Handout 08: Binomial and Geometric Distributions

More Random Variable Results

For two random variables X and Y, the expected value of $X+Y$ is equal to the expected value of X plus the expected value of Y. We cannot prove this, or even rigorously state this result until we formalize the concept of joint distributions. However, we can state a weaker formula that will be very useful and insightful.

Theorem 1 (Weak Linearity of Expectation) *Let g and h both be functions that map the random variable X into the real line. Then $\mathbb{E}(g(X)+h(X))$ is equivalent to $\mathbb{E}g(X)+\mathbb{E}h(X)$.*

Proof. This is a simple algebraic manipulation:

$$\begin{aligned}
\mathbb{E}(g(X)+h(X)) &= \sum_x p(x) \cdot [h(x)+g(x)] \\
&= \sum_x p(x) \cdot h(x) + p(x) \cdot g(x) \\
&= \mathbb{E}g(x) + \mathbb{E}h(x)
\end{aligned}$$

Where the summations are over the support of X ■.

As an example of the utility of this result, note that there is often a simpler way to compute the variance of a random variable.

Theorem 2 *Let X be a random variable with finite variance. Then, $Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2$.*

Proof. This is another simple algebraic manipulation. Set $\mu = \mathbb{E}X$, then:

$$\begin{aligned}
\mathbb{E}(X-\mu)^2 &= \mathbb{E}\left(X^2+\mu^2-2X\mu\right) \\
&= \mathbb{E}X^2 + \mu^2 - 2\mu\mathbb{E}X \\
&= \mathbb{E}X^2 + \mu^2 - 2\mu^2 \\
&= \mathbb{E}X^2 - \mu^2
\end{aligned}$$

Using the linearity of the expectation ■.

Often it will be easier to compute $\mathbb{E}X^2$ instead of directly calculating the variance.

Some Discrete Distributions

A Bernoulli trial is a random variable that only takes on the values 0 and 1. It can be uniquely specified via the probability that it is equal to 1 (usually denoted by p), as then the probability that it is equal to 0 is just $1-p$. We saw this on the last worksheet. We say that such a random variable has a Bernoulli distribution and write this as:

$$X \sim \text{Bernoulli}(p)$$

On the last worksheet you found the expected value and variance of this random variable. We will now define three important distributions in terms of Bernoulli trials. The definitions depend on a sequence of independent trials, a concept that we have not formally defined, but should make intuitive sense.

Consider a sequence of independent Bernoulli trials X_1, X_2, and so forth, then:

- The sum $Y = \sum_{i=1}^{n} X_i$ has a *Binomial* distribution, with parameters n and p. We write $Y \sim Bin(n, p)$.
- The number of 0's, Y, obtained before the first 1 has a *Geometric* distribution. We write $Y \sim Geom(p)$.
- The number of 0's, Y, obtained before the first r 1's has a *Negative Binomial* distribution. We write $Y \sim NB(k, p)$.

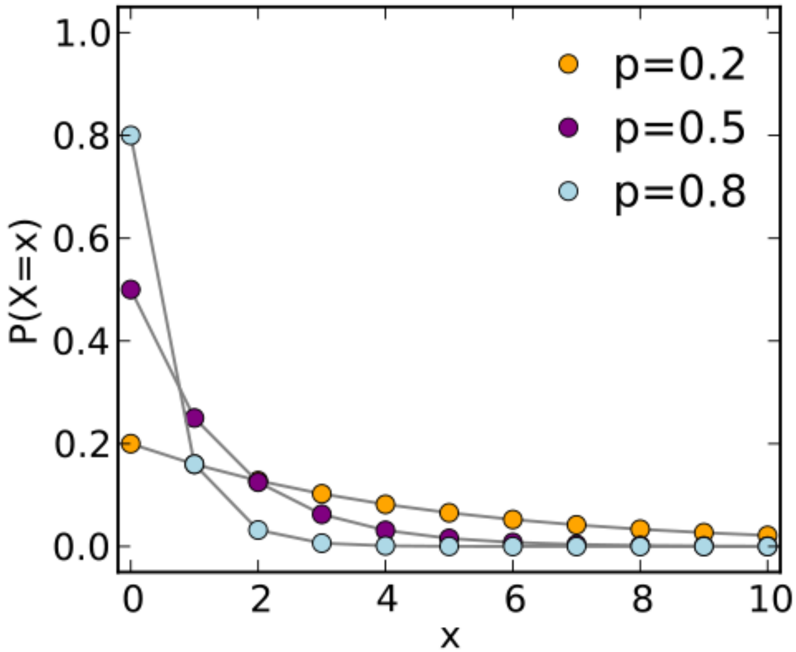
Of course, $Geom(p)$ is the same as $NB(1, p)$, but the geometric distribution is important enough to have its own name.

