Quiz Section 4

Review

1) Probability Mass. For every random variable $X$, we have $\sum_x P(X = x) = \underline{1}$.

2) Expectation. $E(X) = \underline{}$.

3) Linearity of expectation. For any random variables $X_1, \ldots, X_n$, and real numbers $a_1, \ldots, a_n$,

$$E(a_1 X_1 + \cdots + a_n X_n) = \underline{}$$

4) Variance. $Var(X) = \underline{}$.

5) Independence. Two random variables $X$ and $Y$ are independent if $\underline{}$.

6) Variance and Independence. For any two independent random variables $X$ and $Y$,

$$Var(X + Y) = \underline{}$$

Task 1 – Random Variables

Assume that we roll a fair 3-sided die three times. Here, the sides have values 1, 2, 3.

a) Describe the PMF of the random variable $X$ giving the sum of the first two rolls.

- We have $P(X = 2) = 1/9$, $P(X = 3) = 2/9$, $P(X = 4) = 3/9$, $P(X = 5) = 2/9$, and $P(X = 6) = 1/9$.

b) Give the expectation $E(X)$.

- We give a direct proof here, and note that

$$E(X) = 1/9 \cdot (2 + 6) + 2/9 \cdot (3 + 5) + 3/9 \cdot 4 = (8 + 16 + 12)/9 = 4$$

- $\underline{e)}$ Compute $P(X > 3)$.

$$P(X > 3) = 3/9 + 2/9 + 1/9 = 6/9 = 2/3$$

- $\underline{d)}$ Let $Y$ be the random variable describing the sum of the three rolls. Describe the joint PMF of $X$ and $Y$.

- \[
\begin{array}{c|cccccccc}
X / Y & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
2 & 1/27 & 1/27 & 1/27 & 0 & 0 & 0 & 0 \\
3 & 0 & 2/27 & 2/27 & 2/27 & 0 & 0 & 0 \\
4 & 0 & 0 & 3/27 & 3/27 & 3/27 & 0 & 0 \\
5 & 0 & 0 & 0 & 2/27 & 2/27 & 2/27 & 0 \\
6 & 0 & 0 & 0 & 0 & 1/27 & 1/27 & 1/27 \\
\end{array}
\]

- $\underline{e)}$ Compute $P(X = 5 \mid Y = 7)$.
First, \( P(X = 5 \mid Y = 7) = P(X = 5, Y = 7) / P(Y = 7) \). Then, \( P(X = 5, Y = 7) = 2/27 \), whereas
\[
\]
Thus, \( P(X = 5 \mid Y = 7) = 2/27 \cdot 9/2 = 1/3 \).

### Task 2 – Servers

A web service uses \( m \) identical servers for load balancing. Every web request is assigned to one of the servers independently and uniformly at random. Assume the web service receives \( n \) requests.

**a)** For any \( i \in \{1, \ldots, m\} \) and \( j \in \{1, \ldots, n\} \), define the Bernoulli random variable \( X_{i,j} \) which is one if and only the \( j \)-th request is assigned to server \( i \), and zero otherwise.

What is the expected values \( E(X_{i,j}) \)?

Clearly, \( P(X_{i,j} = 1) = \frac{1}{m} \), and thus \( E(X_{i,j}) = 1/m \).

**b)** What is the expected load of server \( i \in \{1, \ldots, m\} \)?

The load of server \( i \) is \( \sum_{j=1}^{n} X_{i,j} \), and using linearity of expectation, we have
\[
E\left( \sum_{j=1}^{n} X_{i,j} \right) = \sum_{j=1}^{n} E(X_{i,j}) = \frac{n}{m}.
\]

### Task 3 – More Linearity

**a)** Alice rolls a fair, six-sided die \( n \) times, what is the expectation of the sum of the \( n \) outcomes?

Let \( X_i \) denote the outcome of the \( i \)-th die, i.e., \( P(X_i = j) = 1/6 \) for all \( i \in [n] \), and \( j \in [6] \). Then, we know that \( E(X_i) = 3.5 \) for all \( i \in [n] \). Also, the sum of the outcomes is defined as \( Z = \sum_{i=1}^{n} X_i \), and thus
\[
E(Z) = \sum_{i=1}^{n} E(X_i) = 3.5 \cdot n.
\]

**b)** Bob plays a game where a die is rolled in each round, until 6 comes out. He wins $3 every time 6 does not appear. How much does Bob expect to win?

We first consider a random variable \( R \) which gives us the number of rounds Bob plays. Clearly, \( R \) is geometric, with parameter \( p = 1/6 \). Thus, \( E(R) = 1/p = 6 \). Now, bob will always win \( W = 3(R - 1) \) dollars, and we can now easily compute \( E(W) = E(3R - 3) = 3E(R) - 3 = 15 \), i.e., he expects to win $15.

**c)** In a room with \( n \) people, how many groups of three people are expected to have the same birthday?

(Assuming birthdays are independent, and equally liked for each of the \( n \) people. Further, assume there are only 365 days.)

