MEAN/VARIANCE OF BINOMIAL AND NEGATIVE BINOMIAL RANDOM VARIABLE

JAMES MARSHALL REBER

The reason for writing these notes is that I am unable to find nice proofs online for these facts, so I'd like to have a centralized place to reference them. All definitions can be found in my notes or on Wikipedia. The goal will be to prove the mean of a binomial/negative binomial random variable via algebraic manipulations. We proceed with the calculations now.

Claim. The mean of a binomial random variable $X \sim \mathrm{Bin}(n,p)$ is

$$\mathbb{E}(X) = np.$$

Proof. The nicest proof I've found is on Wikipedia, which uses the linearity of expectation. The gist is to rewrite the binomial random variable as a sum of i.i.d. Bernoulli random variables, and then use the fact that a Bernoulli random variable has mean p. To be consistent with the proof for the negative binomial random variable, we will try to instead prove it from scratch.

Let's first just write the definition. We have

$$\mathbb{E}(X) = \sum_{k=0}^{\infty} k\binom{n}{k}p^k(1-p)^{n-k}.$$

The first observation I'll make here is that we will define $\binom{n}{k} = 0$ for $k > n$ (so that this series is defined). The next observation to make is to note that at $k = 0$ this sum is just zero, so without loss of generality we can just start the series at $k = 1$. The final observation is to notice that

$$\begin{aligned} k\binom{n}{k} &= k \cdot \frac{n!}{k!(n-k)!} = \frac{n!}{(k-1)!(n-k)!} = \frac{n-k+1}{n-k+1}\frac{n!}{(k-1)!(n-k)!} \\ &= n-k+1\frac{n!}{(k-1)!(n-k+1)!} = (n-k+1)\binom{n}{k-1}. \end{aligned}$$

We use this substitution to rewrite the above as

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=1}^{\infty}(n-k+1)\binom{n}{k-1}p^k(1-p)^{n-k} \\ &= \sum_{k=1}^{\infty} n\binom{n}{k-1}p^k(1-p)^{n-k} - \sum_{k=1}^{\infty}(k-1)\binom{n}{k-1}p^k(1-p)^{n-k}. \end{aligned}$$

In the term on the left, this is almost just the sum of the pmf from $k = 1$ to ∞. Pulling out terms, we have

$$\begin{aligned} \sum_{k=1}^{\infty} n\binom{n}{k-1}p^k(1-p)^{n-k} &= \frac{np}{1-p}\sum_{k=1}^{\infty}\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1} \\ &= \frac{np}{1-p}\sum_{k=0}^{\infty}\binom{n}{k}p^k(1-p)^{n-k} = \frac{np}{1-p}. \end{aligned}$$

Now examining the term on the right, we almost have the expected value. Let's pull out terms again. This gives us

$$\begin{aligned}\sum_{k=1}^{\infty}(k-1)\binom{n}{k-1}p^k(1-p)^{n-k} &= \frac{p}{1-p}\sum_{k=1}^{\infty}(k-1)\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1} \\ &= \frac{p}{1-p}\sum_{k=0}^{\infty}k\binom{n}{k-1}p^k(1-p)^{n-k} = \frac{p}{1-p}\mathbb{E}(X).\end{aligned}$$

Putting it all together, we have

$$\mathbb{E}(X) = \frac{np}{1-p} - \frac{p}{1-p}\mathbb{E}(X).$$

We can add $p/(1-p)\mathbb{E}(X)$ to both sides to get

$$\left(\frac{p}{1-p}+1\right)\mathbb{E}(X) = \frac{\mathbb{E}(X)}{1-p} = \frac{np}{1-p}.$$

Multiply both sides by $1-p$ to get

$$\mathbb{E}(X) = np,$$

as desired. □

Of course in the above proof it would have been much faster to just use the linearity trick, however the advantage of going about the proof this way is that we can adapt it for the negative binomial random variable.

Claim. The mean of a negative binomial random variable $X \sim \text{NB}(k,p)$ is

$$\mathbb{E}(X) = \frac{kp}{1-p}.$$

Remark. There are many definitions for the negative binomial random variable. We are using the one with the pmf defined by

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

In other words, we are measuring the probability of n success before $k-1$ failures occur. See Wikipedia for the alternative formulations and their means.

