X^k)\right]_{s=0} = \mathbb{E}(X^p) + \sum_{k=p+1}^{\infty} 0 = \mathbb{E}(X^p).$$

Consider a random variable Z with moment-generating function

$$M_Z(s) = \frac{a - 3s}{s^2 - 6s + 8} \qquad \text{for } |s| < 2.$$

Calculate the following quantities:

a. The numerical value of the parameter a.

b. $\mathbb{E}(Z)$.

c. $\mathrm{var}(Z)$.

Solution:

a. By definition, we know that $M_Z(s) = \mathbb{E}(e^{sZ})$, so we must have

$$M_Z(0) = \mathbb{E}(e^{0Z}) = 1 = \frac{a}{8},$$

from which it follows that $a = 8$.

b. We find the first moment as

$$\mathbb{E}(Z) = \left[\frac{d}{ds} M_Z(s)\right]_{s=0} = \left[\frac{2}{(4-s)^2} + \frac{1}{(2-s)^2}\right]_{s=0} = \frac{3}{8}.$$

c. We find that $\mathrm{var}(Z) = \frac{11}{64}$, where the second moment is

$$\mathbb{E}(Z^2) = \left[\frac{d^2}{ds^2} M_Z(s)\right]_{s=0} = \left[\frac{4}{(4-s)^3} + \frac{2}{(2-s)^3}\right]_{s=0} = \frac{5}{16}.$$

2. **Transforms and Independence**

In this problem, we will make use of multivariate moment generating functions, defined for a random vector $X = (X_1, \dots, X_n)$ as

$$M_X(t) = \mathbb{E}(e^{t \cdot X}) = \mathbb{E}(e^{\sum_{i=1}^n t_i X_i})$$

for $t = (t_1, \dots, t_n) \in \mathbb{R}^n$. You may assume that MGFs are unique: if $M_X(t) = M_Y(t)$ for all t, then $X \sim Y$.

Consider the random vector $X = (X_1, \dots, X_n)$. Show that $X_1, \dots, X_n$ are independent if and only if for all $t \in \mathbb{R}^n$,

$$M_X(t) = \prod_{i=1}^n M_{X_i}(t_i).$$

Solution: First suppose that the X_i are independent. Then for any $t \in \mathbb{R}^n$, we have

$$M_X(t) = \mathbb{E}(e^{t \cdot X}) = \mathbb{E}\left(\prod_{i=1}^n e^{t_i X_i}\right) = \prod_{i=1}^n \mathbb{E}(e^{t_i X_i}) = \prod_{i=1}^n M_{X_i}(t_i).$$

Conversely, suppose the product identity holds. Let $X_i' \sim X_i$ be copies such that $X_1', \dots, X_n'$ are independent. For the corresponding random vector $X' = (X_1', \dots, X_n')$, we have

$$M_X(t) = \prod_{i=1}^n \mathbb{E}(e^{t_i X_i}) = \prod_{i=1}^n \mathbb{E}(e^{t_i X_i'}) = M_{X'}(t)$$

for all $t \in \mathbb{R}^n$. But by the uniqueness property of MGFs, this implies that $(X_1, \dots, X_n) \stackrel{\mathsf{d}}{=} (X_1', \dots, X_n')$. Hence $X_1, \dots, X_n$ are independent.

3. **Coupon Collector Bounds**

   Recall the coupon collector's problem, in which there are n different types of coupons. Every box contains a single coupon, and we let the random variable X be the number of boxes bought until one of every type of coupon is obtained. The expected value of X is nH_n, where $H_n := \sum_{i=1}^n \frac{1}{i}$ is the *harmonic number* of order n, which satisfies the inequality

   $$\ln n \le H_n \le \ln n + 1.$$

   a. Use Markov's inequality in order to show that

   $$\mathbb{P}(X > 2nH_n) \le \frac{1}{2}.$$

   b. Use Chebyshev's inequality in order to show that

   $$\mathbb{P}(X > 2nH_n) \le \frac{\pi^2}{6(\ln n)^2}.$$

   Note: You can use Euler's solution to the Basel problem, the identity $\sum_{i=1}^\infty \frac{1}{i^2} = \frac{\pi^2}{6}$.

   c. Define appropriate events and use the union bound in order to show that

   $$\mathbb{P}(X > 2nH_n) \le \frac{1}{n}.$$

   Note: $a_n = (1 - \frac{1}{n})^n$ is a strictly increasing sequence with limit e^{-1}.

   Solution:

   a. We are given $\mathbb{E}(X) = nH_n$, so

   $$\mathbb{P}(X > 2nH_n) \le \frac{\mathbb{E}(X)}{2nH_n} = \frac{1}{2}.$$

   b. We can write X as an independent sum $\sum_{i=1}^n X_i$, where $X_i \sim \text{Geometric}(\frac{n-i+1}{n})$, so

   $$\text{var}(X) = \sum_{i=1}^n \text{var}(X_i) < \sum_{i=1}^n \left(\frac{n}{n-i+1}\right)^2 = \sum_{i=1}^n \left(\frac{n}{i}\right)^2 < \frac{\pi^2 n^2}{6}.$$

   Using Chebyshev's inequality, we have that

   $$\mathbb{P}(X > 2nH_n) \le \mathbb{P}(|X - nH_n| > nH_n) \le \frac{\text{var}(X)}{(nH_n)^2} < \frac{\pi^2}{6H_n^2} \le \frac{\pi^2}{6(\ln n)^2}.$$

   c. Let A_i be the event that we fail to get box i after $2nH_n$ tries.

