CS 206: Practice problems - Set III
(Solutions and answers to problems)
Summer 2018
(Based on lectures 10-13)

Problem 1

We repeatedly throw two fair dice until two sixes are rolled. What is the expected number of throws required? If $X$ is the number of throws needed what is $Var(X)$?

Solution sketch: This can be modelled as a geometric distribution with parameter $p = 1/36$. The expected value is $1/p = 36$, and the variance is

$$\frac{1}{p^2} - \frac{1}{p} = (36)^2 - 36.$$

Problem 2

Suppose we keep throwing a dice till we see a 6. Let $Y$ be the total number of throws needed, and let $X$ be the number of times we observe an even number.

1. What is $P(Y = 5)$?


3. What is $E[X | Y = 8]$? Solution sketch: Follow example at the end of Section 2 in Lecture 12.


5. Let $B$ be the event that the very first roll results in an odd number. Write $E[X | B]$ in terms of $E[X]$. Answer: $E[X | B] = E[X]$.


1
7. Compute \( E[X] \). \textbf{Hint:} Use the law of total expectation.

\textbf{Solution sketch:} Notice that \( A, B, \) and \( C \) form a partition of the sample space. Hence, using the law of total expectation,

\[
E[X] = P(A)E[X|A] + P(B)E[X|B] + P(C)E[X|C] = \frac{1}{6} \cdot 1 + \frac{3}{6} (E[X]) + \frac{2}{6} (1+E[X]).
\]

Solving this linear equation for \( E[X] \), gives \( E[X] = 3 \).

\textbf{Problem 3}

You have 10 pairs of socks (so 20 socks total), with each pair being a different color. You put them in the washing machine, but the washing machine eats 4 of the socks at random. What is the expected number of complete pairs you have left?

\textbf{Hint:} Define an indicator random variable for every pair.

\textbf{Solution sketch:} Let’s define 10 indicator random variables, \( X_1, \ldots, X_{10} \), one for each pair of socks, such that \( X_i \) is 1 if and only if both socks in the \( i^{th} \) pair survive. Then, the total number of pairs that survive is \( X = X_1 + \ldots + X_{10} \), and using linearity of expectation

\[
E[X] = \sum_{i=1}^{10} E[X_i].
\]

Also, \( E[X_i] = P(X_i = 1) = P(\text{Both socks in pair } i \text{ survive}) \). The latter can be shown to be equal to (check yourself)

\[
\left( \frac{18}{4} \right) \cdot \left( \frac{20}{4} \right),
\]

and so \( E[X] = \left( \frac{18}{4} \right) \cdot \left( \frac{20}{4} \right) \cdot 10 \).

\textbf{Problem 4}

\( n \) people, \( P_1, \ldots, P_n \) enter a room. Each pair \( \{P_i, P_j\} \) decides to shake hands independently with probability 0.5. A group of three people \( \{P_i, P_j, P_k\} \) is called a trio if every person in the group has shaken hands with the other two. Let \( T \) be the total number of trios. What is the range of the random variable \( T \)? Can you find \( E[T] \)?

\textbf{Hint:} Define an indicator random variable for every group of three people.

\textbf{Solution sketch:} The range of \( T \) is \( \{0, \ldots, \binom{n}{3}\} \) since the max number of trios that can be formed among \( n \) people is \( \binom{n}{3} \) (number of ways of forming groups of 3 from \( n \) people). Let us define \( \binom{n}{3} \) random variables, \( X_{\{i,j,k\}} \) where \( 1 \leq i < j < k \leq n \), such that \( X_{\{i,j,k\}} = 1 \) if \( P_i, P_j, P_k \) form a trio. Clearly, \( T = \sum_{1\leq i<j<k \leq n} X_{\{i,j,k\}} \), and using linearity of expectation

\[
E[T] = \sum_{1\leq i<j<k \leq n} E[X_{\{i,j,k\}}].
\]
Since $X_{i,j,k}$ is an indicator random variable

$$
\mathbb{E}[X_{i,j,k}] = P(X_{i,j,k} = 1)
= P(P_i, P_j, P_k \text{ form a trio}) = (0.5)^3
$$

(Convince yourself as to why this probability is $(0.5)^3$). This means that

$$
\mathbb{E}[T] = (0.5)^3 \binom{n}{3}.
$$

**Problem 5**

Answer True or False and provide explanations:

