Section 01: Solutions

1. Expectation

(a) You've just started a new exercise regimen. You start on the 2nd floor of CSE1, and then make a random choice:

- With probability p_1 you run up 2 flights of stairs.
- With probability p_2 you run up 1 flight of stairs.
- With probability p_3 you walk down 1 flight of stairs.

Where $p_1 + p_2 + p_3 = 1$.

You will do two iterations of your exercise scheme (with each draw being independent). Let X be the floor you're on at the end of your exercise routine. Recall you start on floor 2.

(i) Let Y be the expected difference between your ending floor and your starting floor in one iteration. What is $\mathbb{E}[Y]$ (in terms of p_1, p_2, p_3)?

Solution:

Recall for a random variable $X, \mathbb{E}[X] = \sum_i x_i \cdot p_i$.
So $\mathbb{E}[Y] = 2 \cdot p_1 + 1 \cdot p_2 + (-1) \cdot p_3$

(ii) What is $\mathbb{E}[X]$ (use your answer from the previous part)

Solution:

Since we start at floor 2, we can take 2 and add the difference ($\mathbb{E}[Y]$) twice to get our expected floor at the end of the routine.
$\mathbb{E}[X] = 2 + \mathbb{E}[Y] + \mathbb{E}[Y] = 2 + 2\mathbb{E}[Y]$

(iii) You change your scheme: instead of doing two independent iterations, you decide the second iteration of your regimen will just use the same random choice as your first (in particular they are no longer independent!). Does $\mathbb{E}[X]$ change?

Solution:

No! We can say using the same choice as the first will effectively double Y, thus by linearity of expectation, $\mathbb{E}[X] = 2 + \mathbb{E}[2Y] = 2 + 2\mathbb{E}[Y]$

(b) Let $X_1, X_2, \ldots, X_n$ be independent random variables drawn uniformly from $(0, 1)$. Define the random variable $Y = \max\{X_1, \ldots, X_n\}$. What is $\mathbb{E}[Y]$?

Solution:

Recall the definition of $\mathbb{E}[Y]$ is $\int_{-\infty}^{\infty} y f(y) dy$, where $f(y)$ is the pdf for Y, so we need the PDF for Y. It's easier here to calculate the CDF and then take its derivative to find the PDF. Let $F(y)$ be the CDF for Y. Recall that a CDF asks the question "what is the probability that the value of Y is at most y?" For our Y that is asking, "what is the probability that all the X_i are less than y, but we can calculate that!

For $y \in (0, 1)$, that's just the chance that n independent Unif(0,1) random variables evaluated to at most y. The probability of that is just y^n. To turn that into a PDF, we just need to take the derivative with respect to y, so we get a CDF of:

$$f(y) = \begin{cases} 0 & \text{if } y \le 0 \text{ or } y \ge 1 \\ ny^{n-1} & \text{otherwise} \end{cases}$$

We can now evaluate the integral:

$$\begin{aligned} \mathbb{E}[Y] &= \int_{-\infty}^{\infty} y \cdot f(y) dy \\ &= \int_0^1 y \cdot ny^{n-1} dy \\ &= n \cdot \frac{y^{n+1}}{n+1}\Big|_0^1 \\ &= \frac{n}{n+1} \end{aligned}$$

Let's sanity check that solution – Y is bounded between 0 and 1, and so is the formula we got. As n increases, we're taking the max of more things, so we should expect $\mathbb{E}[Y]$ to increase, and it does. For $n = 1$, we're just asking for the expectation of one Unif(0,1) variable, so we should get $1/2$ (and we do). Seems sensible!

2. Linearity and Independence

Suppose we have two random variables X and Y, such that $\mathbb{E}[X] = \mathbb{E}[Y] = 2$. For each of the following quantities either:

- State the value of the quantity if we have enough information to find it, or
- Give examples of two different values the quantity could take if we do not.

(a) $\mathbb{E}[X + Y]$
Solution:

By linearity of expectation, $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] = 4$

(b) $\mathbb{E}[XY]$
Solution:

We cannot compute $\mathbb{E}[XY]$.
Let's say $P(X = 0) = P(X = 4) = 1/2$, $P(Y = 0|X = 4) = 1$, $P(Y = 4|X = 0) = 1$.
That is, we pick X to be 0 or 4 randomly and set Y to be 4 or 0, respectively depending on X.
Then $\mathbb{E}[XY] = 0$ since one of them will always take on a value of 0.

Alternatively, if $P(X = 2) = P(Y = 2) = 1$, then $E[XY] = 4$.

