Math 3501 - Probability and Statistics I

2.6 - The negative binomial distribution

Negative binomial distribution

Observe a sequence of independent Bernoulli trials until exactly r successes occur, where r is a fixed positive integer.

Let X denote the number of trials needed to observe the rth success.

The pmf of X is

r successes x-r failures

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

we need to repeat the experiment at least r times to observe r successes

and we say that X has a *negative binomial distribution* (with parameters r and p).

ways in which r-1 successes may occur among 1st x-1 trials

$$g(x) = P(X = x) = \binom{x-1}{r-1} p^r \cdot (1-p)^{x-r}$$

we need x repetitions to observe r successes → r successes → x-r failures

on 1st x-1 trials, r-1 successes r th success

F F F S S ...... F S

x Bernoulli trials

$$\frac{(x-1)!}{(r-1)!(x-r)!} = \binom{x-1}{r-1}$$

Calculus II

Recall: Recall that the Binomial series, i.e., the Maclaurin's series expansion for the function

$$h(w) = (1-w)^{-r} = \frac{1}{(1-w)^r}\ , \qquad r \in \mathbb{R}\ ,$$

generalizes the binomial expansion

$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

n positive integer

n+1 terms

is given by

$$(1-w)^{-r} = \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} w^k = \sum_{k=0}^{\infty} \binom{r+k-1}{r-1} w^k, \quad |w| < 1\ ,$$

where

$$\binom{r+k-1}{r-1} = \frac{(r+k-1)(r+k-2)(r+k-3)\cdots r}{k!}$$

are (generalized) binomial coefficients.

Setting $x = k + r$ (or, equivalently, $k = x - r$), we obtain

series general term is very similar to the negative binomial pmf

$$(1-w)^{-r} = \sum_{x=r}^{\infty} \binom{x-1}{r-1} w^{x-r}, \quad |w| < 1\ .$$

alternative representation for Binomial series

Consequence: the pmf of the negative binomial distribution satisfies

Note that

$g(x) = \binom{x-1}{r-1} p^r \cdot (1-p)^{x-r} > 0$

for all $x = r, r+1, \ldots$

$$\sum_{x=r}^{\infty} g(x) = \sum_{x=r}^{\infty} \binom{x-1}{r-1} p^r q^{x-r}$$

$$= p^r \underbrace{\sum_{x=r}^{\infty} \binom{x-1}{r-1} q^{x-r}}_{} = \underbrace{\overbrace{p^r (1-q)^{-r}}^{p^r \cdot p^{-r}}}_{} = 1$$

Binomial series from previous slide with w replaced by q

$\Downarrow$

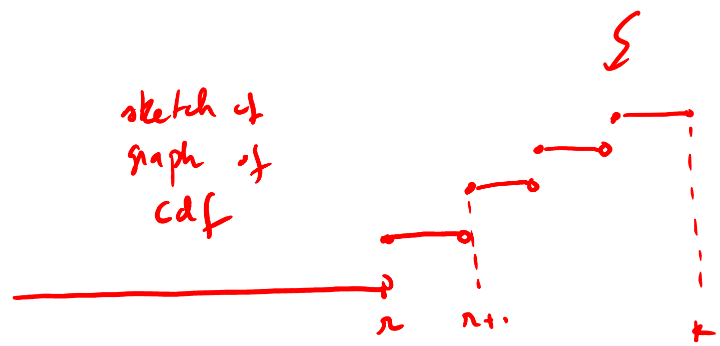
sum equals $(1-q)^{-r}$

$q = 1-p$

$1-q = p$

$g(x)$ is indeed a pmf !!

sketch of graph of cdf



Geometric distribution

→ Special case of negative binomial distribution with $r = 1$ } 1st success

If $r = 1$ in the negative binomial distribution, we say that X has a *geometric distribution*, with pmf given by

$$\boxed{g(x) = p(1-p)^{x-1}\ , \quad x = 1, 2, 3, \dots}$$

} X gives # of trials until 1st success is observed!