1. Let $X$ be the number rolled when a fair dice is thrown. Then $\text{Range}(X) = \{1, \ldots, 6\}$. **Answer: True**

2. If $X$ is an indicator random variable for an event $E$, then $\mathbb{E}[X] = P(E)$. **Answer: True**

3. A coin is tossed $n$ times (it’s fair). Let $X$ be number of heads seen and $Y$ the number of tails. Then $X$ and $Y$ are independent. **Answer: False (why?)**

4. If $X_1, \ldots, X_n$ are 5-wise independent then they are also pairwise independent. (Assume $n > 1000$) **Answer: This is true. Follows from the definition.**

5. If $X_1, \ldots, X_n$ are 5-wise independent then they are also 10-wise independent. (Assume $n > 1000$) **Answer: False, 10-wise independence is stronger than 5-wise independence (see definition).**

6. Suppose 21 jobs are assigned to 20 processors randomly. Let $X$ be the number of processors that are assigned at least two jobs. Then $P(X > 0) = 1$. **Answer: True. Pinhole principle: two jobs must always end up in the same processor.**

7. If $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ then $X$ and $Y$ must be independent. **Solution sketch:** We discussed this in class: it’s not true; only the reverse direction is true — independence implies that $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$. Let’s see an example where $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ yet $X$ and $Y$ are not independent:

Let $X$ and $Y$ be such that $\text{Range}(X) = \{-1, 0, 1\}$, and $P(X = -1) = \frac{1}{2}, P(X = 0) = \frac{1}{2}, P(X = 1) = \frac{1}{2}$, and let $Y = X^2$. Then clearly, $X$ and $Y$ are not independent (why?), yet $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ (check this).

8. $\text{Var}(-X) = \text{Var}(X)$. **Answer: True**

Problem 6

Let \(X_1, X_2, X_3, X_4\) be pairwise independent random variables such that for every \(1 \leq i \leq 4\), \(E[X_i] = 0.8\). What is \(E[X_1X_2 + X_1X_3 + X_1X_4 + X_2X_3 + X_2X_4 + X_3X_4]\)?

**Solution sketch:** Using linearity of expectation and the fact that for independent random variables \(E[XY] = E[X]E[Y]\), we have that

\[
E[X_1X_2 + X_1X_3 + X_1X_4 + X_2X_3 + X_2X_4 + X_3X_4] = E[X_1]E[X_2] + E[X_1]E[X_3] + \ldots + E[X_3]E[X_4].
\]

Now solve the problem using the fact that \(E[X_i] = 0.8\).

Problem 7

We want to choose a committee of 10 people from a group of 100 people, 60 men and 40 women. Suppose we do this randomly, i.e. we choose 10 random people from the group of 100 people so that all outcomes are equally likely. Let \(X\) be the total number of women chosen.

1. What is \(P(X = 6)\)? **Answer:** \(\binom{40}{6}\binom{60}{4}\binom{100}{10}\)

2. Let’s represent the women as \(W_1, \ldots, W_{40}\). Let \(X_i\) (for \(1 \leq i \leq 40\)) be the indicator random variable for the event of \(W_i\) being chosen for the committee. Then is it true that \(X_i\) and \(X_j\) (for \(i \neq j\)) are independent random variables. Why or why not? **Answer:** Clearly not, because, for example, given that the first 10 women are all chosen, i.e. the event \([X_1 = 1 \cap \ldots \cap X_{10} = 1]\), the probability that \(W_{11}\) is chosen is zero, i.e. the probability of the event \([X_{11} = 1]\) is zero, which is not equal to \(P(X_{11} = 1) = \frac{\binom{99}{9}}{\binom{100}{10}}\). If \(X_1, \ldots, X_{40}\) were mutually independent, we would expect that

\[
P(X_{11} = 1 | X_1 = 1 \cap X_2 = 1 \cap \ldots \cap X_{40} = 1) = P(X_{11} = 1).
\]

3. What is \(E[X]\)?

**Solution sketch:** \(E[X] = \sum_{i=1}^{40} E[X_i] = 40 \frac{\binom{99}{9}}{\binom{100}{10}}\)

Problem 8

At a grocery store there are \(X\) number of people that shop each day. Here \(X\) is a random variable with the following probability mass function: \(P(X = k) = \frac{\mu^ke^{-\mu}}{k!}\) for any integer \(k \geq 0\). Here \(\mu > 0\) is a fixed constant. Compute \(E[X]\).

