CS 70 Discrete Mathematics and Probability Theory
Summer 2022 Jingjia Chen, Michael Psenka and Tarang Srivastava HW 7

Due: Thursday 8/4, 11:59 PM
Grace period until Friday 8/5, 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Law of Large Numbers

Recall that the *Law of Large Numbers* holds if, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left[\left|\frac{1}{n}S_n - \mathbb{E}\left[\frac{1}{n}S_n\right]\right| > \varepsilon\right] = 0.$$

In class, we saw that the Law of Large Numbers holds for $S_n = X_1 + \cdots + X_n$, where the X_i's are i.i.d. random variables. This problem explores if the Law of Large Numbers holds under other circumstances.

Packets are sent from a source to a destination node over the Internet. Each packet is sent on a certain route, and the routes are disjoint. Each route has a failure probability of $p \in (0,1)$ and different routes fail independently. If a route fails, all packets sent along that route are lost. You can assume that the routing protocol has no knowledge of which route fails.

For each of the following routing protocols, determine whether the Law of Large Numbers holds when S_n is defined as the total number of received packets out of n packets sent. Answer **Yes** if the Law of Large Number holds, or **No** if not. Give a justification of your answer. (Whenever convenient, you can assume that n is even.)

(a) **Yes** or **No**: Each packet is sent on a completely different route.

(b) **Yes** or **No**: The packets are split into $n/2$ pairs of packets. Each pair is sent together on its own route (i.e., different pairs are sent on different routes).

(c) **Yes** or **No**: The packets are split into 2 groups of $n/2$ packets. All the packets in each group are sent on the same route, and the two groups are sent on different routes.

(d) **Yes** or **No**: All the packets are sent on one route.

2 Practical Confidence Intervals

(a) It's New Year's Eve, and you're re-evaluating your finances for the next year. Based on previous spending patterns, you know that you spend \$1500 per month on average, with a standard deviation of \$500, and each month's expenditure is independently and identically distributed. As a college student, you also don't have any income. How much should you have in your bank account if you don't want to run out of money this year, with probability at least 95%?

(b) As a UC Berkeley CS student, you're always thinking about ways to become the next billionaire in Silicon Valley. After hours of brainstorming, you've finally cut your list of ideas down to 10, all of which you want to implement at the same time. A venture capitalist has agreed to back all 10 ideas, as long as your net return from implementing the ideas is positive with at least 95% probability.

Suppose that implementing an idea requires 50 thousand dollars, and your start-up then succeeds with probability p, generating 150 thousand dollars in revenue (for a net gain of 100 thousand dollars), or fails with probability $1-p$ (for a net loss of 50 thousand dollars). The success of each idea is independent of every other. What is the condition on p that you need to satisfy to secure the venture capitalist's funding?

(c) One of your start-ups uses error-correcting codes, which can recover the original message as long as at least 1000 packets are received (not erased). Each packet gets erased independently with probability 0.8. How many packets should you send such that you can recover the message with probability at least 99%?

3 Just One Tail, Please

Let X be some random variable with finite mean and variance which is not necessarily non-negative. The *extended* version of Markov's Inequality states that for a non-negative function $\phi(x)$ which is monotonically increasing for $x > 0$ and some constant $\alpha > 0$,

$$\mathbb{P}[X \geq \alpha] \leq \frac{\mathbb{E}[\phi(X)]}{\phi(\alpha)}$$

Suppose $\mathbb{E}[X] = 0$, $\text{Var}(X) = \sigma^2 < \infty$, and $\alpha > 0$.

(a) Use the extended version of Markov's Inequality stated above with $\phi(x) = (x+c)^2$, where c is some positive constant, to show that:

$$\mathbb{P}[X \geq \alpha] \leq \frac{\sigma^2 + c^2}{(\alpha + c)^2}$$

(b) Note that the above bound applies for all positive c, so we can choose a value of c to minimize the expression, yielding the best possible bound. Find the value for c which will minimize the RHS expression (you may assume that the expression has a unique minimum).

We can plug in the minimizing value of c you found in part (b) to prove the following bound:

$$\mathbb{P}[X \geq \alpha] \leq \frac{\sigma^2}{\alpha^2 + \sigma^2}.$$

This bound is also known as Cantelli's inequality.

