Dis 7A: Bounds on Probability

Monday, 3 August 2020 4:08 PM

Let's say you have a Q that goes:
… Show that Pr(A) <= 1 / k

What tools do we have to show something like this?

Union Bound

Retrieval Practice + Derivation of the fundamentals

In **Markov's Inequality**, we have a random variable X, positive constant c, and the statement:

$Pr(X > c) \leq \frac{E[X]}{c}$,

$Pr(X > c) \geq 0.5$

i) What kind of mathematical object is X > c?
ii) Suppose X > c ⊆ A and the two both belong to the same sample space. What is true about Pr(A)?
iii) What's missing from this description of Markov's inequality?

has to be non-negative

[Skip]
Suppose X is **not** a non-negative RV. Give an example of a value for X(w) where the proof below breaks:

Alternative proof of Theorem 17.1. Since X is a nonnegative random variable and $c > 0$, we have, for all $\omega \in \Omega$,

$$X(\omega) \geq c\,I\{X(\omega) \geq c\}, \qquad (1)$$

since the right hand side is 0 if $X(\omega) < c$ and is c if $X(\omega) \geq c$. Multiplying both sides by $\mathbb{P}[\omega]$ and summing over $\omega \in \Omega$ gives

$$\mathbb{E}[X] \geq c\,\mathbb{E}[I\{X \geq c\}] = c\,\mathbb{P}[X \geq c],$$

where the first inequality follows from (1) and the fact that $\mathbb{P}[\omega] \geq 0$ for all $\omega \in \Omega$, while the last equality follows from the fact that $I\{X \geq c\}$ is an indicator random variable. □

[Skip]
Practice:
Take $Y = (X - E[X])^2$, where X is any RV. Bound the value of $Pr(Y > c^2)$.

What did you just do?

How you use this to 'derive' something you've seen (or not)?

Practice with bounding probabilities of random variables

[2CQ]
$\boxed{|X – c| = 1}$ *is equivalent to:*
A) X – c = 1 and - (X – c) = 1
B) X – c = 1 or - (X – c) = 1

$$\Pr(|X - E(X)| > c) \le \frac{Var(X)}{c^2}$$

X is a RV. E[X] = 1, Var[X] = 1.
Bound $\boxed{Pr(X >= 1.2)}$ *via Chebyshev's Inequality.*
Hint 1: if A <= B and B <= C, then A <= C. What might be have for A, B, C?
Hint 2: Proofs via Chebyshev's IE usually take the form:
Pr(X >= x) <= Pr(|X – E[X]| >= c) <= Var[X] / c^2 (your goal is to figure out the 'correct' 'c' value.

$$\Pr(X \ge 1.2) \le \Pr(|X-1| \ge 0.2) \le \frac{Var(X)}{0.2^2} \le \frac{1}{(0.2)^2}$$

$$X \ge 1.2 \subseteq (X \ge 1.2 \cup x \le 0.8)$$

Questions

1 Probabilistic Bounds

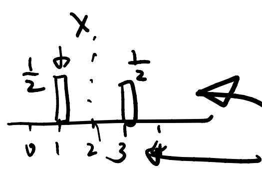
A random variable X has variance $\mathrm{Var}(X) = 9$ and expectation $\mathbb{E}[X] = 2$. Furthermore, the value of X is never greater than 10. Given this information, provide either a proof or a counterexample for the following statements.

(a) $\mathbb{E}[X^2] = 13$.

$$E(X^2) = Var(X) + E^2(X) = 9 + 4 = 13$$

(b) $\mathbb{P}[X = 2] > 0$.

$$\Pr(X = 5) = \frac{1}{2}$$
$$\Pr(X = -1) = \frac{1}{2}$$



$$E(X) = \frac{1}{2} \times 5 + \frac{1}{2} \times -1 = 2 \checkmark$$
$$E(X^2) = \frac{1}{2} \times 25 + \frac{1}{2} \times 1 = 13$$

$$X = \begin{cases} a & \text{wp } p \\ b & \text{wp } 1-p \end{cases}$$

(c) $\mathbb{P}[X \ge 2] = \mathbb{P}[X \le 2]$.

$\boxed{X < 10}$, $E(X) = 2$, $Var(X) = 9$, $E(X^2) = 13$

$$X = \begin{cases} a & \text{wp } p \\ b & \text{wp } (1-p) \end{cases}$$

$$X = \begin{cases} 5 & \text{wp } \frac{1}{2} \\ -1 & \text{wp } \frac{1}{2} \end{cases}$$

satisfies conditions

$E(X) =$
$E(X^2) =$

$E[X] = 2$, $Var(X) = 9$, $X \le 10$ → $0 \le 10 - X$ $= Y$

(d) $\mathbb{P}[X \le 1] \le 8/9$.

