

Chapter 3. Discrete Random Variables

3.5: Zoo of Discrete Random Variables Part II

[Slides \(Google Drive\)](#)

Alex Tsun

[Video \(YouTube\)](#)

3.5.1 The Uniform (Discrete) Random Variable

In this lecture we will continue to expand our zoo of discrete random variables. The next one we will discuss is the uniform random variable. This models situations where the probability of each value in the range is equally likely, like the roll of a fair die.

Definition 3.5.1: Uniform Random Variable

X is a uniform random variable, denoted $X \sim \text{Unif}(a, b)$, where $a < b$ are integers, if and only if X has the following probability mass function

$$p_X(k) = \begin{cases} \frac{1}{b-a+1}, & k \in \{a, a+1, \dots, b\} \\ 0, & \text{otherwise} \end{cases}$$

X is equally likely to take on any value in $\Omega_X = \{a, a+1, \dots, b\}$. This set contains $b-a+1$ integers, which is why $\mathbb{P}(X = k)$ is always $\frac{1}{b-a+1}$.

Additionally,

$$\mathbb{E}[X] = \frac{a+b}{2} \text{ and } \text{Var}(X) = \frac{(b-a)(b-a+1)}{12}$$

As you might expect, the expected value is just the average of the endpoints that the uniform random variable is defined over.

Proof of Expectation and Variance of Uniform.

Suppose $X \sim \text{Unif}(a, b)$. We need to use the fact that $\sum_{i=1}^n i = \frac{n(n+1)}{2}$ and $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$ to compute some quantities. I will skip some steps as it is pretty tedious and just algebra, but just focus on the setup.

$$\mathbb{E}[X] = \sum_{k=a}^b k \cdot p_X(k) = \sum_{k=a}^b k \cdot \frac{1}{b-a+1} = \frac{1}{b-a+1} \sum_{k=a}^b k = \dots = \frac{a+b}{2}$$

$$\mathbb{E}[X^2] = \sum_{k=a}^b k^2 \cdot p_X(k) = \sum_{k=a}^b k^2 \cdot \frac{1}{b-a+1} = \frac{1}{b-a+1} \sum_{k=a}^b k^2 = \dots$$

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{(b-a)(b-a+1)}{12}$$

□

This variable models situations like rolling a fair six sided die. Let X be the random variable whose value is the number face up on a die roll. Since the die is fair each outcome is equally likely, which means that $X \sim \text{Unif}(1, 6)$ so

$$p_X(k) = \begin{cases} \frac{1}{6}, & k \in \{1, 2, \dots, 6\} \\ 0, & \text{otherwise} \end{cases}$$

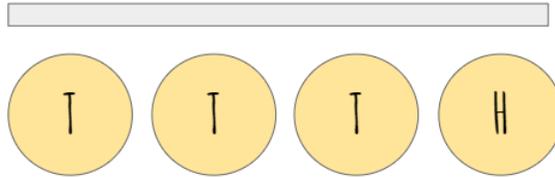
This is fairly intuitive, but is nice to have these formulas in our zoo so we can make computations quickly, and think about random processes in an organized fashion. Using the equations above we can find that

$$\mathbb{E}[X] = \frac{1+6}{2} = 3.5 \quad \text{and} \quad \text{Var}(X) = \frac{(6-1)(6-1+2)}{12} = \frac{35}{12}$$

3.5.2 The Geometric Random Variable

Another random variable that arises from the Bernoulli process is the Geometric random variable. It models situations that can be thought of as the number of trials up to and including the first success.

For example, suppose we are betting on how many **independent** flips it will take for a coin to land heads for the first time. The coin lands heads with a probability p , and you feel confident that it will take four flips to get your first head. The only way that this can occur is with the following sequence of flips (since our first head must have been on the fourth trial, we know everything before must be tails):



Let X be the random variable that represents the number of **independent** coin flips up to and including your first head. Lets compute $\mathbb{P}(X = 4)$. $X = 4$ occurs exactly when there are 3 tails followed by a head. So,

$$\mathbb{P}(X = 4) = \mathbb{P}(TTTH) = (1-p)(1-p)(1-p)p = (1-p)^3 p$$

In general,

$$p_X(k) = (1-p)^{k-1} p$$

This is because there must be $k-1$ tails in a row followed by a head occurring on the k th trial.

Let's also verify that the probabilities sum to 1.

$$\begin{aligned} \sum_{k=1}^{\infty} p_X(k) &= \sum_{k=1}^{\infty} (1-p)^{k-1} p && \text{[Geometric PMF]} \\ &= p \sum_{k=1}^{\infty} (1-p)^{k-1} && \text{[take out constant]} \\ &= p \sum_{k=0}^{\infty} (1-p)^k && \text{[reindex to 0]} \\ &= p \left(\frac{1}{1-(1-p)} \right) && \left[\text{geometric series formula: } \sum_{i=0}^{\infty} r^i = \frac{1}{1-r} \right] \\ &= p \cdot \frac{1}{p} = 1 \end{aligned}$$

The second last step used the geometric series formula - this may be why this random variable is called Geometric!