(c) $\mathbb{E}[X^2]$
Solution:

We cannot compute $\mathbb{E}[X^2]$.
Recall Var$(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

Since we don't know Var(X), it can take on any value $\mathbb{E}[X^2] = \text{Var}(X) + \mathbb{E}[X]^2$.
For example, $\mathbb{E}[X^2]$ could equal 4 (for $P(X = 2) = 1$) or 5 (for $\mathcal{N}(2, 1)$)

(d) $\mathbb{E}[X]^2$
Solution:

$\mathbb{E}[X]^2 = 2^2 = 4$

Suppose we additionally know that X and Y are independent. Do any of the answers change?
Solution:

Yes, if X and Y are independent, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] = 4$.

3. Variance and Concentration

Anna and Kevin want to see if the students in the course like probability theory. You (because you're so friendly) know that 200 out of the 250 students in the course say they like probability theory, but Anna and Kevin don't believe you. They decide to use the following process to estimate the number of people who like probability theory:

- Choose a student uniformly at random (and independent from any previous choices).
- Record $X_i = \begin{cases} 1 & \text{if the student likes probability} \\ 0 & \text{otherwise} \end{cases}$

They will choose 30 such students this way, and they define $X = \frac{\sum_{i=1}^{30} X_i}{30}$, the average of the X_i.

(a) What is $\mathbb{E}[X_1]$?

Solution:

$\frac{200}{250} \cdot 1 + \frac{50}{250} \cdot 0 = \frac{4}{5}$

(b) What is Var(X_1)? Hint: $p(1 - p)$ is the variance of a Bernoulli random variable with probability of success p.

Solution:

Following the hint, it's $\left(\frac{4}{5} \cdot \frac{1}{5}\right)$.

(c) What is $\mathbb{E}[X]$?

Solution:

$$\mathbb{E}\left[\frac{1}{30}\sum_{i=1}^{30} X_i\right] = \frac{1}{30} \cdot \sum_{i=1}^{30} \mathbb{E}\left[X_i\right] = \frac{1}{30} \cdot 30 \cdot \frac{4}{5} = \frac{4}{5}$$

(d) What is Var(X)?

Solution:

$$
\begin{aligned}
\text{Var}(X) &= \text{Var}\left(\frac{1}{30}\sum_{i=1}^{30} X_i\right) \\
&= \frac{1}{30^2}\text{Var}\left(\sum_{i=1}^{30} X_i\right) \\
&= \frac{1}{30^2} 30 \cdot \text{Var}(X_i) \\
&= \frac{4}{30 \cdot 25}
\end{aligned}
$$

We're using the independence of X_i to use that the variance of the sum equals the sum of the variances.

Kevin and Anna are worried that less than half the course likes probability theory. They will stop being worried if $X \geq 0.5$. Use Chebyshev's inequality to give a lower bound on the probability that they stop worrying.

Theorem 1 (Chebyshev's Inequality)**.** *If X is a random variable with finite mean μ and finite variance σ^2, then for any real number $k > 0$:*

$$
Pr\left[|X - \mu| \geq k\sigma\right] \leq \frac{1}{k^2}
$$

Solution:

We're trying to find an upper bound on $X < 0.5$. To apply Chebyshev, we need to rephrase that event in terms of something like $|X - \mu| \geq k\sigma$. This won't be an exact match, but we can still find an upper bound:

$\text{Pr}[X < 0.5] \leq \text{Pr}\left[\left|X - \frac{4}{5}\right| \geq \left(\frac{4}{5} - \frac{1}{2}\right)\right]$

Now we just need to find the value of k so that we can substitiute $k\sigma$ where we have $\frac{4}{5} - \frac{1}{2}$.

Plugging and chugging: $k = \left(\frac{4}{5} - \frac{1}{2}\right) \cdot \sqrt{\frac{30 \cdot 25}{4}} \approx 4.107$

Now applying Chebyshev we have:

$$
\begin{aligned}
\text{Pr}[X < 0.5] &\leq \text{Pr}\left[\left|X - \frac{4}{5}\right| \geq \left(\frac{4}{5} - \frac{1}{2}\right)\right] \\
&\leq \text{Pr}\left[|X - \mu| \geq 4.107\sigma\right] \\
&\leq 1/4.107^2 \\
&\leq 0.059
\end{aligned}
$$

So the chances that Anna and Kevin are worried is definitely less than 6%. Hoeffding's Inequality (which we used in lecture 1) applies to this problem as well, and would give a tighter bound. You could also observe that the number of people who say "yes" is a binomial random variable and calculate the exact probability that way, but "concentration inequalities" (like Chebyshev and Hoeffding) are easier to use as n changes and gets much larger.