Section 3.8 Negative Binomial Distribution

Tuesday, September 27, 2016 12:53 PM

Section 3.8 Negative ...

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. □

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 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 rth 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

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

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

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

RATIONALE: The form of $p_Y(y)$ can be explained intuitively. If the rth success occurs on the yth trial, then $r-1$ successes must have occurred during the 1st $y-1$ 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 1st $y-1$ 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 yth trial, we observe the rth success (this occurs with probability p). Because the yth 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 ($p = 0.30$). 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?

"success" labels "infected"

SOLUTION. step 1: Let Y denote the tree on which she observes the 3rd infected tree. Then, step2: $Y \sim \text{nib}(r = 3, p = 0.3)$. We want to step3: compute $P(Y = 6 \text{ or } Y = 7)$. The nib(3, 0.3) pmf gives

→ the 3rd "success"

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

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

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

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

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

Geometric dist

1st success

_ _ _ _ _ _

$\underbrace{\qquad\qquad}_{\text{Geo}}$

Neg. Binomial

Y: # of trials to observe the rth success

If $r = 1$, Geometric.

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

$P(Y = r) = p^r$

$P(Y = r+1) = p \times P(\text{among } r \text{ } (r-1) \text{ successes})$

$= p \times \binom{r}{r-1} p^{r-1}(1-p)$ ← ??? , rth success on $(r+1)$th trial

_ _ _ _ _ _ _ _ _ _

among r trials there are $r-1$ successes p

r success

$P(Y = r+2) = p \times \binom{r+1}{r-1} p^{r-1}(1-p)^2$

rth ~~success~~ success

_ _ _ _ _ _ _

$r+1$ trials, $r-1$ successes $(r+2)$th trial

$y \geq r$

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

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

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

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

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

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

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

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.

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

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \\ &= f(a) + \left[\frac{f'(a)}{1!}\right](x-a)^1 + \left[\frac{f''(a)}{2!}\right](x-a)^2 + \cdots . \end{aligned}$$

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

NEGATIVE BINOMIAL MGF: Suppose that $Y \sim \text{nib}(r, p)$. 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

$$\begin{aligned} f'(w) &= r(1-w)^{-(r+1)} \\ f''(w) &= r(r+1)(1-w)^{-(r+2)} \\ &\vdots \end{aligned}$$

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

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

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

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

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

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

$$\begin{aligned} m_Y(t) = E(e^{tY}) &= \sum_{y=r}^{\infty} e^{ty}\binom{y-1}{r-1}p^r q^{y-r} \\ &= \sum_{y=r}^{\infty} e^{t(y-r)}e^{tr}\binom{y-1}{r-1}p^r q^{y-r} \\ &= (pe^t)^r \sum_{y=r}^{\infty}\binom{y-1}{r-1}(qe^t)^{y-r} = (pe^t)^r(1-qe^t)^{-r}. \ \square \end{aligned}$$

REMARK: Showing that the nib(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 $Y \sim \text{nib}(r, p)$, with $q = 1 - p$,

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

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

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

For $Y \sim$ Neg binomial(r, p)

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

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

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

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

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

equivalence

$\Leftrightarrow$ r th success occurs after the yth trial

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

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

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

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

Negative Binomial MGF:

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

$$= \sum_{y \in R} e^{ty} P_Y(y) \qquad e^{ty} = e^{tr} \times e^{t(y-r)}$$

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

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

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

$$= \sum_{y=r}^{\infty}\binom{y-1}{r-1}(e^t p)^r \times \underbrace{\left[e^t(1-p)\right]^{y-r}}_{e^t q} \qquad q = 1-p$$

$$1-z = e^t q$$

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

$e^{-}q$

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

$1-z = e^t q$

$z = 1-e^t q$

$$= \sum_{y=r}^{\infty} \binom{y-1}{r-1} (1-z)^{y-r} \times z^r \times \left(\frac{e^t p}{z}\right)^r$$

$$= \left(\frac{e^t p}{z}\right)^r \underbrace{\sum_{y=r}^{\infty} \binom{y-1}{r-1} (1-z)^{y-r} z^r}_{1}$$

$$= \left(\frac{e^t p}{1-e^t q}\right)^r$$

provided $0<z<1$

$0 < 1-e^t q < 1$

$0 < e^t q < 1$

$t < -\ln q$