Question 1: Rolling Dice

Let $X$ be the random variable that is defined as the greater of the two numbers that appear when a pair of dice is rolled. For example, if you roll 2 and 5, then $X = 5$.

a) Determine the expected value and the standard deviation of $X$ for two fair dice whose results are independent from each other.

Let us build a table showing the value of $X$ for all possible pairs of numbers that we can throw with Die 1 ($D_1$) and Die 2 ($D_2$):

<table>
<thead>
<tr>
<th>$D_2$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

We now need to compute the sum of all 36 values of $X$. To do that, let us count how many times each value of $X$ occurs. We find that value 1 occurs once, 2 occurs 3 times, 3 occurs 5 times, 4 occurs 7 times, 5 occurs 9 times, and 6 occurs 11 times. Then we have:

$$X_{\text{sum}} = 1 \cdot 1 + 3 \cdot 2 + 5 \cdot 3 + 7 \cdot 4 + 9 \cdot 5 + 11 \cdot 6 = 1 + 6 + 15 + 28 + 45 + 66 = 161$$

Since all 36 pairs of number occur equally likely, the expected value of $X$ is simply:

$$E(X) = \frac{161}{36} \approx 4.472$$
The variance is given by:

\[
V(X) = \frac{1 \cdot (1 - 4.472)^2 + 3 \cdot (2 - 4.472)^2 + 5 \cdot (3 - 4.472)^2 + 7 \cdot (4 - 4.472)^2 + 9 \cdot (5 - 4.472)^2 + 11 \cdot (6 - 4.472)^2}{36}
\]

\[
V(X) = \frac{1 \cdot 12.056 + 3 \cdot 6.112 + 5 \cdot 2.167 + 7 \cdot 0.223 + 9 \cdot 0.279 + 11 \cdot 2.335}{36} \approx 1.97
\]

Then the standard deviation is:

\[
\sigma(X) = \sqrt{V(X)} \approx 1.4
\]

b) Someone inadvertently steps on one of the two dice, thereby flattening it so that from now on it can only show the numbers one and six (with equal probability). Determine the expected value (no standard deviation) of \(X\) after this accident.

The table now looks like this:

<table>
<thead>
<tr>
<th>(D_2)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(D_1)</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Then we have:

\[
E(X) = \frac{21 + 36}{12} = \frac{57}{12} = 4.75
\]

c) To make things worse, suddenly a mysterious mechanism comes into effect: The sum of the numbers of the two dice (the flat and the normal one) is never seven any more, while all other possible outcomes, i.e., pairs of numbers, occur with equal probability. What is the new expected value of \(X\)?
If we mark the cases that can no longer occur with Xs, we now get the following table:

<table>
<thead>
<tr>
<th>D₂</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>D₁</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>X</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

The expected value is:

\[
E(X) = \frac{15 + 30}{10} = \frac{45}{10} = 4.5
\]

d) In the final, mysterious state of the dice, are the numbers that appear on the flat die \((X_1)\) and on the normal die \((X_2)\) independent? Prove your answer.

No, they are not independent. We can easily give a counterexample:

- \(p(X_1 = 1) = 0.2\)
- \(p(X_2 = 6) = 0.5\)
- \(p(X_1 = 1 \land X_2 = 6) = 0\)

Clearly, \(p(X_1 = 1) \cdot p(X_2 = 6) \neq p(X_1 = 1 \land X_2 = 6)\), so the events are not independent.

Question 2: The Boston Powerflower

Botanists at UMass Boston recently discovered a new local flower species that they named the Boston powerflower. It has beautiful, blue blossoms, and each flower lives for only one summer. During that time, each flower produces 13 seeds. Three of these seeds will turn into flowers in the following year, and the remaining ten seeds will turn into flowers the year after. As the name powerflower suggests, these seeds always turn into flowers; there is no failure ever.

During the year of this discovery (let us call it year zero), the scientists found three powerflowers on the UMass campus, and in the following year (year one), there were already eight of them.
a) Devise a recurrence relation for the number of flowers $f_n$ in year $n$, and specify the initial conditions.

