CS 70 Spring 2022 — Discrete Mathematics and Probability Theory — Satish Rao and Koushik Sen — HW 12

Due: Saturday 4/16, 4:00 PM
Grace period until Saturday 4/16, 6:00 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 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 $\alpha$ 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.)

# 2 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.

# 3 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%?

# 4 Estimating $\pi$

In this problem, we discuss some interesting ways that you could probabilistically estimate $\pi$, and see how good these techniques are at estimating $\pi$.

**Technique 1:** Buffon's needle is a method that can be used to estimate the value of $\pi$. There is a table with infinitely many parallel lines spaced a distance 1 apart, and a needle of length 1. It turns out that if the needle is dropped uniformly at random onto the table, the probability of the needle intersecting a line is $\frac{2}{\pi}$. We will see a proof of this later on in the notes.

**Technique 2:** Consider a square dartboard, and a circular target drawn inscribed in the square dartboard. A dart is thrown uniformly at random in the square. The probability the dart lies in the circle is $\frac{\pi}{4}$.

**Technique 3:** Pick two integers $x$ and $y$ independently and uniformly at random from 1 to $M$, inclusive. Let $p_M$ be the probability that $x$ and $y$ are relatively prime. Then

$$\lim_{M\to\infty} p_M = \frac{6}{\pi^2}.$$

Let $p_1 = \frac{2}{\pi}, p_2 = \frac{\pi}{4}$, and $p_3 = \frac{6}{\pi^2}$ be the probabilities of the desired events of **Technique 1**, **Technique 2**, and **Technique 3**, respectively. For each technique, we apply each technique $N$ times, then compute the proportion of the times each technique occurred, getting estimates $\hat{p}_1, \hat{p}_2$, and $\hat{p}_3$,, respectively.

(a) For each $\hat{p}_i$, compute an expression $X_i$ in terms of $\hat{p}_i$ that would be an estimate of $\pi$.

(b) Using Chebyshev's Inequality, compute the minimum value of $N$ such that $X_2$ is within $\varepsilon$ of $\pi$ with $1-\delta$ confidence. Your answer should be in terms of $\varepsilon$ and $\delta$.

For $X_1$ and $X_3$, computing the minimum value of $N$ will be more tricky, as the expressions for $X_1$ and $X_3$ are not as nice as $X_2$.

(c) For $i = 1$ and 3, compute a constant $c_i$ such that

$$|X_i - \pi| < \varepsilon \implies |\hat{p}_i - p_i| < c_i\varepsilon + o(\varepsilon^2),$$

where the $o(\varepsilon^2)$ represents terms containing powers of $\varepsilon$ that are 2 or higher (i.e. $\varepsilon^2, \varepsilon^3$, etc.).

(Hint: You may find the following Taylor series helpful: For $x$ close to 0,

$$\frac{1}{a-x} = \frac{1}{a} + o(x)$$
$$\frac{1}{(a-x)^2} = \frac{1}{a^2} + o(x).$$

The $o(x)$ represents terms that have $x^1$ powers or higher. )

In this problem, we assume $\varepsilon$ is close enough to 0 such that $o(\varepsilon^2)$ is 0. In other words,

$$\mathbb{P}\left[|\hat{p}_i - p_i| < c_i\varepsilon + o(\varepsilon^2)\right] = \mathbb{P}\left[|\hat{p}_i - p_i| < c_i\varepsilon\right].$$

Combining with part (c) then gives

$$\mathbb{P}\left[|X_i - \pi| < \varepsilon\right] \le \mathbb{P}\left[|\hat{p}_i - p_i| < c_i\varepsilon\right].$$

(d) For $i = 1$ and 3, use Chebyshev's Inequality and the above work to compute the minimum value of $N$ such that $X_i$ is within $\varepsilon$ of $\pi$ with $1-\delta$ confidence. Your answer should be in terms of $\varepsilon$ and $\delta$.

(e) Which technique required the lowest value for $N$? Which technique required the highest?

# 5 Balls in Bins Estimation

We throw $n > 0$ balls into $m \geq 2$ bins. Let $X$ and $Y$ represent the number of balls that land in bin 1 and 2 respectively.

(a) Calculate $\mathbb{E}[Y \mid X]$. [*Hint*: Your intuition may be more useful than formal calculations.]

(b) What is $L[Y \mid X]$ (where $L[Y \mid X]$ is the best linear estimator of $Y$ given $X$)? [*Hint*: Your justification should be no more than two or three sentences, no calculations necessary! Think carefully about the meaning of the conditional expectation.]

(c) Unfortunately, your friend is not convinced by your answer to the previous part. Compute $\mathbb{E}[X]$ and $\mathbb{E}[Y]$.

(d) Compute $\text{Var}(X)$.

(e) Compute $\text{cov}(X,Y)$.

(f) Compute $L[Y \mid X]$ using the formula. Ensure that your answer is the same as your answer to part (b).

# 6 Gambling Woes

Forest proposes a gambling game to you (uh oh!). Every day, you flip two independent fair coins. If both of the coins come up heads, then your fortune triples on that day. If one coin comes up heads and the other coin comes up tails, then your fortune is cut in half. If both of the coins comes up tails, then game over: you lose all of your money! Forest claims that you can get rich quickly with this scheme, but you decide to calculate some probabilities first.

(a) Let $M_0$ denote your money at the start of the game, and let $M_n$ denote the amount of money you have at the end of the $n$th day. Compute $\mathbb{E}[M_{n+1} \mid M_n]$.

(b) Use the law of iterated expectation to calculate $\mathbb{E}[M_{n+1}]$ in terms of $\mathbb{E}[M_n]$. Solve your recurrence to obtain an expression for $\mathbb{E}[M_{n+1}]$. Do you think this is a fair game?

(c) Calculate $\mathbb{P}(M_n > 0)$. What is the behavior as $n \to \infty$? Would you still play this game?

# 7 Iterated Expectation

In this question, we will try to achieve more familiarity with the law of iterated expectation.

1. You lost your phone charger! It will take $D$ days for the new phone charger you ordered to arrive at your house (here, $D$ is a random variable). Suppose that on day $i$, the amount of battery you lose is $B_i$, where $\mathbb{E}[B_i] = \beta$. Let $B = \sum_{i=1}^{D} B_i$ be the total amount of battery drained between now and when your new phone charger arrives. Apply the law of iterated

expectation to show that $\mathbb{E}[B] = \beta\,\mathbb{E}[D]$. (Here, the law of iterated expectation has a very clear interpretation: the amount of battery you expect to drain is the average number of days it takes for your phone charger to arrive, multiplied by the average amount of battery drained per day.)

2. Consider now the setting of independent Bernoulli trials, each with probability of success $p$. Let $S_i$ be the number of successes in the first $i$ trials. Compute $\mathbb{E}[S_m \mid S_n]$. (You will need to consider three cases based on whether $m > n$, $m = n$, or $m < n$. Try using your intuition rather than proceeding by calculations.)

# 8 In the Moments

Suppose a random variable $X$ satisfies the following conditions:

- $\mathbb{P}[X < 0] = \mathbb{P}[X > 1] = 0$
- $\mathbb{E}[X] \geq 2c$
- $\mathbb{E}\left[X^2\right] \leq \frac{19c^2}{4}$

Prove that

$$\mathbb{P}[X \geq c] \geq \max\left(c, \frac{1}{4}\right)$$

(Hint: Break the inequality into cases; conditional expectation may be helpful in one case)