Problem 1. (10 pts.) Independence.

(a) A and C are not necessarily independent. We can see show this by giving one example. So consider two tosses of a fair coin. Let

\[ A = \text{event that the two tosses are different} = \{HT, TH\} \]
\[ B = \text{event that the first toss is heads} = \{HH, HT\} \]
\[ C = \text{event that the two tosses are same} = \{HH, TT\} \]

Clearly \( P(A) = P(B) = P(C) = \frac{1}{2} \), and

\[ P(A \cap B) = P(B \cap C) = \frac{1}{4} = \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right). \]

(Knowing that the first toss is heads tells you nothing about whether the two tosses are the same or different.) So \( A \) and \( B \) are independent, and \( B \) and \( C \) are independent. But \( A \) and \( C \) are not independent: if we know that one event (two tosses different) occurs, then the other (two tosses the same) is impossible. Formally,

\[ P(A \cap C) = 0 \neq P(A)P(C). \]

This example also illustrates part (d) of the problem; do you see how?

(b) Yes, they are mutually independent. To see this we compute all the needed probabilities from the diagram.

\[ P(A) = \frac{1}{5} + \frac{1}{10} + \frac{1}{15} + \frac{2}{15} = \frac{1}{2} \]
\[ P(B) = \frac{1}{10} + \frac{1}{10} + \frac{1}{15} + \frac{1}{15} = \frac{1}{3} \]
\[ P(C) = \frac{2}{15} + \frac{2}{15} + \frac{1}{15} + \frac{1}{15} = \frac{2}{5} \]
\[ P(A \cap C) = \frac{2}{15} + \frac{1}{15} = \frac{1}{5} \]
\[ P(A \cap B) = \frac{1}{10} + \frac{1}{15} = \frac{1}{6} \]
\[ P(B \cap C) = \frac{1}{15} + \frac{1}{15} = \frac{2}{15} \]
\[ P(A \cap B \cap C) = \frac{1}{15} \]

To see mutual independence we have to see that the probability of every pairwise intersection is given by the product of the probabilities and the probability of the 3-way intersection is the product of the three probabilities.

\[ P(A)P(B) = \frac{1}{6} = P(A \cap B) \quad P(A)P(C) = \frac{1}{5} = P(A \cap C) \]
\[ P(B)P(C) = \frac{2}{15} = P(B \cap C) \quad P(A)P(B)P(C) = \frac{1}{15} = P(A \cap B \cap C) \]

(c) No, they are not mutually independent. The easiest way to see this is that

\[ P(A \cap B \cap C) = 0 \neq P(A)P(B)P(C). \]

(d) The reason this should be true is that independence of \( E \) and \( D \) means that knowing \( E \) has happened tells you nothing about the probability that \( D \) has happened. But knowing whether \( E \) has happened is exactly the same as knowing whether \( E^c \) has happened; so \( D \) and \( E^c \) should also be independent.
Checking that the formal definition does what it should is a short calculation. By the
law of total probability (first equality) and the independence of $E$ and $D$ ($P(E \cap C) = P(E)P(D)$, used in the second equality),

$$P(D) = P(E^c \cap D) + P(E \cap D) = P(E^c \cap D) + P(E)P(D).$$

Moving $P(E)P(D)$ to the left side, we get

$$P(E^c \cap D) = P(D) - P(E)P(D) = (1 - P(E))P(D) = P(E^c)P(D).$$

**Problem 2.** (10 pts.)

(a) First we have to compute the probability table for $X$.

<table>
<thead>
<tr>
<th>$x$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(x)$</td>
<td>1/4</td>
<td>3/16</td>
<td>9/64</td>
<td>27/256</td>
<td>81/256</td>
</tr>
</tbody>
</table>

For example, to get $x = 3$ a player has to roll 'not 1, not 1, 1' and this has probability $(3/4)^2(1/4)$. The case $x = 5$ is special because the turn is going to end no matter what. So, the probability is just $(3/4)^4$.

We put these into R and did the calculations there. The R code is posted in the usual place. Here are the computations written out by hand.

$$\mu = E(X) = \frac{1}{4} \cdot 1 + \frac{3}{16} \cdot 2 + \frac{9}{64} \cdot 3 + \frac{27}{256} \cdot 4 + \frac{81}{256} \cdot 5 \approx 3.051.$$  

$$\text{Var}(X) = E((X-\mu)^2) = \frac{1}{4}(1-\mu)^2 + \frac{3}{16}(2-\mu)^2 + \frac{9}{64}(3-\mu)^2 + \frac{27}{256}(4-\mu)^2 + \frac{81}{256}(5-\mu)^2 \approx 2.556.$$  

(We could have also used $\text{Var}(X) = E(X^2) - \mu^2$.)

$$\sigma = \sqrt{\text{Var}(X)} \approx 1.599$$

(b) We graph the pmf of $X$ as point plot and then as a density histogram. The cdf is a staircase graph.

(c) **Answer:** $-0.2286682$ dollars.

The calculation without any tricks is done in ps3-sol.r. The table below sets up the computations and points to a trick that makes the computation easy. The table is organized as follows:
The top row shows the possible scores (number of rolls) for the house and the left most column shows the possible scores for the player.

Each entry in the main part of the table shows the players net winnings for that pair of outcomes. For example, if the player scores a 4 and the house a 3, the 3,4 entry in the table is +1 because the player paid a dollar to play and won $2 for a net of +$1.

The margins show the probabilities, e.g. the 9/64 in the righth-most column is the probability the player’s number of rolls is 3.