Proof. We use the same trick as before (and since it's the exact same trick, I'm going to be a little careless with details). By definition, the mean is

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} n\binom{n+k-1}{n}p^n(1-p)^{k-1}.$$

Notice

$$n\binom{n+k-1}{n} = n\frac{(n+k-1)!}{n!(k-1)!} = \frac{(n+k-1)!}{(n-1)!(k-1)!} = (n+k-1)\frac{(n-1+k-1)!}{(n-1)!(k-1)!} = (n+k-1)\binom{n+k-2}{n-1}.$$

Rewriting the above, we have

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{n=1}^{\infty}(n+k-1)\binom{n+k-2}{n-1}p^n(1-p)^{k-1} \\
&= \sum_{n=1}^{\infty}(n-1)\binom{n+k-2}{n-1}p^n(1-p)^{k-1} + \sum_{n=1}^{\infty}k\binom{n+k-2}{n-1}p^n(1-p)^{k-1} \\
&= p\sum_{n=1}^{\infty}(n-1)\binom{n+k-2}{n-1}p^{n-1}(1-p)^{k-1} + kp\sum_{n=1}^{\infty}\binom{n+k-2}{n-1}p^{n-1}(1-p)^{k-1} \\
&= p\sum_{n=0}^{\infty}n\binom{n+k-1}{n}p^n(1-p)^{k-1} + kp\sum_{n=0}^{\infty}\binom{n+k-1}{n}p^n(1-p)^{k-1} \\
&= p\mathbb{E}(X) + kp.
\end{aligned}$$

Now we rearrange again to get

$$(1-p)\mathbb{E}(X) = kp.$$

Divide both sides by $1-p$ to get

$$\mathbb{E}(X) = \frac{kp}{1-p}.$$

□

The core behind these proofs is the following observation.

Claim. Suppose $n \geq k \geq 0$. We have that

$$k\binom{n}{k} = (n-k+1)\binom{n}{k-1}.$$

Notice that this trick will also give us the variance for both the binomial and the negative binomial..

Problem 1. If $X \sim \text{Bin}(n,p)$, then

$$\text{Var}(X) = np(1-p).$$

Proof. Details will be sparse in this part because it follows the same trick as before.

Recall the variance can be defined as

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

We know $\mathbb{E}(X)$, so the goal is to figure out $\mathbb{E}(X^2)$. By definition, we see

$$\mathbb{E}(X^2) = \sum_{k=0}^{\infty}k^2\binom{n}{k}p^k(1-p)^{n-k}.$$

Use the trick from before (writing $k^2 = k \cdot k$) to get

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

Now, we can rewrite

$$k(n-k+1) = -k(k-1) + nk.$$

Distributing the sum, we have

$$\mathbb{E}(X^2) = \sum_{k=1}^{\infty}nk\binom{n}{k-1}p^k(1-p)^{n-k} - \sum_{k=1}^{\infty}k(k-1)\binom{n}{k-1}p^k(1-p)^{n-k}.$$

Focusing on the term on the left, we can simplify a little to get

$$\frac{np}{1-p}\sum_{k=1}^{\infty} k\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1}.$$

Ideally we'd like a $(k-1)$ term on the inside instead of k since this would match everything else. We can add and subtract to get this. Notice that

$$\begin{aligned}\frac{np}{1-p}\sum_{k=1}^{\infty} k\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1} &= \frac{np}{1-p}\sum_{k=1}^{\infty} k\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1}\\ -\frac{np}{1-p}\sum_{k=1}^{\infty}\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1} &+ \frac{np}{1-p}\sum_{k=1}^{\infty}\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1}\\ = \frac{np}{1-p}\sum_{k=1}^{\infty}(k-1)\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1} &+ \frac{np}{1-p}\sum_{k=1}^{\infty}\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1}.\end{aligned}$$

Now the term on the far left in this sum can be rewritten as

$$\frac{np}{1-p}\sum_{k=0}^{\infty} k\binom{n}{k}p^{k}(1-p)^{n-k} = \frac{np}{1-p}\mathbb{E}(X).$$

The term on the far right can be rewritten as

$$\frac{np}{1-p}\sum_{k=0}^{\infty}\binom{n}{k}p^{k}(1-p)^{n-k} = \frac{np}{1-p}.$$

Thus

$$\frac{np}{1-p}\sum_{k=1}^{\infty} k\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1} = \frac{np}{1-p}(\mathbb{E}(X)+1).$$

Going back to our original sum, we focus on the term on the right. We have

$$\sum_{k=1}^{\infty} k(k-1)\binom{n}{k-1}p^{k}(1-p)^{n-k}.$$

This is close to having a $(k-1)^1$ term in the middle. If we subtract and add, we can get that term. Doing so, we have

$$\begin{aligned}\sum_{k=1}^{\infty} k(k-1)\binom{n}{k-1}p^{k}(1-p)^{n-k} &= \sum_{k=1}^{\infty} k(k-1)\binom{n}{k-1}p^{k}(1-p)^{n-k}\\ -\sum_{k=1}^{\infty}(k-1)\binom{n}{k-1}p^{k}(1-p)^{n-k} &+ \sum_{k=1}^{\infty}(k-1)\binom{n}{k-1}p^{k}(1-p)^{n-k}\\ = \sum_{k=1}^{\infty}(k-1)^2\binom{n}{k-1}p^{k}(1-p)^{n-k} &+ \sum_{k=1}^{\infty}(k-1)\binom{n}{k-1}p^{k}(1-p)^{n-k}.\end{aligned}$$