**Hint:** Use the fact that \(e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots\) for any \(x\).

**Solution sketch:** Note that the range of \(X\) must obviously be the natural numbers. Using the definition of expectation:

\[
E[X] = \sum_{k \geq 0} k \cdot \frac{\mu^ke^{-\mu}}{k!} = \sum_{k \geq 1} k \cdot \frac{\mu^ke^{-\mu}}{k!} = \sum_{k \geq 1} \frac{\mu^ke^{-\mu}}{(k-1)!}
\]
\[= \mu e^{-\mu} \left( \sum_{k \geq 1} \frac{\mu^{(k-1)}}{(k-1)!} \right) = \mu e^{-\mu} (1 + \mu + \frac{\mu^2}{2!} + \frac{\mu^3}{3!} + \ldots) = \mu e^{-\mu} e^\mu = \mu \]

**Problem 9**

Alice is playing blackjack at a casino. She wins each round with probability \( \frac{1}{3} \), and loses the round with probability \( \frac{2}{3} \). For every round that she wins, she makes $9 dollars and for each round that she loses, she loses $3 dollars. Suppose Alice plays 100 rounds of blackjack and let \( X \) denote her net gain at the end of 100 rounds (net gain = money won - money lost).

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

**Solution sketch:** Let \( X_i \) be the net gain in round \( i \) (there are 100 rounds). Clearly, \( X = \sum_{i=1}^{100} X_i \), and so \( \mathbb{E}[X] = \sum_{i=1}^{100} \mathbb{E}[X_i] \). It’s easy to see that
\[
\mathbb{E}[X_i] = 9 \cdot \frac{1}{3} + (-3) \cdot \frac{2}{3} = 1,
\]
and so
\[
\mathbb{E}[X] = 100.
\]

(b) What is \( \text{Var}(X) \)?

**Solution sketch:** Note that \( \text{Var}(X) = \text{Var}(X_1 + \ldots + X_{100}) = \sum_{i=1}^{100} \text{Var}(X_i) \) since \( X_1, \ldots, X_{100} \) are independent (the rounds are independent). So we basically have to compute \( \text{Var}(X_i) \). Recall that \( \text{Var}(X_i) = \mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2 = \mathbb{E}[X_i^2] - 1 \). Let’s compute \( \mathbb{E}[X_i^2] \):
\[
\mathbb{E}[X_i] = 81 \cdot \frac{1}{3} + 9 \cdot \frac{2}{3} = 33,
\]
and so \( \text{Var}(X_i) = 33 - 1 = 32 \). Thus, \( \text{Var}(X) = 32 \times 100 = 3200 \).

(c) Given the best upper bound you can on the probability that Alice will end up owing money at the end of 100 rounds.

**Solution sketch:** We are interested in upper bounding \( P(X < 0) \), or in other words. Since \( \mathbb{E}[X] = 100 \), the event \( X < 0 \) is the same as the event \( \mathbb{E}[X] - X > 100 \) (why? convince yourself!), and so we are want to upper bound \( P(\mathbb{E}[X] - X > 100) \).

Chebyshev’s inequality says that
\[
P(|X - \mathbb{E}[X]| > 100) \leq \frac{\text{Var}(X)}{(100)^2} = \frac{3200}{10000} = 0.32,
\]
and since \( P(\mathbb{E}[X] - X > 100) \leq P(|X - \mathbb{E}[X]| > 100) \) (why?), we have that
\[
P(X < 0) = P(\mathbb{E}[X] - X > 100) \leq 0.32.
\]
This means that with probability (confidence) at least 68%, Alice will have a non-negative net gain!
Problem 10

Suppose that $X$ is a random variable such that $E[X] = 100$ and $Var(X) = 15$.

1. What is $E[X^2]$? **Solution sketch:** $Var(X) = E[X^2] - (E[X])^2$, and so

$$E[X^2] = 15 + 10000 = 10015.$$

2. What is $E[5X - 2X^2 + 1]$? **Solution hint:** Use linearity of expectation, and what you computed in the first part!