$f_0 = 3$
$f_1 = 8$
$f_n = 3f_{n-1} + 10f_{n-2}$

b) Use the above recurrence relation to predict the number of flowers on the UMass campus in years 2, 3, 4, and 5.

$f_2 = 3 \cdot 8 + 10 \cdot 3 = 54$
$f_3 = 3 \cdot 54 + 10 \cdot 8 = 242$
$f_4 = 3 \cdot 242 + 10 \cdot 54 = 1266$
$f_5 = 3 \cdot 1266 + 10 \cdot 242 = 6218$

c) Find an explicit formula for computing the number of flowers in any given year without requiring iteration, i.e., repeated application of an equation. You should (but do not have to) check the correctness of your formula using some of the results you obtained in (b).

Characteristic equation:

$r^2 - 3r - 10 = 0$

Roots:

$r_1 = 5$
$r_2 = -2$

Then the solution is of the following form:

$f_n = \alpha_1 \cdot 5^n + \alpha_2 \cdot (-2)^n$

In order to find the values for $\alpha_1$ and $\alpha_2$, Let us consider the cases $n = 0$ and $n = 1$, for which we have the initial conditions:

(I) $f_0 = 3 = \alpha_1 \cdot 5^0 + \alpha_2 \cdot (-2)^0 = f_n = \alpha_1 + \alpha_2$

(II) $f_1 = 8 = \alpha_1 \cdot 5 + \alpha_2 \cdot (-2)$

Multiplying Eq. (I) by 2 and adding Eq. (II) to it gives us:
Then we easily see that $\alpha_1 = 2$ and $\alpha_2 = 1$, so the solution becomes:

$$f_n = 2 \cdot 5^n + (-2)^n$$

Let us verify:

$$f_2 = 2 \cdot 5^2 + (-2)^2 = 50 + 4 = 54$$
$$f_3 = 2 \cdot 5^3 + (-2)^3 = 250 - 8 = 242$$
$$f_4 = 2 \cdot 5^4 + (-2)^4 = 1250 + 16 = 1266$$
$$f_5 = 2 \cdot 5^5 + (-2)^5 = 6250 - 32 = 6218$$

Looks good!

**Question 3: Bernoulli and the Red Sox**

Let us assume that the probability of the Boston Red Sox to beat the New York Yankees at a baseball game is 70% and that no ties are possible, i.e., the New York Yankees win 30% of the time. Furthermore, the result of each game is independent of the results of any previous games.

a) To collect some money for charity, the Red Sox and Yankees agree to play ten games against each other. What is the probability that the Red Sox win at most eight of these games?

Here, it is easier to compute the probability of the complementary event $-E$, i.e., the Red Sox winning 9 or 10 games:

$$p(\text{winning 10 games}) = 0.7^{10} \approx 0.0282$$
$$p(\text{winning 9 games}) = \binom{10}{9} \cdot 0.7^9 \cdot 0.3 \approx 0.1211$$

$$p(-E) \approx 0.0282 + 0.1211 = 0.1493$$

$$p(E) = 1 - p(-E) \approx 0.8507 \text{ or } 85.07\%$$

b) In the above charity games, what is the probability that the Yankees win exactly seven games?

$$p = \binom{10}{7} \cdot 0.3^7 \cdot 0.7^3 = 120 \cdot 0.0002187 \cdot 0.343 \approx 0.009 \text{ or } 0.9\%$$

c) What is the probability that the Red Sox “sweep” the Yankees, i.e., win all four games against them in the playoffs?
p = 0.7^4 = 0.2401 or 24.01%

d) What is the probability that the Yankees “sweep” the Red Sox, i.e., win all four games against them in the playoffs?

P = 0.34 = 0.0081 or 0.81%

e) **Bonus:** What is the probability that the Red Sox beat the Yankees in the playoffs, i.e., are the first to win a total of four games before the Yankees do?