   $$\mathbb{P}(A_i) \le \left(\frac{n-1}{n}\right)^{2nH_n} = \left[\left(1 - \frac{1}{n}\right)^n\right]^{2H_n} < e^{-2H_n} \le e^{-2\ln n} = \frac{1}{n^2}.$$

   Now, by the union bound, we can conclude that

   $$\mathbb{P}(X > 2nH_n) = \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \le \sum_{i=1}^n \mathbb{P}(A_i) \le \frac{1}{n}.$$

4. **Confidence Interval Comparisons**

   In order to estimate the probability p of heads of a coin flip, you flip a coin $n \geq 1$ times and count the number of heads S_n. You use the estimator $\hat{p} = \frac{S_n}{n}$, and you choose the sample size n to have the guarantee

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

   Using Chebyshev's inequality, determine n for the following parameters. You should not have p in your final answer.

   a. Compare the value of n when $\varepsilon = 0.05, \delta = 0.1$ to the value of n when $\varepsilon = 0.1, \delta = 0.1$.

   b. Compare the value of n when $\varepsilon = 0.1, \delta = 0.05$ to the value of n when $\varepsilon = 0.1, \delta = 0.1$.

   Solution: By Chebyshev's inequality,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{\text{var}(\frac{S_n}{n} - p)}{\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2}.$$

   Thus, we set $\delta = p(1-p)/(n\epsilon^2)$, i.e. $n = p(1-p)/(\delta\varepsilon^2)$.

   When ε is reduced to half of its original value, n is scaled to 4 times its original value, and when δ is reduced to half of its original value, n will be twice its original value.

   To be more concrete, we can maximize $p(1-p)/(\delta\varepsilon^2)$ at $p = \frac{1}{2}$. Thus, when $\varepsilon = 0.1, \delta = 0.1$, we have $n = 250$. Changing δ to 0.05 results in $n = 500$, while letting $\varepsilon = 0.05$ results in $n = 1000$.

5. **Convergence in L^p**

Let $p \geq 1$. A sequence of random variables $(X_n)_{n\geq 1}$ is said to **converge in** L^p (norm) to a random variable X if

$$\lim_{n\to\infty} \mathbb{E}(|X_n - X|^p) = 0.$$

Prove that if $X_n \to X$ in L^p, then $X_n \to X$ in probability.

Solution: Note that for $p \geq 1$, $x \mapsto x^p$ is a monotonic function. By Markov's inequality,

$$\mathbb{P}(|X_n - X| \geq \varepsilon) = \mathbb{P}(|X_n - X|^p \geq \varepsilon^p) \leq \frac{\mathbb{E}(|X_n - X|^p)}{\varepsilon^p}.$$

If $X_n \to X$ in L^p, i.e. $\mathbb{E}(|X_n - X|^p) \to 0$, then $\mathbb{P}(|X_n - X| \geq \varepsilon) \to 0$ for any $\varepsilon > 0$, which is precisely convergence in probability.

6. **Breaking a Stick**

I break a stick n times, $n \geq 1$, in the following manner: the ith time I break the stick, I keep a fraction $X_i \sim \text{Uniform}((0,1])$ of the remaining stick. Suppose that $X_1, X_2, \ldots, X_n$ 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 at time n.

a. Show that $P_n^{1/n}$ converges almost surely, and find its limit.

b. Compute $\mathbb{E}(P_n)^{1/n}$.

c. Now compute $\mathbb{E}(P_n^{1/n})$. Do you find the same answer as in part b? Is the limit of $\mathbb{E}(P_n^{1/n})$ equal to the limit you found in part a?

Solution:

a. By the Strong Law of Large Numbers,

$$\lim_{n\to\infty} \ln P_n^{1/n} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \ln X_i = \mathbb{E}(\ln X_1) \text{ a.s.}$$

We also find that $\mathbb{E}(\ln X_1) = \int_0^1 \ln x \, \mathrm{d}x = -1$. Thus, we have that

$$\mathbb{P}(\ln P_n^{1/n} \to -1) = \mathbb{P}(P_n^{1/n} \to e^{-1}) = 1,$$

so the almost-sure limit is e^{-1}.

b. By independence,

$$\mathbb{E}(P_n)^{1/n} = \mathbb{E}\left(\prod_{i=1}^n X_i\right)^{1/n} = (\mathbb{E}(X_1)^n)^{1/n} = \mathbb{E}(X_1) = \frac{1}{2}.$$

c. We now find that

$$\mathbb{E}(P_n^{1/n}) = \mathbb{E}(X_1^{1/n}) \cdots \mathbb{E}(X_n^{1/n}) = \mathbb{E}(X_1^{1/n})^n = \left(\int_0^1 x^{1/n} \, \mathrm{d}x\right)^n = \left(\frac{n}{n+1}\right)^n.$$

This differs from part b because the expectation of a nonlinear transformation is not necessarily equal to the nonlinear transformation of the expectation. However,

$$\lim_{n\to\infty} \mathbb{E}(P_n^{1/n}) = \lim_{n\to\infty} \left(\frac{n}{n+1}\right)^n = \lim_{n\to\infty}\left(1 - \frac{1}{n+1}\right)^n = e^{-1},$$

which is indeed the almost-sure limit of $P_n^{1/n}$ we found in part a.