

Section 3.8 Negative Binomial Distribution

Tuesday, September 27, 2016 12:53 PM



CHAPTER 3 STAT/MATH 511, J. TEBBS

Finally,

$$V(Y) = E(Y^2) - [E(Y)]^2 = \frac{1+q}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{q}{p^2} \square$$

NOTE: WMS derive the geometric mean and variance using a different approach (not using the mgf). See pp 116-117. \square

Example 3.17. At an orchard in Maine, "20-lb" bags of apples are weighed. Suppose that four percent of the bags are underweight and that each bag weighed is independent. If Y denotes the number of bags observed to find the first underweight bag, then $Y \sim \text{geom}(p = 0.04)$. The mean of Y is

$$E(Y) = \frac{1}{p} = \frac{1}{0.04} = 25 \text{ bags.}$$

The variance of Y is

$$V(Y) = \frac{q}{p^2} = \frac{0.96}{(0.04)^2} = 600 \text{ (bags)}^2. \square$$

3.8 Negative binomial distribution

NOTE: The negative binomial distribution can be motivated from two perspectives:

- as a generalization of the geometric
- as an "inverse" version of the binomial.

TERMINOLOGY: Imagine an experiment where Bernoulli trials are observed. If Y denotes the trial on which the r th success occurs, $r \geq 1$, then Y has a **negative binomial distribution** with waiting parameter r and probability of success p .

NEGATIVE BINOMIAL PMF: The pmf for $Y \sim \text{nib}(r, p)$ is given by

$$p_Y(y) = \begin{cases} \binom{y-1}{r-1} p^r (1-p)^{y-r}, & y = r, r+1, r+2, \dots \\ 0, & \text{otherwise.} \end{cases}$$

Of course, when $r = 1$, the $\text{nib}(r, p)$ pmf reduces to the $\text{geom}(p)$ pmf.

PAGE 48

$$\sum_{y=r}^{\infty} \binom{y-1}{r-1} p^r (1-p)^{y-r} = 1$$

CHAPTER 3 STAT/MATH 511, J. TEBBS

RATIONALE: The form of $p_Y(y)$ can be explained intuitively. If the r th success occurs on the y th trial, then $r-1$ successes must have occurred during the first $y-1$ trials. The total number of sample points (in the underlying sample space S) where this occurs is given by the binomial coefficient $\binom{y-1}{r-1}$, which counts the number of ways you can choose the locations of $r-1$ successes in a string of the first $y-1$ trials. The probability of any particular such ordering, by independence, is given by $p^{r-1}(1-p)^{y-r}$. Thus, the probability of getting exactly $r-1$ successes in the $y-1$ trials is $\binom{y-1}{r-1} p^{r-1} (1-p)^{y-r}$. On the y th trial, we observe the r th success (this occurs with probability p). Because the y th trial is independent of the previous $y-1$ trials, we have

$$P(Y = y) = \underbrace{\binom{y-1}{r-1} p^{r-1} (1-p)^{y-r}}_{\text{pertains to 1st } y-1 \text{ trials}} \times p = \binom{y-1}{r-1} p^r (1-p)^{y-r}.$$

Example 3.18. A botanist is observing oak trees for the presence of a certain disease. From past experience, it is known that 30 percent of all trees are infected ($p = 0.30$). "success" labels "infected"

Treating each tree as a Bernoulli trial (i.e., each tree is infected/not), what is the probability that she will observe the 3rd infected tree ($r = 3$) on the 6th, or 7th observed tree?

SOLUTION. Let Y denote the tree on which she observes the 3rd infected tree. Then, step 1: $\text{nib}(r=3, p=0.3)$. We want to compute $P(Y = 6 \text{ or } Y = 7)$. The $\text{nib}(3, 0.3)$ pmf gives step 2: $\text{nib}(r=3, p=0.3)$. We want to compute $P(Y = 6 \text{ or } Y = 7)$. The $\text{nib}(3, 0.3)$ pmf gives

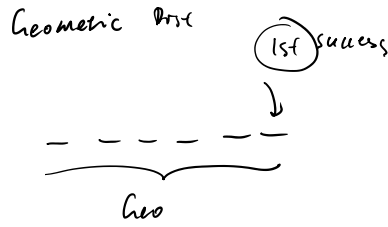
$$p_Y(6) = P(Y = 6) = \binom{6-1}{3-1} (0.3)^3 (1-0.3)^{6-3} = 0.0926$$

$$p_Y(7) = P(Y = 7) = \binom{7-1}{3-1} (0.3)^3 (1-0.3)^{7-3} = 0.0972$$

Thus, $P_Y(6) = 0.3 \times \text{binompdf}(6-1, .3, 3-1)$

$$P(Y = 6 \text{ or } Y = 7) = P(Y = 6) + P(Y = 7) = 0.0926 + 0.0972 = 0.1898. \square$$