Definition 3.5.2: Geometric Random Variable

X is a Geometric random variable, denoted $X \sim \text{Geo}(p)$, if and only if X has the following probability mass function (and range $\Omega_X = \{1, 2, \dots\}$):

$$p_X(k) = (1-p)^{k-1}p, \quad k = 1, 2, 3, \dots$$

Additionally,

$$\mathbb{E}[X] = \frac{1}{p} \quad \text{and} \quad \text{Var}(X) = \frac{1-p}{p^2}$$

Proof of Expectation and Variance of Geometric.

Suppose $X \sim \text{Geo}(p)$. The expectation is pretty complicated and uses a calculus trick, so don't worry about it too much. Just understand the first two lines, which are the setup! But before that, what do you think it should be? For example, if $p = 1/10$, how many flips do you think it would take until our first head? Possibly 10? And if $p = 1/7$, maybe 7? So seems like our guess will be $\mathbb{E}[X] = \frac{1}{p}$. It turns out this intuition is actually correct!

$$\begin{aligned} \mathbb{E}[X] &= \sum_{k \in \Omega_X} k \cdot p_X(k) && \text{[def of expectation]} \\ &= \sum_{k=1}^{\infty} k(1-p)^{k-1}p \\ &= p \sum_{k=1}^{\infty} k(1-p)^{k-1} && [p \text{ is a constant with respect to } k] \\ &= p \sum_{k=1}^{\infty} \frac{d}{dp} (-(1-p)^k) && \left[\frac{d}{dy} y^k = ky^{k-1}, \text{ and chain rule of calculus} \right] \\ &= -p \left(\frac{d}{dp} \sum_{k=1}^{\infty} (1-p)^{k-1} \right) && \text{[swap sum and integral]} \\ &= -p \left(\frac{d}{dp} \frac{1}{1-(1-p)} \right) && \left[\text{geometric series formula: } \sum_{i=0}^{\infty} r^i = \frac{1}{1-r} \right] \\ &= -p \left(\frac{d}{dp} \frac{1}{p} \right) \\ &= -p \left(-\frac{1}{p^2} \right) \\ &= \frac{1}{p} \end{aligned}$$

We'll actually have a much nicer proof of this fact in 5.3 using the law of total expectation, so look forward to that! I hope you'll take my word that $\mathbb{E}[X^2]$ is even worse, so I will not provide that proof. But the expectation and variance are here for you to cite! \square

Example(s)

Let's say you buy lottery tickets every day, and the probability you win on a given day is 0.01, independently of other days. What is the probability that after a year (365 days), you still haven't won? What is the expected number of days until you win your first lottery?

Solution If X is the number of days until the first win, then $X \sim \text{Geo}(p = 0.01)$. Hence, the probability we don't win after a year is (using the PMF)

$$\mathbb{P}(X \geq 365) = 1 - \mathbb{P}(X < 365) = 1 - \sum_{k=1}^{364} \mathbb{P}(X = k) = 1 - \sum_{k=1}^{364} (1 - 0.01)^{k-1} 0.01$$

This is great, but for the geometric, we can actually get a closed-form formula by thinking of what it means that $X \geq 365$ in English. $X \geq 365$ happens if and only if we lose for the first 365 days, which happens with probability 0.99^{365} . If you evaluated that nasty sum above and this quantity, you would find that they are equal!

Finally, we can just cite the expectation of the Geometric RV:

$$\mathbb{E}[X] = \frac{1}{p} = \frac{1}{0.01} = 100$$

This is the point of the zoo! We do all these generic calculations so we can use them later anytime. \square

Example(s)

You gamble by flipping a fair coin independently up to and including the first head. If it takes k tries, you earn $\$2^k$ (i.e., if your first head was the flip, you would earn $\$8$). How much would you pay to play this game?

Solution Let X be the number of flips to the first head. Then, $X \sim \text{Geo}(\frac{1}{2})$ because its a fair coin, and

$$p_X(k) = \left(1 - \frac{1}{2}\right)^{k-1} \left(\frac{1}{2}\right) = \frac{1}{2^k} \quad k = 1, 2, 3, \dots$$

It is usually unwise to gamble, especially if your expected earnings are lower than the price to play. So, let Y be your expected earnings. Note that $Y = 2^X$ because the amount you win depends the number of flips it takes to get a heads. We will use LOTUS to compute $\mathbb{E}[Y] = \mathbb{E}[2^X]$. Recall $\mathbb{E}[2^X] \neq 2^{\mathbb{E}[X]} = 2^2 = 4$ as we've seen many times now.

$$\mathbb{E}[Y] = \mathbb{E}[2^X] = \sum_{k=1}^{\infty} 2^k p_X(k) = \sum_{k=1}^{\infty} 2^k \frac{1}{2^k} = \sum_{k=1}^{\infty} 1 = \infty$$

So, you are expected to win an infinite amount of money!

