1 Probabilistic Bounds

A random variable X has variance $\text{Var}(X) = 9$ 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) $\mathbb{E}[X^2] = 13$.

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

(c) $\mathbb{P}[X \geq 2] = \mathbb{P}[X \leq 2]$.

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

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

Solution:

(a) TRUE. Since $9 = \text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - 2^2$, we have $\mathbb{E}[X^2] = 9 + 4 = 13$.

(b) FALSE. It is not necessary for a random variable to be able to take on its mean as a value. Construct a random variable X that satisfies the conditions in the question but does not take on the value 2. A simple example would be a random variable that takes on 2 values, where $\mathbb{P}[X = a] = \mathbb{P}[X = b] = 1/2$, and $a \neq b$. The expectation must be 2, so we have $a/2 + b/2 = 2$. The variance is 9, so $\mathbb{E}[X^2] = 13$ (from Part (a) and $a^2/2 + b^2/2 = 13$. Solving for a and b, we get $\mathbb{P}[X = -1] = \mathbb{P}[X = 5] = 1/2$ as a counterexample.

(c) FALSE. The median of a random variable is not necessarily the mean, unless it is symmetric. Construct a random variable X that satisfies the conditions in the question but does not have an equal chance of being less than or greater than 2. A simple example would be a random variable that takes on 2 values, where $\mathbb{P}[X = a] = p, \mathbb{P}[X = b] = 1 - p$. Here, we use the same approach as part (b) except with a generic p, since we want $p \neq 1/2$. The expectation must be 2, so we have $pa + (1 - p)b = 2$. The variance is 9, so $\mathbb{E}[X^2] = 13$ and $pa^2 + (1 - p)b^2 = 13$. Solving for a and b, we find the relation $b = 2 \pm 3/\sqrt{x}$, where $x = (1 - p)/p$. Then, we can find an example by plugging in values for x so that $a, b \leq 10$ and $p \neq 1/2$. One such counterexample is $\mathbb{P}[X = -7] = 1/10, \mathbb{P}[X = 3] = 9/10$.

(d) TRUE. Let $Y = 10 - X$. Since X is never exceeds 10, Y is a non-negative random variable. By Markov's inequality,

$$\mathbb{P}[10 - X \geq a] = \mathbb{P}[Y \geq a] \leq \frac{\mathbb{E}[Y]}{a} = \frac{\mathbb{E}[10 - X]}{a} = \frac{8}{a}.$$

Setting $a = 9$, we get $\mathbb{P}[X \leq 1] = \mathbb{P}[10 - X \geq 9] \leq 8/9$.

(e) TRUE. Chebyshev's inequality says $\mathbb{P}[|X - \mathbb{E}[X]| \geq a] \leq \text{Var}(X)/a^2$. If we set $a = 4$, we have

$$\mathbb{P}[|X - 2| \geq 4] \leq \frac{9}{16}.$$

Now we observe that $\mathbb{P}[X \geq 6] \leq \mathbb{P}[|X - 2| \geq 4]$, because the event $X \geq 6$ is a subset of the event $|X - 2| \geq 4$.

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.

Solution:

Let X be the total number of points you receive in the class. Then $X = X_1 + X_2$ where X_1 is the number points received on Homework 1 and X_2 is the number of points received on Homework 2. Your Homework 1 score is generated as $X_1 = 3Y_1$, where the r.v. Y_1 represents the integer that the GSI chose when grading it. Similarly, $X_2 = 4Y_2$. The problem statement tells us that Y_1 and Y_2 are independent, both with mean 5 and variance 1, so $\mathbb{E}[Y_1] = \mathbb{E}[Y_2] = 5$ and $\text{Var}(Y_1) = \text{Var}(Y_2) = 1$. Thus,

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] = 3\mathbb{E}[Y_1] + 4\mathbb{E}[Y_2] = 35,$$
$$\text{Var}(X) = \text{Var}(3Y_1) + \text{Var}(4Y_2) = 9\text{Var}(Y_1) + 16\text{Var}(Y_2) = 25.$$

Using Chebyshev's Inequality, we get

$$\mathbb{P}[X \geq 60] \leq \mathbb{P}[|X - 35| \geq 25] \leq \frac{\text{Var}(X)}{25^2} = \frac{1}{25}.$$

Unfortunately, any student will have at most a 4% chance of getting an A.

Note that although we calculated a bound for $\mathbb{P}[|X - 35| \geq 25]$, which is the probability that you will get 60 or above or 10 or below, we cannot divide by 2 to refine our bound unless the distribution is symmetric about its mean. In this case, the distribution is not symmetric.

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.

Solution:

(a) 10 tosses. By LLN, the sample mean should have higher probability to be close to the population mean as n increases. Therefore the average proportion of coins that are heads should be closer to 0.50, and has a lower chance of being greater than 0.60 if there are 100 tosses (compared with 10 tosses).

(b) 100 tosses. Again, by LLN, the sample mean should have higher probability to be close to the population mean as n increases. Therefore the average proportion of coins that are heads should be closer to 0.50, and has a lower chance of being smaller than 0.40 if there are 100 tosses. A lower chance of being smaller than 0.40 is the desired result.

(c) 100 tosses. Again, by LLN, the average proportion of coins that are heads should be closer to 0.50, and has a lower chance of being both smaller than 0.40 if there are 100 tosses. Similarly, there is a lower chance of being larger than 0.60 if there are 100 tosses. Lower chances of both of these events is desired if we want the fraction of heads to be between 0.4 and 0.6.

(d) 10 tosses. Compare the probability of getting equal number of heads and tails between $2n$ and

$2n+2$ tosses.

$$\mathbb{P}[n \text{ heads in } 2n \text{ tosses}] = \binom{2n}{n}\frac{1}{2^{2n}}$$

$$\mathbb{P}[n+1 \text{ heads in } 2n+2 \text{ tosses}] = \binom{2n+2}{n+1}\frac{1}{2^{2n+2}} = \frac{(2n+2)!}{(n+1)!(n+1)!} \cdot \frac{1}{2^{2n+2}}$$

$$= \frac{(2n+2)(2n+1)2n!}{(n+1)(n+1)n!n!} \cdot \frac{1}{2^{2n+2}}$$

$$= \frac{2n+2}{n+1} \cdot \frac{2n+1}{n+1}\binom{2n}{n} \cdot \frac{1}{2^{2n+2}} < \left(\frac{2n+2}{n+1}\right)^2\binom{2n}{n} \cdot \frac{1}{2^{2n+2}}$$

$$= 4\binom{2n}{n} \cdot \frac{1}{2^{2n+2}} = \binom{2n}{n}\frac{1}{2^{2n}} = \mathbb{P}[n \text{ heads in } 2n \text{ tosses}]$$

As we increment n, the probability will always decrease. Therefore, the larger n is, the less probability we'll get exactly 50% heads. $\square$

Note: By Stirling's approximation, $\binom{2n}{n}2^{-2n}$ is roughly $(\pi n)^{-1/2}$ for large n.

See https://github.com/dingyiming0427/CS70-demo/ for a code demo.