Since the player’s and the house’ scores are independent the probability of the outcome in each table entry is just the product of the two corresponding marginal probabilities. For example, the probability of the +1 in the 4,3 entry is \( \frac{27}{256} \times \frac{9}{64} \).

We could compute the expectation by summing up each table value times its probability, but for this particular table there is a trick: all the off-diagonal entries cancel, e.g. the +1 in the 4,3 entry and the −1 in the 3,4 entry have the same probability. So, when we do the sum for expectation these two terms will cancel.

All that’s left is to sum the diagonal entries. We get

\[
-(\frac{1}{4})^2 - (\frac{3}{16})^2 - (\frac{9}{64})^2 - (\frac{27}{256})^2 - (\frac{8}{256})^2 = -0.2286682.
\]

Again we note that the cancellation is particular to this game. Still it is worth saying in words where it comes from: The player and the house are playing exactly the same way and independently; so each outcome for the player beating the house (like \( X = 3, H = 1 \) has exactly the same probability as an outcome for the house beating the player (here \( H = 3, X = 1 \). In the first of these the player gains $1 (winnings minus the fee to play) and in the second the player loses $1 (the fee to play). So these outcomes cancel each other and contribute zero to the expected gain. The other contribution to the expected gain is from ties. In such a case the player loses $1.

Problem 3. (10 pts.) Analyzing data.
For (a) and (b) the R-code is posted on our R-code page in ps3-sol.r

(a) Mean \( \approx 2.554528 \), variance \( \approx 4.3018 \).
Note: we used the R function \texttt{var()} to compute the variance. This uses the following formula for the variance of \( n \) points of data:

\[
\frac{\sum_{j=1}^{5000} (x_j - \mu)^2}{4999}.
\]

We will learn the reason for dividing by 4999 = 5000 − 1 instead of 5000 when we do the statistics portion of the class. In this case, there is very little difference between dividing by 5000 or 4999.

(b)

(c) Looking at the distribution we see it is bimodal with a spike at 5 years. About half the patients die in the first year but about half live more than 2.5 years with over 20% still alive after 5 years. The spike is because everyone who survives to 5 years is lumped into that category. The average of 2.5 years is not that meaningful because there seem to be two categories of patients. This is reflected in the large standard deviation.

(d) The treatment appears to be effective for about half the patients. More research would be needed to understand what characteristics of the disease or patients predict the treatment will be effective.

**Problem 4.** (10 pts.) (a) Let \( X \) be the number of feet I am from the east side of my roof. Since I’m equally likely to be anywhere on the roof, \( X \) follows a Uniform(0,40) distribution, i.e. it has pdf \( f_X(x) = 1/40 \) and cdf \( F_X(x) = x/40 \).

Let \( Y \) be the depth of the snow at a random spot. We are given that \( Y = 8 + X^2/40 \). That is, \( Y \) is a transform of \( X \). So we can use change of variables to compute the pdf of \( Y \). It’s probably easiest to work from the cdf in this case. First, since \( X \) has range \([0,40]\) we see that \( Y \) has range \([8,48]\). For \( 8 \leq y \leq 48 \) we get:

\[
F_Y(y) = P(Y \leq y) = P(8 + X^2/40 \leq y) = P(X \leq \sqrt{40(y - 8)}) = F_X(\sqrt{40(y - 8)}) = \frac{\sqrt{40(y - 8)}}{40}
\]
This is the cdf of $Y$. To find the pdf we just differentiate:

$$f_Y(y) = \frac{1}{2\sqrt{40}(y-8)}.$$ 

The R code for these plots is posted in ps3-sol.r

(b) We are asked for $P(Y \leq 15)$. This is just $F_Y(15)$, which we can compute from part (a). Answer: $P(Y \leq 15) \approx 0.4183$.

**Problem 5.** (10 pts.) **Gallery of continuous random variables.**

(a) (i) You don’t really need R for this: $P(Z \leq 0) = 0.5$.

(ii) $P(Z > 1.5) = 1 - \text{pnorm}(1.5) = 0.0668072$.

(iii) $P(|Z| \leq 1.5) = \text{pnorm}(1.5) - \text{pnorm}(-1.5) = 0.8663856$.

(b) What you were supposed to notice is that these answers are identical to the ones in part (a). This is because $Z = (X - \mu)/\sigma$ is standard normal.

(c) The pdf is $f_Y(y) = \lambda e^{-\lambda y}$. The range of $Y$ starts at 0, so we compute the cdf by an integration

$$F_Y(a) = P(Y \leq a) = \int_0^a \lambda e^{-\lambda y} dy = 1 - e^{-\lambda y}.$$ 

We use this to compute the probability asked for:

$$P(Y \leq 1/\lambda) = F_Y(1/\lambda) = 1 - e^{-1} \approx 0.6321206.$$ 

**Problem 6.** (10 pts.) **Galton-Binet-Stanford.**

(a) The code for these plots is in ps3-sol.r
(b) Let $X$ be the IQ of a randomly chosen person. We want to find $P(X \leq 92)$. We use R to compute this:

$$P(X \leq 92) = 1 - \text{pnorm}(92, 100, 15) = 0.2969014.$$ 

(c) The problem asks us to find a cutoff $c$ such that $P(X > c) = 0.95$. This is the same as finding $c$ so that $P(X \leq c) = 0.05$. We could find this be testing $\text{pnorm}()$ on a lot of values, But R has the $\text{qnorm}()$ function which does exactly what we want.

$$c = \text{qnorm}(0.05, 100, 15) = 75.3272.$$