There are \( \binom{n}{3} \) groups of three people, i.e., every group of people is represented by a subset \( I \subseteq [n] \) with size \( |I| = 3 \). Now, define \( X_I \) to be the Bernoulli distributed so that it is one if all three people in \( I \) have the same Birthday, and 0 if not. We have that
\[
P(X_I = 1) = \frac{1}{365}, \quad 365 = \frac{1}{365^2},
\]
because there are 365 dates on which the three Birthdays can collide. Thus, \( \mathbb{E}(X_i) = \frac{1}{365} \), and the expected numbers of groups of three people with same Birthdays is

\[
\mathbb{E} \left( \sum_{i} X_i \right) = \sum_{i} \mathbb{E}(X_i) = \left( \frac{n}{3} \right) \cdot 1/(365)^2 .
\]

**Task 4 – Expectations, Independence, and Variance**

a) Give random variables \( X \) and \( Y \) (via their joint PMF) such that \( \mathbb{E}(X \cdot Y) \neq \mathbb{E}(X) \cdot \mathbb{E}(Y) \).

Imagine we flip two correlated coins – they are both heads with probability 1/2 and both tails with probability 1/2. Let \( X \) (resp. \( Y \)) be 1 if and only if the first (resp. second) coin is heads, and 0 otherwise. Then, \( \mathbb{P}(X \cdot Y = 1) = \frac{1}{4} \), and \( \mathbb{P}(X \cdot Y = 0) = \frac{1}{2} \). Thus, \( \mathbb{E}(X \cdot Y) = \frac{1}{2} \).

On the other hand, the individual coins are uniform, and thus \( \mathbb{E}(X) = \mathbb{E}(Y) = \frac{1}{2} \). And, \( \frac{1}{2} \neq \frac{1}{2} \cdot \frac{1}{2} \).

b) Give a random variable \( X \) with range \( \{-1, 1\} \) such that \( \mathbb{E}(X)^2 \neq \mathbb{E}(X^2) \).

We can have for instance \( X \) such that \( \mathbb{P}(X = -1) = \mathbb{P}(X = 1) = \frac{1}{2} \). Then, \( \mathbb{E}(X) = 0 \) and thus \( \mathbb{E}(X)^2 = 0 \). On the other hand, \( X^2 \) is 1 with probability 1, and thus \( \mathbb{E}(X^2) = 1 \).

c) Let \( U \) be a random variable which is uniform over the set \( \{n\} = \{1, 2, \ldots, n\} \), i.e, \( \mathbb{P}(U = i) = \frac{1}{n} \) for all \( i \in \{n\} \). Compute \( \mathbb{E}(U^2) \) and \( \text{Var}(U) \).

**Hint:** \( \sum_{i=1}^{n} i = \frac{n(n+1)}{2} \) and \( \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6} \).

First off, note that

\[
\mathbb{E}(U^2) = \frac{1}{n} \sum_{i=1}^{n} i^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}.
\]

by the hint. Also, note that

\[
\mathbb{E}(U) = \frac{1}{n} \sum_{i=1}^{n} i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.
\]

Therefore

\[
\text{Var}(U) = \mathbb{E}(U^2) - \mathbb{E}(U)^2 = \frac{(n+1)(2n+1)}{6} - \left( \frac{n+1}{2} \right)^2
\]

\[
= \frac{n+1}{12} \cdot (4n + 2 - 3n - 3) = \frac{(n+1)(n-1)}{12}.
\]

d) Let \( Y_1 \) and \( Y_2 \) be the independent outcomes of two dice rolls, and let \( Z = Y_1 + Y_2 \). Then, compute \( \mathbb{E}(Z^2) \) and \( \text{Var}(Z) \).

**Hint:** Try to use an indirect solution using linearity and independence, without the need of explicitly giving the distribution of \( Z^2 \).

First note that by linearity and independence,

\[
\mathbb{E}(Z^2) = \mathbb{E}(Y_1^2) + \mathbb{E}(Y_2^2) + 2\mathbb{E}(Y_1 \cdot Y_2) = \mathbb{E}(Y_1^2) + \mathbb{E}(Y_2^2) + 2\mathbb{E}(Y_1) \mathbb{E}(Y_2) .
\]
We know that \( E (Y_1) = E (Y_2) = 21/6 \). We also know that \( E (Y_1^2) = E (Y_2^2) = 91/6 \) (from class). Thus,
\[
E (Z^2) = 91/3 + 2 \cdot 21^2/36 = 91/3 + 147/6 = 329/6.
\]

On the other hand, we know that \( E (Z) = 7 \). Therefore,
\[
\text{Var}(Z) = E (Z^2) - E (Z)^2 = 329/6 - 294/6 = 35/6.
\]

We could also have used \( \text{Var}(Z) = \text{Var}(Y_1 + Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2) = 35/12 \cdot 2 = 35/6 \), using the calculation from class for the individual variances.