RELATIONSHIP WITH THE BINOMIAL: Recall that in a binomial experiment, we fix the number of Bernoulli trials, n , and we observe the number of successes. In a

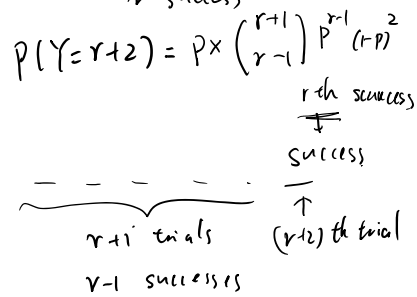
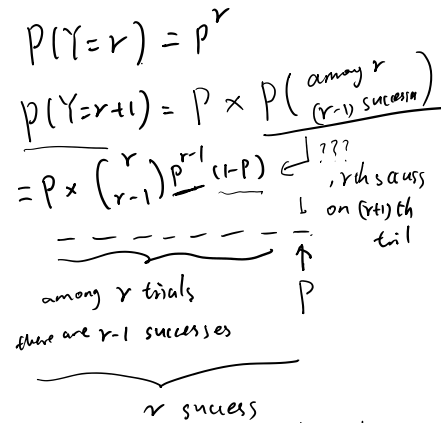


Neg. Binomial

Y : # of trials to observe the r th success

If $r=1$, Geometric.

the support = $\{r, r+1, r+2, \dots, \infty\}$



$y \geq r$

$P(Y=y) = p \times \binom{y-1}{r-1} p^{r-1} (1-p)^{y-r}$

$= p \times \binom{y-1}{r-1} p^{r-1} (1-p)^{y-r}$

If $X \sim \text{Binomial}(n, p)$

$P(X=x) = \text{binompdf}(n, p, x)$

For $Y \sim \text{Neg Binomid}(r, p)$

$P(Y=y) = p \times \text{binompdf}(y-1, p, r-1)$

$$P(Y=6 \text{ or } Y=7) = P(Y=6) + P(Y=7) = 0.0926 + 0.0972 = 0.1898. \square$$

RELATIONSHIP WITH THE BINOMIAL: Recall that in a binomial experiment, we fix the number of Bernoulli trials, n , and we observe the number of successes. In a negative binomial experiment, we fix the number of successes we are to observe, r , and we continue to observe Bernoulli trials until we reach that numbered success. In this sense, the negative binomial distribution is the "inverse" of the binomial distribution.

For $Y \sim \text{Neg Bin}(r, p)$
 $P(Y=y) = P(X \text{ binompdf}(y-1, p, r-1))$

$$P(Y \leq y) = \sum_{a=r}^y P(Y=a)$$

$$P(Y \leq 100) = \sum_{a=3}^{100} P(Y=a) \text{ tedious}$$

$$P(Y \leq y) = 1 - P(Y > y) = 1 - \text{binomcdf}(y, p, r-1)$$

$Y > y$: # of trials to observe the r th success is greater than y

equivalent

\Leftrightarrow r th success occurs after the y th trial

\Leftrightarrow among the y trials, there are at most $r-1$ successes

Define X to be the # of successes among the y trials

$$X \sim \text{Binomial}(y, p)$$

$$P(X \leq r-1) = \text{binomialcdf}(y, p, r-1)$$

Negative Binomial MGF:

$$M_Y(t) = E[e^{tY}]$$

$$= \sum_{y=r}^{\infty} e^{ty} P_Y(y)$$

$$e^{ty} = e^{tr} \times e^{t(y-r)}$$

$$= \sum_{y=r}^{\infty} e^{ty} \binom{y-1}{r-1} p^r (1-p)^{y-r}$$

$$\sum_{y=r}^{\infty} \binom{y-1}{r-1} \frac{p^r}{z^r (1-z)^{y-r}} = 1 \text{ for any } 0 < z \leq 1$$

$$= \sum_{y=r}^{\infty} \binom{y-1}{r-1} (e^{tp})^r \times (e^{t(y-r)} \times (1-p)^{y-r})$$

$$= \sum_{y=r}^{\infty} \binom{y-1}{r-1} (e^{tp})^r \times [e^{t(1-p)}]^{y-r}$$

$$q = 1-p$$

$$e^{tq}$$

$$1-z = e^{tq}$$

RECALL: Suppose that the real function $f(x)$ is infinitely differentiable at $x = a$. The Taylor series expansion of $f(x)$ about the point $x = a$ is given by

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n = f(a) + \left[\frac{f'(a)}{1!} \right] (x-a)^1 + \left[\frac{f''(a)}{2!} \right] (x-a)^2 + \dots$$

When $a = 0$, this is called the **McLaurin series expansion** of $f(x)$.

