Discrete Expected Value and Variance

1. **If Z is some random variable, how do you find the *expected value* of Z?**

   The expected value of a random variable is a *weighted average* of the values that Z can take, where the weights are given by the probability of obtaining that value. So the formula is:

   $$E(Z) = \sum k \cdot P(Z = k)$$

   How do you find the *variance* of Z (two ways)?

   We can think of the variance in a random variable as telling us how much we vary from the expected value of the random variable (which is often written as μ). This can be expressed as:

   $$\mathrm{Var}(Z) = E\left((X-\mu)^2\right)$$

   However, we can manipulate this equation to get another formula that can be easier to use in practice:

   $$\mathrm{Var}(Z) = E\left(X^2\right) - E(X)^2$$

2. **Suppose you roll two, fair 4-sided dice. Let X denote the minimum of the two rolls and Y the maximum of the two rolls. Are these two random variables *independent*?**

   No! There are several different ways to see this. For instance, consider the events $X = 4$ and $Y = 1$. Each of these occurs with probability $\frac{1}{16}$ – since there is only one way to have the minimum value of the two rolls be 4, and the same for the maximum being 1. So independence demands that $P(X = 4 \text{ and } Y = 4) = \frac{1}{16} \cdot \frac{1}{16}$. However, this joint event is impossible, i.e. it has probability 0. So the events are **not independent**, and so the random variables are also **not independent**.

   For the above random variables, find the following:

   a) The expected value of X, $E(X)$

   We can find the pmf of X by simple counting. Then the formula for expected value of X gives us:

   $$E(X) = 1 \cdot \frac{7}{16} + 2 \cdot \frac{5}{16} + 3 \cdot \frac{3}{16} + 4 \cdot \frac{1}{16} = \frac{30}{16}$$

   b) $E(Y)$
   This can be computed as we did in the previous part.

   $$E(Y) = 1 \cdot \frac{1}{16} + 2 \cdot \frac{3}{16} + 3 \cdot \frac{5}{16} + 4 \cdot \frac{7}{16} = \frac{50}{16}$$

   Note that $E(X) + E(Y) = E(X+Y) = \frac{80}{16} = 5$. Is there another way to see this?

   c) $E(2X + 3Y)$
   Expectation of linear combinations can be split up into pieces.

   $$E(2X+3Y) = 2E(X) + 3E(Y) = 2 \cdot \frac{30}{16} + 3 \cdot \frac{50}{16} = \frac{210}{16}$$

d) $E(XY)$
We **cannot** simply multiply the two expected values because the random variables are not independent. First, we can find the pmf of the product:

XY	Probability
1	$1/16$
2	$2/16$
3	$2/16$
4	$3/16$
6	$2/16$
8	$2/16$
9	$1/16$
12	$2/16$
16	$1/16$

So multiplying across each row and then adding these values up gives us $E(XY) = \frac{100}{16}$.

e) The variance of X, $\text{Var}(X)$
We can use the second formula for variance. Which means we first need to find:

$$E\left(X^2\right) = 1^2 \cdot \frac{7}{16} + 2^2 \cdot \frac{5}{16} + 3^2 \cdot \frac{3}{16} + 4^2 \cdot \frac{1}{16} = \frac{70}{16}$$

Then remember to subtract off $E(X)^2$! So we get that:

$$\text{Var}(X) = \frac{70}{16} - \left(\frac{30}{16}\right)^2 = \frac{55}{64}$$

f) $\text{Var}(2Y)$
We will use the important fact that $\text{Var}(2Y) = 2^2\text{Var}(Y)$. Then this is as before:

$$\text{Var}(2Y) = 4 \cdot \left[\frac{170}{16} - \left(\frac{50}{16}\right)^2\right] = 4 \cdot \frac{55}{64}$$

Notice that $\text{Var}(X) = \text{Var}(Y)$; is there a way to see this without calculating?

3. **You have a weighted coin that lands on tails twice as often as it lands on heads. If you flip this coin 10 times, what is the expected number of heads you get? What is the standard error of this random variable?**

This is a **binomial random variable** with the probability of success $p = 1/3$ and the number of trials $n = 10$. So the the expected number of successes is just $np = 10/3$. The standard error is $\sqrt{np(1-p)} = \frac{\sqrt{20}}{3}$.

4. **Someone flips a biased coin an unknown number of times. They tell you that the average number of heads is 10 with a variance of 5. Can you figure out both the chance that the coin lands on heads *and* the number of times the coin was flipped?**

This is also a **binomial random variable**. however, now we do not know the parameters, p and n; that's what we are to find. Instead, we know that the expected value is $np = 10$. Also, the variance is $np(1-p) = 5$. Then from these two equations, we can solve for our two unknowns. We end up with $p = 0.5$ and $n = 20$, so it was actually a fair coin that was flipped 20 times.