The Geometric Distribution

Let $X \sim Geom(p)$. Using the logic from the past worksheet we see quickly that the probability mass function of X is given by:

$$\mathbb{P}(X = k) = (1-p)^k \cdot p$$

This just uses the naïve definition of counting (there is only one way to obtain k 0's before the first 1). Using this definition we can calculate the expected value and variance using some straightforward, but subtle, summation tricks. Here is the probability mass function:



Theorem 3 (Geometric Expectation and Variance) *Let $X \sim Geom(p)$. Then $\mathbb{E}X$ is equal to $\frac{1-p}{p}$.*

Proof. By definition:

$$\mathbb{E}(X) = \sum_{k=0}^{\infty} kq^k p$$

where q = 1 − p. This sum looks unpleasant; it's not a geometric series because of the extra k multiplying each term. But we notice that each term looks similar to $k \cdot q^{(k-1)}$, the derivative of q^k (with respect to q), so let's start there:

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}.$$

This geometric series converges since $0 < q < 1$. Differentiating both sides with respect to q, we get

$$\sum_{k=0}^{\infty} kq^{k-1} = \frac{1}{(1-q)^2}.$$

Finally, if we multiply both sides by pq, we recover the original sum we wanted to find:

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=0}^{\infty} kq^k p \\ &= pq \sum_{k=0}^{\infty} kq^{k-1} \\ &= pq \frac{1}{(1-q)^2} \\ &= \frac{q}{p}.\end{aligned}$$

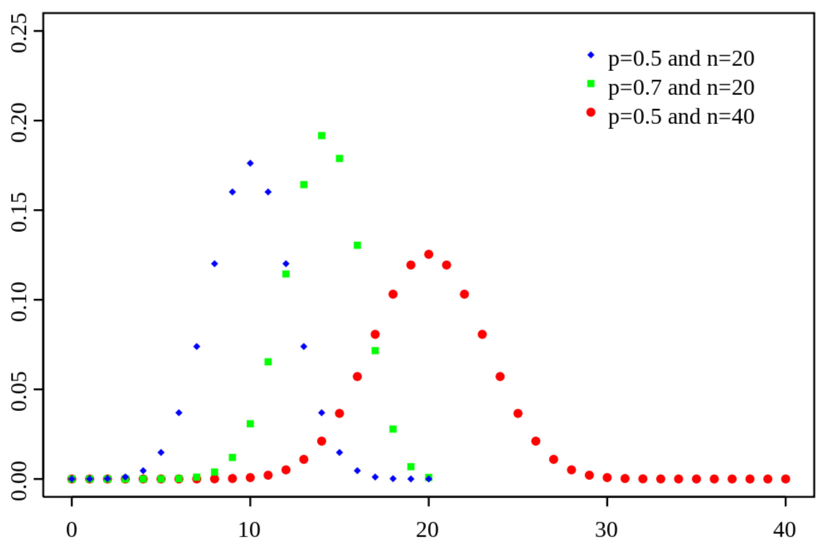
Substituting back for q finishes the result ■.

Binomial Distribution

A similar story yields the probability mass function for a binomial random variable. If $X \sim Bin(n,p)$, then:

$$\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

This is, again, just a counting problem. There really was a reason for learning all of those counting rules! Here is the probability mass function:



The next two results establish closed-form formulas for the mean and variance of the binomial distribution.

Theorem 4 *If $X \sim B(n,p)$, then $EX = np$.*

Proof. Proving this results requires combining a number of simple steps:

$$
\begin{aligned}
\mathbb{E}X &= \sum_{x=0}^{n} x \cdot \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} x \cdot \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, \text{ as above is zero when } x \text{ is zero} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x}, \text{ canceling first term of factorial} \\
&= np \cdot \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x}
\end{aligned}
$$

Making the substitution $y = x - 1$:

$$
\begin{aligned}
\mathbb{E}X &= np \cdot \sum_{y=0}^{n-1} \frac{(n-1)!}{(y)!(n-1-y)!} p^y (1-p)^{n-1-y} \\
&= np \cdot \sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{(n-1)-y} \\
&= np
\end{aligned}
$$

The final line comes because the sum is the distribution of a variable with a $B(n-1,p)$ distribution $\blacksquare$.

We will see that many probability results can be established by rewriting a summand (and later, integrals) as known probability mass functions.

Theorem 5 *If $X \sim B(n,p)$, then $V(X) = np(1-p)$.*

Proof. We start by seeing that the expected value of $\mathbb{E}(X(X-1))$ can be derived very similarly to the expected value calculation:

$$
\begin{aligned}
\mathbb{E}(X(X-1)) &= \sum_{x=0}^{n} x \cdot (x-1) \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=2}^{n} x \cdot (x-1) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, \text{ as above is zero when } x < 2 \\
&= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x}, \text{ canceling first terms of factorial} \\
&= n(n-1)p^2 \cdot \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x}
\end{aligned}
$$

Making the substitution $y = x - 2$:

$$\begin{aligned}
\mathbb{E}(X(X-1)) &= n(n-1)p^2 \cdot \sum_{y=0}^{n-2} \frac{(n-2)!}{(y)!(n-2-y)!} p^y (1-p)^{n-2-y} \\
&= n(n-1)p^2 \cdot \sum_{y=0}^{n-2} \binom{n-2}{y} p^y (1-p)^{(n-2)-y} \\
&= n(n-1)p^2
\end{aligned}$$

The final line comes, this time, because the sum is the distribution of a variable with a $B(n-2, p)$ distribution. Then, the variance can be computed as:

$$\begin{aligned}
V(X) &= \mathbb{E}X^2 - (\mathbb{E}X)^2 \\
&= \mathbb{E}(X^2 - X + X) - (\mathbb{E}X)^2 \\
&= \mathbb{E}(X \cdot (X-1) + X) - (\mathbb{E}X)^2 \\
&= \mathbb{E}(X \cdot (X-1)) + \mathbb{E}X - (\mathbb{E}X)^2 \\
&= n(n-1)p^2 + np - n^2p^2 \\
&= np - np^2 \\
&= np(1-p)
\end{aligned}$$

As desired $\blacksquare$.