NEGATIVE BINOMIAL MGF: Suppose that $Y \sim \text{nb}(r, p)$. The mgf of Y is given by

$$m_Y(t) = \left(\frac{pe^t}{1-qe^t} \right)^r$$

where $q = 1-p$, for all $t < -\ln q$. Before we prove this, let's state and prove a lemma.

LEMMA. Suppose that r is a nonnegative integer. Then,

$$\sum_{y=r}^{\infty} \binom{y-1}{r-1} (qe^t)^{y-r} = (1-qe^t)^{-r}$$

Proof of lemma. Consider the function $f(w) = (1-w)^{-r}$, where r is a nonnegative integer. It is easy to show that

$$f'(w) = r(1-w)^{-(r+1)}$$

$$f''(w) = r(r+1)(1-w)^{-(r+2)}$$

$$\vdots$$

In general, $f^{(z)}(w) = r(r+1) \dots (r+z-1)(1-w)^{-(r+z)}$, where $f^{(z)}(w)$ denotes the z th derivative of f with respect to w . Note that

$$f^{(z)}(w) \Big|_{w=0} = r(r+1) \dots (r+z-1)$$

Now, consider writing the McLaurin Series expansion of $f(w)$; i.e., a Taylor Series expansion of $f(w)$ about $w = 0$; this expansion is given by

$$(1-w)^{-r} = f(w) = \sum_{z=0}^{\infty} \frac{f^{(z)}(0)w^z}{z!} = \sum_{z=0}^{\infty} \frac{r(r+1) \dots (r+z-1)}{z!} w^z = \sum_{z=0}^{\infty} \binom{z+r-1}{r-1} w^z$$

Letting $w = qe^t$ and $z = y-r$, the lemma is proven for $0 < q < 1$. \square

Now that we are finished with the lemma, let's find the mgf of $Y \sim \text{nb}(r, p)$. With $q = 1-p$, we have

$$m_Y(t) = E(e^{tY}) = \sum_{y=r}^{\infty} e^{ty} \binom{y-1}{r-1} p^r q^{y-r}$$

$$= \sum_{y=r}^{\infty} e^{t(y-r)} \binom{y-1}{r-1} p^r q^{y-r}$$

$$= (pe^t)^r \sum_{y=r}^{\infty} \binom{y-1}{r-1} (qe^t)^{y-r} = (pe^t)^r (1-qe^t)^{-r}. \square$$

REMARK: Showing that the $\text{nb}(r, p)$ pmf sums to one can be done by using a similar series expansion as above. We omit it for brevity.

MEAN AND VARIANCE: For $Y \sim \text{nb}(r, p)$, with $q = 1-p$,

$$E(Y) = \frac{r}{p} \text{ and } V(Y) = \frac{rq}{p^2}$$

3.9 Hypergeometric distribution

SETTING: Consider a collection of N objects (e.g., people, poker chips, plots of land, etc.) and suppose that we have two dichotomous classes, Class 1 and Class 2. For example, the objects and classes might be

Poker chips red/blue

1. definition of Neg Binom
2. $r \cdot p$
3. $P(Y=y), P(Y \leq y), P(r \leq Y \leq s)$
4. Mean, Variance
5. mgf

etc.) and suppose that we have two dichotomous classes, Class 1 and Class 2. For example, the objects and classes might be

Poker chips	red/blue
People	infected/not infected
Plots of land	respond to treatment/not.

From the collection of N objects, we sample n of them (without replacement), and record Y , the number of objects in Class 1.

REMARK: This sounds like a binomial setup! However, the difference here is that N , the **population size**, is finite (the population size, theoretically, is assumed to be infinite in the binomial model). Thus, if we sample from a population of objects **without replacement**, the "success" probability changes from trial to trial. This, violates the binomial

$$\begin{aligned}
 &= \sum_{y=r}^{\infty} \binom{y-1}{r-1} \underbrace{(e^{tq})^{y-r}}_{e^{-z}} \times \underbrace{(e^{tp})^r}_{e^{-z}} \quad \begin{matrix} e^{-z} \\ 1-z=e^{tq} \\ z=1-e^{tq} \end{matrix} \\
 &= \sum_{y=r}^{\infty} \binom{y-1}{r-1} (1-z)^{y-r} \times z^r \times \left(\frac{e^{tp}}{z}\right)^r \\
 &= \left(\frac{e^{tp}}{z}\right)^r \underbrace{\sum_{y=r}^{\infty} \binom{y-1}{r-1} (1-z)^{y-r} z^r}_1 \\
 &= \left(\frac{e^{tp}}{1-e^{tq}}\right)^r \quad \text{provided } 0 < z < 1 \\
 & \quad 0 < 1-e^{tq} < 1 \\
 & \quad 0 = e^{tq} < 1 \\
 & \quad t < -\ln q
 \end{aligned}$$