In order for the Red Sox to win, there needs to be a series of either 4, 5, 6, or 7 games in which the Red Sox win exactly 4 games and the Yankees fewer than 4. Also, it is clear that this series has to end with a win by the Red Sox. So the possibilities are:

- Red Sox win 4 games straight: \( p_1 = 0.7^4 = 0.2401 \)
- Red Sox win 3 out of 4 games, followed by another Red Sox win. \( p_2 = C(4, 3) \cdot 0.7^4 \cdot 0.3 = 4 \cdot 0.2401 \cdot 0.3 = 0.28812 \)
- Red Sox win 3 out of 5 games, followed by another Red Sox win. \( p_3 = C(5, 3) \cdot 0.7^4 \cdot 0.3^2 = 10 \cdot 0.2401 \cdot 0.09 = 0.21609 \)
- Red Sox win 3 out of 6 games, followed by another Red Sox win. \( p_4 = C(6, 3) \cdot 0.7^4 \cdot 0.3^3 = 20 \cdot 0.2401 \cdot 0.027 = 0.12965 \)

Because none of these cases can occur at the same time (i.e., the events are disjoint), we can simply add these probabilities:

\[
p = p_1 + p_2 + p_3 + p_4 \approx 0.874 \text{ or } 87.4\%
\]

f) **Bonus:** Let us assume (the unlikely case that) the Yankees won the first two games against the Red Sox in the playoffs. What is the probability that the Red Sox will still beat the Yankees, i.e., win four games before the Yankees win another two?

In this case, the Red Sox have to either win 4 games straight or win 3 out of 4 games, followed by another win. Using the terms from (e), we get:

\[
p = p_1 + p_2 = 0.2401 + 0.2881 = 0.5282 \text{ or } 52.82\%
\]
Question 4: Urns and Probabilities

a) An urn contains three red balls and four blue balls. You draw two balls from the urn, one after the other, without putting them back. Event A is that the first ball is red, and event B is that the second ball is red. Are events A and B independent? Prove your answer.

\[ p(A) = \frac{3}{7} \]

If A occurs, i.e., the first ball is red, then \( p(B) = \frac{2}{6} = \frac{1}{3} \), otherwise \( p(B) = \frac{3}{6} = \frac{1}{2} \). Since \( p(A) = \frac{3}{7} \), we have:

\[ p(B) = \frac{3}{7} \cdot \frac{1}{3} + \frac{4}{7} \cdot \frac{1}{2} = \frac{1}{7} + \frac{2}{7} = \frac{3}{7} \]

From our reasoning above know that \( p(A \cap B) = \frac{3}{7} \cdot \frac{1}{3} = \frac{1}{7} \)

\[ p(A) \cdot p(B) = \frac{9}{49} \]

Therefore, \( p(A \cap B) \neq p(A) \cdot p(B) \), and the events are not independent.

b) This time you are using the same urn and balls, but you draw three balls, again one after the other. Event C is that the first two balls you draw have different colors, and Event D is that the third ball is blue. Are events C and D independent? Again, prove your answer.

The probability of C is the probability of drawing a red ball and then a blue ball plus the probability of drawing a blue ball and then a red ball:

\[ p(C) = \frac{3}{7} \cdot \frac{2}{3} + \frac{4}{7} \cdot \frac{1}{2} = \frac{2}{7} + \frac{2}{7} = \frac{4}{7} \]

We can actually simplify the computation of \( p(D) \). In part (a), we found that the probabilities of the first and second balls being red were equal. That is no coincidence. Let us imagine that we picked all seven balls, one by one, without replacing them. In other words, we are producing random permutations of the set of all balls, and there is no reason that probabilities could differ across positions in the sequence. So the probability that the first ball is red is the same as the second, third, and so on, being red. In each case, as we have three red and four blue balls, the probability of picking a red ball will always be \( \frac{3}{7} \). Therefore:
p(D) = 4/7.

Both C and D occur for two sequences: red – blue – blue and blue – red – blue:

\[ p(C \cap D) = \frac{3}{7} \cdot \frac{4}{6} \cdot \frac{3}{5} + \frac{4}{7} \cdot \frac{3}{6} \cdot \frac{3}{5} = \frac{36}{210} + \frac{36}{210} = \frac{72}{210} = \frac{12}{35} \]

\[ p(C) \cdot p(D) = \frac{4}{7} \cdot \frac{4}{7} = \frac{16}{49} \]

Therefore, \( p(C \cap D) \neq p(C) \cdot p(D) \), and the events are not independent.