Some might say they would be willing to pay any finite amount of money to play this game. Think about why that would be unwise, and what this means regarding the modeling tools we have provided you so far. \square

3.5.3 The Negative Binomial Random Variable

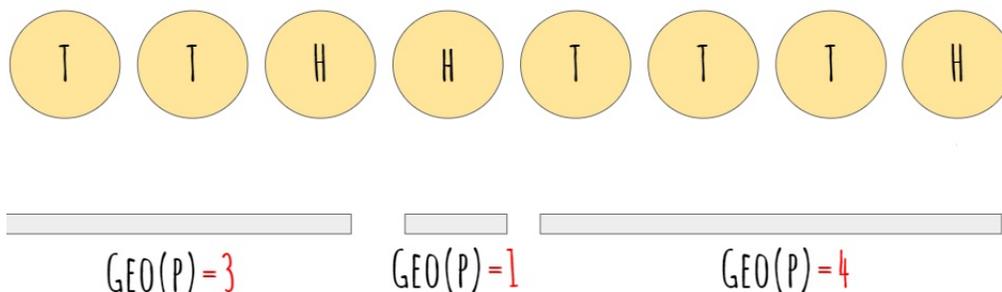
Consider the situation where we are not just betting on the first head, but on the first r heads. How could we use a random variables to model this scenario?

If you'll recall from the last lecture, multiple Bernoulli random variables sum together to produce a more complicated random variable, the binomial. We might try to do something similar with geometric random variables.

Let X be a random variable that represents the number of coin flips it takes to get our r^{th} head.

$$X = \sum_{i=1}^r X_i$$

where X_i is a geometric random variable that represents the number of flips it takes to get the i th head after $i - 1$ heads have already occurred. Since all the flips are independent, so are the rvs X_1, \dots, X_r . For example, if $r = 3$ we might observe the following sequence of flips



In this case, $X_1 = 3$ and represents the number of trials between the 0th to the 1st head; $X_2 = 1$ and represents the number of trials between the 1st to the 2nd head; $X_3 = 4$ and represents the number of trials between the 2nd and the 3rd head. Remember this fact for later!

How do we find $\mathbb{P}(X = 8)$? There must be exactly 5 heads and 3 tails, so it is reasonable to expect $(1 - p)^5 p^3$ to come up somewhere in our final formula, but how many ways can we get a valid sequence of flips? Note that the last coin flip *must* be a heads, otherwise we would've gotten our r heads earlier than our 8th flip. From here, any 2 of the first 7 flips can be heads, and 5 of must be tails. Thus, there are $\binom{7}{2}$ valid sequences of coin flips.

Each of these 7 flip sub-sequences (of the 8 total flips) occurs with probability $(1 - p)^5 p^2$ and there is no overlap. However, we need to include the probability that the last coin flip is a heads. So,

$$p_X(8) = \mathbb{P}(X = 8) = \binom{7}{2} (1 - p)^5 p^2 \cdot p = \binom{7}{2} (1 - p)^5 p^3$$

We can generalize as follows:

$$p_X(k) = \binom{k-1}{r-1} (1 - p)^{k-r} p^r$$

Again, the interpretation is that our r^{th} head must come at the k^{th} trial exactly; so in the first $k - 1$ we can get $r - 1$ heads anywhere (hence the binomial coefficient), and overall we have r heads and $k - r$ tails.

If we are interested in finding the expected value of X we might try the brute force approach directly from the definition of expected value

$$\mathbb{E}[X] = \sum_{k \in \Omega_X} k p_X(k) = \sum_{k \in \Omega_X} k \binom{k-1}{r-1} (1 - p)^{k-r} p^r$$

but this approach is overly complicated, and there is a much simpler way using linearity of expectation! Suppose $X_1, \dots, X_r \sim \text{Geo}(p)$ are independent. As we showed earlier, $X = \sum_{i=1}^r X_i$, and we showed that each $\mathbb{E}[X_i] = 1/p$. Using linearity of expectation, we can derive the following:

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^r X_i\right] = \sum_{i=1}^r \mathbb{E}[X_i] = \sum_{i=1}^r \frac{1}{p} = \frac{r}{p}$$

Using a similar technique and the (yet unproven) fact that $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$, we can find the variance of X from the sum of the variances of multiple geometric random variables

$$\text{Var}(X) = \text{Var}\left(\sum_{i=1}^r X_i\right) = \sum_{i=1}^r \text{Var}(X_i) = \sum_{i=1}^r \frac{1-p}{p^2} = \frac{r(1-p)}{p^2}$$

This random variable is called the negative binomial random variable. It is quite common so it too deserves a special place in our zoo.