Now, focusing on the term on the left in this sum, we have

$$\sum_{k=1}^{\infty}(k-1)^2\binom{n}{k-1}p^{k}(1-p)^{n-k} = \frac{p}{1-p}\sum_{k=1}^{\infty}(k-1)^2\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1} = \frac{p}{1-p}\mathbb{E}(X^2).$$

Focusing on the term on the right in this sum, we have

$$\sum_{k=1}^{\infty}(k-1)\binom{n}{k-1}p^k(1-p)^{n-k} = \frac{p}{1-p}\sum_{k=1}^{\infty}(k-1)\binom{n}{k-1}p^{k-1}(1-p)^{n-k+1} = \frac{p}{1-p}\mathbb{E}(X).$$

Thus

$$\sum_{k=1}^{\infty}k(k-1)\binom{n}{k-1}p^k(1-p)^{n-k} = \frac{p}{1-p}\mathbb{E}(X^2) + \frac{p}{1-p}\mathbb{E}(X).$$

Putting it all together, we see

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

Using the fact that $\mathbb{E}(X) = np$ and solving for $\mathbb{E}(X^2)$, we have

$$\mathbb{E}(X^2) = np(np - p + 1).$$

Thus

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = (np)^2 - np^2 + np - (np)^2 = np(1-p).$$

$\square$

Claim. If $X \sim \mathrm{NB}(k, p)$, then

$$\mathrm{Var}(X) = \frac{kp}{(1-p)^2}.$$

Proof. Again, details will be sparse in this part because it follows the same trick as before.

Recall the variance can be defined as

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

We see that

$$\mathbb{E}(X^2) = \sum_{n=0}^{\infty} n^2\binom{n+k-1}{n}p^n(1-p)^{k-1}.$$

Notice we can apply the trick once to get

$$\mathbb{E}(X^2) = \sum_{n=1}^{\infty} n(n+k-1)\binom{n+k-2}{n-1}p^n(1-p)^{k-1}.$$

Expanding out, we have $n(n+k-1) = n(n-1) + nk$. Thus rewriting the above yields

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{n=1}^{\infty}(n(n-1)+nk)\binom{n+k-2}{n-1}p^n(1-p)^{k-1}\\ &= \sum_{n=1}^{\infty} n(n-1)\binom{n+k-2}{n-1}p^n(1-p)^{k-1} + k\sum_{n=1}^{\infty} n\binom{n+k-2}{n-1}p^n(1-p)^{k-1}.\end{aligned}$$

Focusing on the term on the left, we need to do a common analysis trick of adding zero. The motivation for this is that the term looks almost like $(n-1)^2$, which would be great since that gives us back $\mathbb{E}(X^2)$. So we will subtract off whatever is necessary to get this, and then add it back on to ensure we still have the same term. Doing so, we have the term is equal to

$$\begin{aligned}
&\sum_{n=1}^{\infty} n(n-1)\binom{n+k-2}{n-1}p^n(1-p)^{k-1} - \sum_{n=1}^{\infty}(n-1)\binom{n+k-2}{n-1}p^n(1-p)^{k-1} \\
&\qquad + \sum_{n=1}^{\infty}(n-1)\binom{n+k-2}{n-1}p^n(1-p)^{k-1} \\
&= \sum_{n=1}^{\infty}(n-1)^2\binom{n+k-2}{n-1}p^n(1-p)^{k-1} + \sum_{n=1}^{\infty}(n-1)\binom{n+k-2}{n-1}p^n(1-p)^{k-1} \\
&= p\mathbb{E}(X^2) + p\mathbb{E}(X).
\end{aligned}$$

Focusing on the term on the right, we repeat the same trick. We notice it is close to $(n-1)$, so we subtract off whatever we need to in order to get that and then add it back on. Doing so, we have the term is equal to

$$\begin{aligned}
&k\sum_{n=1}^{\infty} n\binom{n+k-2}{n-1}p^n(1-p)^{k-1} - k\sum_{n=1}^{\infty}\binom{n+k-2}{n-1}p^n(1-p)^{k-1} + k\sum_{n=1}^{\infty}\binom{n+k-2}{n-1}p^n(1-p)^{k-1} \\
&\qquad k\sum_{n=1}^{\infty}(n-1)\binom{n+k-2}{n-1}p^n(1-p)^{k-1} + k\sum_{n=1}^{\infty}\binom{n+k-2}{n-1}p^n(1-p)^{k-1} \\
&= kp\mathbb{E}(X) + kp.
\end{aligned}$$

Putting it all together, we have

$$\mathbb{E}(X^2) = p\mathbb{E}(X^2) + (kp+p)\mathbb{E}(X) + kp,$$

or

$$\mathbb{E}(X^2) = \frac{kp(kp+1)}{(1-p)^2}.$$

Now we need to subtract off $\mathbb{E}(X)^2$ in order to get the variance. Doing so, we have

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{kp}{(1-p)^2}.$$

□