5. **Suppose you are drawing cards from a standard deck of 52 cards. What is the expected number of cards you will need until you draw the Ace of Spades?**

The random variable in this case can take any value from 1 to 52 (we could draw it immediately, or it could be at the bottom of the deck). Next we want to figure out the probabilities that each of these happens. By symmetry, the Ace of Spades is as likely to be in one position as any other, so this is just a **uniform distribution**. Then the expected number of draws is:

$$\sum_{k=1}^{52} = k \cdot \frac{1}{52} = \frac{53}{2} = 26.5$$

What is the variance of the number of cards until you get the Ace of Spades?
For variance of the uniform distribution, we could use the formula fro the sum of squares. Or we could just use the formula as it was given in class. If X is uniformly distributed over $\{1, \cdots, n\}$, then $\text{Var}(X) = \frac{n^2-1}{12}$. So for us, the variance is just $\frac{52^2-1}{12}$.

6. **On average, your friend has a 10% of making a certain trick shot. If you agree to let them shoot until they make it, how many tries do you expect them to take? What is the standard error of this random variable?**

This is a **geometric distribution**, as we are counting the number of attempts until a success; the parameter is $p = 0.1$ chance of success. Note that the distribution we use in class only counts the number of failures, so we will have to add 1 at the end to account for the try that succeeds. Then using the known formulas, we get that the expected number of failures is $\frac{0.9}{0.1} = 9$, so we expect them to take 10 tries. The variance for this distribution is $\frac{0.9}{0.1^2} = 90$, so the standard error is $\sqrt{90}$ or about 9.5.

($\star\star$) **How many tries would you expect them to need to make the shot twice?**
We can try to calculate this expected value by just using the infinite sum:

$$\sum_{k=1}^{\infty} k \cdot P(X = k)$$

However, another way to write this, which can make it easier to figure out is:

$$E(X) = \sum_{k=1}^{\infty} P(X \geq k).$$

Now, we just need to find the probability of taking at least k tries to make the shot twice. This is the same as saying that in the first $k-1$ shots, there were 0 or 1 success. So $P(X \geq k) = (1-p)^{k-1} + (k-1)p(1-p)^{k-2}$, where the first term means they failed every time, and the second term means they made the shot once. So we can write:

$$E(X) = \sum_{k=1}^{\infty} \left((1-p)^{k-1} + (k-1)p(1-p)^{k-2}\right)$$

Then to evaluate these, note the infinite sum from the first term is geometric and equals $\frac{1}{p}$. The second infinite sum is just the expected number of trials to get one success, and so also equals $\frac{1}{p}$. So the total expected tries to get two successes when there is a 10% chance of making the shot is $\frac{2}{0.1} = 20$. This is twice the expected number of tries to get one success; why is that?

7. **Suppose you draw 5 cards from a standard deck of 52 cards. What is the expected number of queens you get? What is the standard error of the number of queens that you get?**

   One extremely useful technique that can be used to find the expected number of queens in a 5-card hand is splitting the random variable up into simpler pieces. Let's call Y the number of queens in our hand. We can let Y_1 be the random variable that is 0 if the first card is not a queen and 1 if the first card is a queen. Similarly, Y_2 will be 0 or 1 depending on whether the second card is not a queen or is a queen. Likewise for Y_3, Y_4, and Y_5. Then we can separate $Y = Y_1 + Y_2 + Y_3 + Y_4 + Y_5$ and we have that:

   $$E(Y) = E(Y_1) + E(Y_2) + E(Y_3) + E(Y_4) + E(Y_5) = 5 \cdot \frac{1}{13}$$

   This is because

   $$E(Y_1) = 0 \cdot P(\text{first card is not a queen}) + 1 \cdot P(\text{first card is a queen})$$

   So the expected number of queens is $\frac{5}{13} \approx 0.38$.

   We could also observe that this is a **hypergeometric distribution**. There are 52 total objects, 4 of which are *good*. Then we pull 5 objects and what to know the number of *good* objects we get. As was provided in class, we can find that:

   $$SE(Y) = \sqrt{\text{Var}(Y)} = \sqrt{5 \cdot \frac{4}{52}\frac{48}{52}\frac{47}{51}} \approx 0.57$$

8. **You observe that a radioactive sample releases an emission on average once every ten seconds. What is the probability that there is at least one emission over the course of a minute? What is the expected number of emissions in a minute? What is the standard error for the number of emissions in a minute?**

   This is a **Poisson random variable**, as we are not considering some sequence of discrete trials that can succeed or fail. Rather we are waiting some time period and recording the number of successes. The Poisson parameter $\lambda = 6$, since this is the average number of events that happen over 60 seconds. Then we also have that this is both the expected value for our radioactive decay and the variance. And so the standard error is $\sqrt{6} \approx 2.45$.