Department of Mathematics
College of William & Mary

MATH 451/551

Chapter 4. Common Discrete Distributions

4.4 Negative Binomial Distribution

GuanNan Wang
gwang01@wm.edu

Negative Binomial Distribution

- The negative binomial distribution models the number of failures before the rth success in repeated, independent Bernoulli trials, each with probability of success p.
- **Support:** $\mathcal{A} = \{0, 1, 2, \ldots\}$

lb = 0
ub → ∞

PMF

The probability of r successes and x failures in a specified order, for example

FFFSSFFFS

FFSFFFFSS

3rd successes 5 fails

$x+r-1$

$r-1$

associated with $r = 3$ and $x = 5$, is

$p^r(1-p)^x$

$$p^r(1-p)^x.$$

There are $\binom{x+r-1}{r-1}$ different sequences of failures and successes associated with x failures prior to the rth success.

Negative Binomial Distribution

$X \sim$ NegBinomial (r, p)

- **PMF:** A discrete random variable X with PMF

$X \sim NB(r,p)$

$$f(x) = \binom{x+r-1}{r-1} p^r(1-p)^x, \qquad x = 0, 1, 2, \ldots$$

for some positive integer r and $0 < p < 1$ is a *Negative Binomial*(r, p) random variable.

Mean

$\mathcal{X} = \{0, 1, 2, \cdots\} \Rightarrow X$ is discrete

$$f(x) = \binom{x+r-1}{r-1} p^r (1-p)^x$$

$$E(X) = \sum_{x=0}^{\infty} x f(x) = \sum_{x=0}^{\infty} x \frac{(x+r-1)!}{(r-1)!\{(x+r-1)-(r-1)\}!} p^r (1-p)^x$$

$$= \sum_{x=1}^{\infty} \frac{\cancel{x}(x+r-1)!}{(r-1)!\,\cancel{x!}\,(x-1)!} p^r (1-p)^x = \sum_{x=1}^{\infty} \frac{(x+r-1)!\, r}{r!\,(x-1)!} p^{r+1} \frac{1-p}{p} (1-p)^{x-1}$$

let $\underline{\underline{y = x-1}}$, $m = r+1$

$$\underbrace{\sum_{y=0}^{\infty} \frac{(m+y-1)!}{(m-1)!\,y!} p^m (1-p)^y}_{1} \cdot \boxed{r \frac{1-p}{p}}$$

$$= r \frac{1-p}{p}$$

Variance

$$V(X) = \underbrace{E(X^2)} - \{E(X)\}^2 = E\{\underbrace{X(X-1)}_{?}\} + \underbrace{E(X)}_{\frac{r(1-p)}{p}} - \{\underbrace{E(X)}\}^2$$

$$\underbrace{E\{X(X-1)\}}_{g(x)} = \sum_{x=0}^{\infty} x(x-1)\frac{(x+r-1)!}{(r-1)!\,x!}p^r(1-p)^x$$

$$= \sum_{x=2}^{\infty} \cancel{x(x-1)}\frac{(x+r-1)!}{(r-1)!\,\cancel{x!}_{(x-2)!}}p^r(1-p)^x$$

$$= \sum_{x=2}^{\infty} \frac{(x+r-1)!\,r(r+1)}{(r+1)!\,(x-2)!}p^{r+2}\frac{1}{p^2}(1-p)^{x-2}(1-p)^2$$

let $\dfrac{y = x-2}{m = r+2}$ $\underbrace{\sum_{y=0}^{\infty}\frac{(m+y-1)!}{(m-1)!\,y!}p^m(1-p)^{x-2}}_{1} \times \boxed{r(r+1)\frac{(1-p)^2}{p^2}}$

$$= r(r+1)\frac{(1-p)^2}{p^2}$$

$$V(X) = \frac{r(r+1)(1-p)^2 + rp(1-p) - r^2(1-p)^2}{p^2} = \frac{r(1-p)\{(r+1)(1-p) + p \cancel{- r(1-p)}\}}{p^2} = \frac{r(1-p)}{p^2}$$

MGF

$X_1, X_2, \cdots, X_r.$

$Y = X_1 + X_2 + \cdots + X_r. \sim NB(r, p)$

$M_Y(t) = E(e^{tY}) = E\left(e^{t(X_1+X_2+\cdots+X_r)}\right)$

$= E\left(e^{tX_1} e^{tX_2} \cdots e^{tX_r}\right)$

$= \underbrace{E(e^{tX_1})}\ \underbrace{E(e^{tX_2})} \cdots \underbrace{E(e^{tX_r})}$

$= \left\{\dfrac{p}{1-(1-p)e^t}\right\}^r \qquad (1-p)e^t < 1$

Skewness

$$E\left\{\left(\frac{X-\mu}{\sigma}\right)^3\right\} = \frac{2-p}{\sqrt{r(1-p)}}$$

Kurtosis

$$E\left\{\left(\frac{X-\mu}{\sigma}\right)^4\right\} = \frac{p^2-6p-3pr+3r+6}{r(1-p)}$$

Example 1

Eric is making cold sales calls. The probability of a sale on each call is 0.4. The calls may be considered independent Bernoulli trials.

1. What is the probability that he has exactly five failed calls before his second successful sales call?
2. What is the probability that he has fewer than five failed calls before his second successful sales call?

let X = # failures before Eric's second success.

$X \sim NB(2, 0.4)$ $\quad f_X(x) = \binom{x+2-1}{2-1}(0.4)^2(0.6)^x, \quad x = 0,1,2\ldots$

① $P(X=5) = \binom{5+2-1}{2-1} 0.4^2 \, 0.6^5 \approx 0.07465$

② $P(X<5) = f(0) + f(1) + f(2) + f(3) + f(4) \approx 0.76672$

R Functions

Function	Returned Value
d~~geom~~nbinom(x,r,p)	calculates the probability mass function $f(x)$
p~~geom~~nbinom(x,r,p)	calculates the cumulative mass function $F(x)$
q~~geom~~nbinom(u,r,p)	calculates the percentile (quantile) $F^{-1}(u)$
r~~geom~~nbinom(m,r,p)	generates m random variates

d/p/q/r nbinom (? , r, p)

Alternative Definition

A negative binomial random variable X can also model the **trial number** of the rth success in a sequence of repeated, mutually independent, and identically distributed Bernoulli trials.

rth success

- **Support:** $\mathcal{A} = \{r, r+1, r+2, \ldots\}$
- **PMF:** A discrete random variable X with PMF

  # trial

  $$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \qquad x = r, r+1, \ldots$$

  for some positive integer r and $0 < p < 1$ is a *Negative Binomial*(r, p) random variable.
- **Population Mean:** $\mu = E(X) = \frac{r}{p}$.
- The population variance, skewness, and kurtosis remain the same.

Remark

- Alias: Pascal distribution
- Shorthand: $X \sim NB(r, p)$
- The geometric distribution is a special case of the negative binomial distribution when $r = 1$
- A negative binomial random variable can be thought of as the concatenation of r random experiments associated with the geometric distribution

Thank You

THANK YOU!