$E[X]=2$, $Var(X)=9$ $\quad X \le 10 \rightarrow 0 \le -X+10 = Y$

(d) $\mathbb{P}[X \le 1] \le 8/9$.

$P(X=1) = 0.95$

$Pr(Y \ge c) \le \frac{E(Y)}{c}$

$E(X) = 0.95 \times 1 + \frac{0.05 \times 10}{0.5}$

$= \cancel{1.4}\ 1.45 < 2$

$-X = Y-10$

$X = 10 - Y \le 1$

$\boxed{P(X \le 1)} = Pr(\boxed{9 \le Y})$

$= Pr(Y \ge 9) \le \frac{E(Y)}{9} = \frac{8}{9}$

$9 \le Y$

$E[Y] = E(-X)+10$

$= -2+10 = 8$

(e) $\mathbb{P}[X \ge 6] \le 9/16$.

$P(X \ge 6) \le Pr(|X-2| \ge 4) \le \frac{Var(X)}{4^2}$

$\le \frac{9}{16}$

2 Easy A's

A friend tells you about a course called "Laziness in Modern Society" that requires almost no work. You hope to take this course next semester to give yourself a well-deserved break after working hard in CS 70. At the first lecture, the professor announces that grades will depend only on two homework assignments. Homework 1 will consist of three questions, each worth 10 points, and Homework 2 will consist of four questions, also each worth 10 points. He will give an A to any student who gets at least 60 of the possible 70 points.

However, speaking with the professor in office hours you hear some very disturbing news. He tells you that, in the spirit of the class, the GSIs are very lazy, and to save time the grading will be done as follows. For each student's Homework 1, the GSIs will choose an integer randomly from a distribution with mean $\mu = 5$ and variance $\sigma^2 = 1$. They'll mark each of the three questions with that score. To grade Homework 2, they'll again choose a random number from the same distribution, independently of the first number, and will mark all four questions with that score.

If you take the class, what will the mean and variance of your total class score be? Use Chebyshev's inequality to conclude that you have less than a 5% chance of getting an A when the grades are randomly chosen this way.

T

See if you can figure this out in a systematic manner. Essentially, you want to get things into the form of Q1, where you had all the math you need. Then, proceed like you'd proceed with Q1 e)

$T = 3X + 4Y$, $\quad E[X]=5$, $Var(X)=1$

$E[Y]=5$, $Var(Y)=1$

$E(T) = 15+20 = 35$

$Var(T) = 9+16 = 25$

if X, Y are indep

$Var(aX+bY) = a^2 Var(X) + b^2 Var(Y)$

$Pr(T \ge 60) \overset{?}{<} 0.05$

$Pr(T \geq 60) \stackrel{?}{<} 0.05$

$= Pr(T - 35 \geq 60 - 35)$

$\leq Pr(|T-35| \geq 25) \leq \frac{Var(T)}{25^2} = \frac{25}{25^2}$

$= \frac{1}{25} = 0.04$

< 0.05

Intuitive meaning of the Law of Large Numbers?
The more times you repeat an experiment and take an average, the more likely you'll get closer to the real average.

I find it easier to think of LLN with a concrete yet simple example.

Suppose you roll a fair die a bunch (k) of times. Let's consider the average value rolled in these k flips. As k increases, the **actual average value of the k rolls** is more likely to be:

I) further from 3.5
II) closer to 3.5

[Notebook Demo – share laptop screen]

In formal probability terms:

Theorem 17.4 (Law of Large Numbers). *Let $X_1, X_2, \ldots,$ be a sequence of i.i.d. (independent and identically distributed) random variables with common finite expectation $\mathbb{E}[X_i] = \mu$ for all i. Then, their partial sums $S_n = X_1 + X_2 + \cdots + X_n$ satisfy*

$$\mathbb{P}\left[\left|\tfrac{1}{n}S_n - \mu\right| < \varepsilon\right] \to 1 \qquad \text{as } n \to \infty,$$

for every $\varepsilon > 0$, however small.

3 Working with the Law of Large Numbers

(a) A fair coin is tossed multiple times and you win a prize if there are more than 60% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(b) A fair coin is tossed multiple times and you win a prize if there are more than 40% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(c) A fair coin is tossed multiple times and you win a prize if there are between 40% and 60% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(d) A fair coin is tossed multiple times and you win a prize if there are exactly 50% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.