(c) Recall that Chebyshev's inequality provides a two-sided bound. That is, it provides a bound on $\mathbb{P}[|X - \mathbb{E}[X]| \geq \alpha] = \mathbb{P}[X \geq \mathbb{E}[X] + \alpha] + \mathbb{P}[X \leq \mathbb{E}[X] - \alpha]$. If we only wanted to bound the probability of one of the tails, e.g. if we wanted to bound $\mathbb{P}[X \geq \mathbb{E}[X] + \alpha]$, it is tempting to just divide the bound we get from Chebyshev's by two.

   (i) Why is this not always correct in general?

   (ii) Provide an example of a random variable X (does not have to be zero-mean) and a constant α such that using this method (dividing by two to bound one tail) is not correct, that is, $\mathbb{P}[X \geq \mathbb{E}[X] + \alpha] > \frac{\mathrm{Var}(X)}{2\alpha^2}$ or $\mathbb{P}[X \leq \mathbb{E}[X] - \alpha] > \frac{\mathrm{Var}(X)}{2\alpha^2}$.

Now we see the use of the bound proven in part (b) - it allows us to bound just one tail while still taking variance into account, and does not require us to assume any property of the random variable. Note that the bound is also always guaranteed to be less than 1 (and therefore at least somewhat useful), unlike Markov's and Chebyshev's inequality!

(d) Let's try out our new bound on a simple example. Suppose X is a positively-valued random variable with $\mathbb{E}[X] = 3$ and $\mathrm{Var}(X) = 2$.

   (i) What bound would Markov's inequality give for $\mathbb{P}[X \geq 5]$?

   (ii) What bound would Chebyshev's inequality give for $\mathbb{P}[X \geq 5]$?

   (iii) What bound would Cantelli's Inequality give for $\mathbb{P}[X \geq 5]$? (*Note*: Recall that Cantelli's Inequality only applies for zero-mean random variables.)

4 Markov Chains: Prove/Disprove

Prove or disprove the following statements.

(a) There exists an irreducible, finite Markov chain for which there exist initial distributions that converge to different distributions.

(b) There exists an irreducible, aperiodic, finite Markov chain for which $\mathbb{P}(X_{n+1} = j \mid X_n = i) = 1$ or 0 for all i, j.

(c) There exists an irreducible, non-aperiodic Markov chain for which $\mathbb{P}(X_{n+1} = j \mid X_n = i) \neq 1$ for all i, j.

(d) For an irreducible, non-aperiodic Markov chain, any initial distribution not equal to the invariant distribution does not converge to any distribution.

5 Playing Blackjack

You are playing a game of Blackjack where you start with \$100. You are a particularly risk-loving player who does not believe in leaving the table until you either make \$400, or lose all your money. At each turn you either win \$100 with probability p, or you lose \$100 with probability $1-p$.

(a) Formulate this problem as a Markov chain; i.e. define your state space, transition probabilities, and determine your starting state.

(b) Compute the probability that you end the game with \$400.

6 Reflecting Random Walk

Alice starts at vertex 0 and wishes to get to vertex n. When she is at vertex 0 she has a probability of 1 of transitioning to vertex 1. For any other vertex i, there is a probability of $1/2$ of transitioning to $i+1$ and a probability of $1/2$ of transitioning to $i-1$.

(a) What is the expected number of steps Alice takes to reach vertex n? Write down the hitting-time equations, but do not solve them yet.

(b) Solve the hitting-time equations. [*Hint*: Let R_i denote the expected number of steps to reach vertex n starting from vertex i. As a suggestion, try writing R_0 in terms of R_1; then, use this to express R_1 in terms of R_2; and then use this to express R_2 in terms of R_3, and so on. See if you can notice a pattern.]

7 Continuous Intro

(a) Is

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

a valid density function? Why or why not? Is it a valid CDF? Why or why not?

(b) Calculate $\mathbb{E}[X]$ and $\text{Var}(X)$ for X with the density function

$$f(x) = \begin{cases} \frac{1}{\ell}, & 0 \le x \le \ell, \\ 0, & \text{otherwise.} \end{cases}$$

(c) Suppose X and Y are independent and have densities

$$f_X(x) = \begin{cases} 2x, & 0 \le x \le 1, \\ 0, & \text{otherwise,} \end{cases}$$

$$f_Y(y) = \begin{cases} 1, & 0 \le y \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

What is their joint distribution? (Hint: for this part and the next, we can use independence in much the same way that we did in discrete probability)

(d) Calculate $\mathbb{E}[XY]$ for the above X and Y.