Definition 3.5.3: Negative Binomial Random Variable

X is a negative binomial random variable, denoted $X \sim \text{NegBin}(r, p)$, if and only if X has the following probability mass function (and range $\Omega_X = \{r, r+1, \dots\}$):

$$p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, \dots$$

X is the sum of r independent $\text{Geo}(p)$ random variables.

Additionally,

$$\mathbb{E}[X] = \frac{r}{p} \quad \text{and} \quad \text{Var}(X) = \frac{r(1-p)}{p^2}$$

Also, note that $\text{Geo}(p) \equiv \text{NegBin}(1, p)$, and that if X, Y are independent such that $X \sim \text{NegBin}(r, p)$ and $Y \sim \text{NegBin}(s, p)$, then $X + Y \sim \text{NegBin}(r + s, p)$ (waiting for $r + s$ heads).

3.5.4 Exercises

1. You are a hardworking boxer. Your coach tells you that the probability of your winning a boxing match is 0.25, independently of every other match.
 - (a) How many matches do you expect to fight until you win once?
 - (b) How many matches do you expect to fight until you win ten times?
 - (c) You only get to play 12 matches every year. To win a spot in the Annual Boxing Championship, a boxer needs to win at least 10 matches in a year. What is the probability that you will go to the Championship this year?
 - (d) Let q be your answer from the previous part. How many times can you expect to go to the Championship in your 20 year career?

Solution:

- (a) Let X be the matches you have to fight until you win once. Then, $X \sim \text{Geo}(p = 0.25)$, so $\mathbb{E}[X] = \frac{1}{p} = \frac{1}{0.25} = 4$.
- (b) Let Y be the matches you have to fight until you win ten times. Then, $Y \sim \text{NegBin}(r = 10, p = 0.25)$, so $\mathbb{E}[Y] = \frac{r}{p} = \frac{10}{0.25} = 40$.
- (c) Let Z be the number of matches you win out of 12. Then, $Z \sim \text{Bin}(n = 12, p = 0.25)$, and we want

$$\mathbb{P}(Z \geq 10) = \sum_{k=10}^{12} \binom{12}{k} 0.2^k (1 - 0.2)^{12-k}$$

- (d) Let W be the number of times we make it to the Championship in 20 years. Then, $W \sim \text{Bin}(n = 20, p = q)$, and

$$\mathbb{E}[W] = np = 20q$$

2. You are in music class, and your cruel teacher says you cannot leave until you play the 1000-note song Fur Elise correctly 5 times. You start playing the song, and if you play an incorrect note, you immediately start the song over from scratch. You play each note correctly independently with probability 0.999.

- (a) What is the probability you play the 1000-note song Fur Elise correctly immediately? (i.e., the first 1000 notes are all correct).
- (b) What is the probability you take exactly 20 attempts to correctly play the song 5 times?
- (c) What is the probability you take at least 20 attempts to correctly play the song 5 times?
- (d) (Challenge) What is the expected number of **notes** you play until you finish playing Fur Elise correctly 5 times?

Solution:

- (a) Let X be the number of correct notes we play in Fur Elise in one attempt, so $X \sim \text{Bin}(1000, 0.999)$. We need $\mathbb{P}(X = 1000) = 0.999^{1000} \approx 0.3677$.
- (b) If Y is the number of attempts until we play the song correctly 5 times, then $Y \sim \text{NegBin}(5, 0.3677)$, and so

$$\mathbb{P}(Y = 20) = \binom{20-1}{5-1} 0.3677^5 (1 - 0.3677)^{15} \approx 0.0269$$

- (c) We can actually take two approaches to this. We can either take our Y from earlier, and compute

$$\mathbb{P}(Y \geq 20) = 1 - \mathbb{P}(Y < 20) = 1 - \sum_{k=5}^{19} \binom{k-1}{4} 0.3677^5 (1 - 0.3677)^{k-5} \approx 0.1161$$

Notice the sum starts at 5 since that's the lowest possible value of Y . This would be exactly the probability of the statement asked. We could alternatively rephrase the question as: what is the probability we play the song correctly at most 4 times correctly in the first 19 times? Check that these questions are equivalent! Then, we can let $Z \sim \text{Bin}(19, 0.3677)$ and instead compute

$$\mathbb{P}(Z \leq 4) = \sum_{k=0}^4 \binom{19}{k} 0.3677^k (1 - 0.3677)^{19-k} \approx 0.1161$$

- (d) We will have to revisit this question later in the course! Note that we could have computed the expected number of **attempts** to finish playing Fur Elise though, as it would follow a $\text{NegBin}(5, 0.3677)$ distribution with expectation $\frac{5}{0.3677} \approx 13.598$.