Recall:

- Bernoulli Random Variable, $X \sim Bern(p)$
- Binomial $X \sim Bino(n, p)$
- Geometric $X \sim Geo(p)$
- Expectation $E[X]$
- Hypergeometric, $X \sim Hyge(n, M, N)$ where a finite population of size N is divided into two types, say S and F, where there are M S's (and hence $N - M$ F's), you sample (SRS) with size n, and count the number of S's, x.
- $E[X] \equiv \sum x_i p(x_i)$

Expected Value of a Binomially Distributed Random Variable

Let an experiment consist of a fixed number of independent trials, say n. For each trial, let the trial have two possible outcomes, success and failure, with constant probabilities p and $1-p$. Note that the sample space here consists of sequences of length n of the form, e.g. SFSSSF... Define the random variable X as a function which associates a sequence with the number of successes in that sequence.

As we have seen from the previous examples, the probability distribution function is of the form

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k = 0, 1, 2, \ldots, n$.

We now prove that the expected value of a binomially distributed random variable is given by $E = np$.

proof

$$\begin{aligned}
E &\equiv \sum_{x_i} x_i \ f(x_i) \\
E &= \sum_{k=0}^{n} k \ f(k) \\
E &= \sum_{k=0}^{n} k \ \binom{n}{k} \ p^k \ (1-p)^{n-k} \\
E &= \sum_{k=0}^{n} k \ \frac{n!}{k!(n-k)!} \ p^k \ (1-p)^{n-k}
\end{aligned}$$

$$
\begin{aligned}
E &= \sum_{k=1}^{n} k \, \frac{n!}{k!(n-k)!} \, p^k \, (1-p)^{n-k} \\
E &= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} \, p^k \, (1-p)^{n-k} \\
E &= \sum_{k=1}^{n} \frac{n(n-1)!}{(k-1)!(n-k)!} \, pp^{k-1} \, (1-p)^{n-k} \\
E &= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} \, p^{k-1} \, (1-p)^{n-k} \\
E &= np \sum_{k=1}^{n} \frac{(n-1)!}{(n-1-(k-1))!(k-1)!} \, p^{k-1} \, (1-p)^{n-1-(k-1)} \\
E &= np \sum_{k=1}^{n} \tbinom{n-1}{k-1} \, p^{k-1} \, (1-p)^{n-1-(k-1)}
\end{aligned}
$$

Note that the sum equals 1. (hint: use the binomial formula for $(p+(1-p))^{n-1}$).

Thus

$$E = np(1) = np$$

Expected Value of a Geometric Random Variable:
Example: You will roll a fair die until it comes up SIX. What is the probability that SIX will show up for the first time on the

1. First roll?

2. Second roll?

3. Third roll?

4. Fourth roll?

5. i^{th} roll?

Sampling Without Replacement: Hypergeometric Random Variable Compare the binomial case above with the following. Suppose you have a shipment of 26 hard drives. You will want to do a spot check on these drives to see if there are defects. Because of time constraints you will only test five of them. Suppose that 4 of the 26 drives are defective. Graph the probability distribution function and cumulative distribution function for the random variable which maps individual samples (outcomes) to the number of defects in that sample. On average, how many defective drives would you expect to detect?

The Expected Value of a Function of a Discrete Random Variable Very often we are interested in a function of a random variable. Two interesting examples:

- Suppose you determine that the radii of tree trunks have a textbook distribution, e.g. they are normally distributed. What would the distribution be for their **areas**? Can it be normal?
- You have a two sided "spinner" placed in the plane at the point $(0, 1)$. If the angle that the spinner makes clockwise from the horizontal is uniformly distributed from 0 to π, what does the distribution of points on the real line "being pointed to" look like?

As a simpler example, imagine that you are playing the usual game of tossing a fair die and being paid $1 for HEADS and losing $1 for TAILS. You will play 6 times. Let X be the random variable which counts your earnings. What is the probability function of X? What is its expected value?

Now let Y be defined as $Y \equiv X^2$. What is the probability function of Y? What is its expected value? In particular, is it true that $E[X^2] = E[X]^2$?

Result Suppose a random variable Y is a function of the random variable X. Suppose further that X has a probability mass function $p(x)$ and a set of possible values D. If $Y = h(X)$ then we may compute

$$E[Y] = E[h(X)] = \sum_{x \in D} h(x) \cdot p(x)$$

Is this what we did above? Why is this result important?

The Variance of a Discrete Random Variable Just as with a set of data, we wish to find a measure of spread for a probability distribution. That is, are we likely to get values near one another, or are they likely to be relatively far from each other? Analogously to the data case, define the variance of a discrete random variable X as

$$V[X] \equiv E[(X - E[X])^2] = \sum (x_i - \mu)^2 p(x_i)$$

- Compute the variance of a Bernoulli Random Variable.
- Compute the variance of a Binomial Random Variable.

Examples: Taken from *Introduction to Probability Models* by Sheldon Ross.

1. Suppose items produced by a certain machine are defective with probability 0.1, independently of one another. What is the probability that in a sample of 3 items, at most one will be defective?

2. Suppose that an airplane engine will fail, when in flight, with probability $1 - p$ independently from engine to engine. The flight is successful if at least 50% of its engines remain operative. For what values of p is a four engine plane preferable to a two engine plane?