Note: The pmf of a geometric distribution consists of terms of a geometric series

Recall: the sum of a geometric series is given by

Calculus II

$$\sum_{k=0}^{\infty} ar^k = \sum_{k=1}^{\infty} ar^{k-1} = \frac{\overset{\text{1st term}}{a}}{1-\underset{\text{ratio}}{r}} \quad \text{provided } |r| < 1\ .$$

For the geometric distribution, we have

$$\sum_{x=1}^{\infty} g(x) = \sum_{x=1}^{\infty} (1-p)^{x-1} p = \frac{p}{1-(1-p)} = 1\ .$$

$\frac{p}{1-1+p}$

ratio is $q = (1-p) \in (0,1)$

Still from the sum of a geometric series, we get that whenever k is ~~an~~ positive integer

$$P(X > k) = \sum_{x=k+1}^{\infty} \overbrace{(1-p)^{x-1}}^{g(x)} p = \frac{(1-p)^k p}{\underbrace{1-(1-p)}_{p}} = (1-p)^k = q^k \ .$$

Need to know the process to sum geometric series

Thus, the value of the cdf of the geometric distribution at a positive integer k is

$$P(X \leq k) = 1 - P(X > k) = 1 - (1-p)^k = 1 - q^k \ .$$

cdf $F(x) = P(X \leq k) = \sum_{x=1}^{k} g(x) = \ldots$

$P(\bar{A}) = 1 - P(A)$

Example

Biology students were checking eye color in a large number of fruit flies. For an individual fly, suppose that the probability of white eyes is $1/4$ and the probability of red eyes is $3/4$, and that we may treat these observations as independent Bernoulli trials.

Find the probability that:

a) • at least four flies have to be checked for eye color to observe a white-eyed fly.

b) • the first fly with white eyes is the fourth fly.

Define the r.v. X to be the number of flies that need to be checked until the first fly with white eyes is observed

Recall that the probability of observing white eyes in one of the flies is $1/4$

Thus, X has a geometric distribution with prob of success $p = \frac{1}{4}$

that is, the pmf of X is $\underbrace{g(x) = \frac{1}{4} \cdot \left(\frac{3}{4}\right)^{x-1}}$, $x = 1, 2, 3, \ldots$

a) $\underbrace{P(X \geq 4)}_{\substack{1^{st} \text{ white-eyed fly} \\ \text{is observed after} \\ \text{at least } 4 \text{ trials}}} = \sum_{x=4}^{\infty} g(x) = \sum_{x=4}^{\infty} \frac{1}{4}\left(\frac{3}{4}\right)^{x-1} =$

geometric series with 1^{st} term equal to $\frac{1}{4}\left(\frac{3}{4}\right)^{4-1}$ and ratio $\frac{3}{4}$

$$= \frac{\frac{1}{4}\left(\frac{3}{4}\right)^{4-1}}{1 - \frac{3}{4}}$$

b) $P(X=4) = g(4) = \frac{1}{4} \cdot \left(\frac{3}{4}\right)^3$

$$= \frac{\frac{1}{4}\left(\frac{3}{4}\right)^3}{\frac{1}{4}} = \left(\frac{3}{4}\right)^3$$

additional questions

$$P(X \geq 4) = 1 - P(X < 4) = 1 - P(X \leq 3) = 1 - \sum_{x=1}^{3} \frac{1}{4}\left(\frac{3}{4}\right)^{x-1}$$

$$\underbrace{P(X \leq 6)} = 1 - P(X > 6) = 1 - P(X \geq 7) = 1 - \sum_{x=7}^{\infty} \frac{1}{4}\left(\frac{3}{4}\right)^{x-1}$$

$$= 1 - \frac{\frac{1}{4}\left(\frac{3}{4}\right)^6}{1 - \frac{3}{4}} = 1 - \left(\frac{3}{4}\right)^6 \checkmark$$

What if we really want to evaluate $P(X \le 6)$ as $\underbrace{\sum_{x=1}^{6} \frac{1}{4} \cdot \left(\frac{3}{4}\right)^{x-1}}_{\text{6 terms to sum ....}}$

$$\frac{1}{4} \cdot \frac{1 - \left(\frac{3}{4}\right)^6}{1 - \frac{3}{4}}$$

$$= 1 - \left(\frac{3}{4}\right)^6$$

OR the partial sum for a geometric series:

$$S = \sum_{k=1}^{N} a r^{k-1} \Rightarrow S = a + ar + ar^2 + ar^3 + \dots + ar^{N-2} + ar^{N-1}$$

$$\Rightarrow rS = ar + ar^2 + ar^3 + ar^4 + \dots + ar^{N-1} + ar^N$$

$$\Rightarrow S - rS = a - ar^N \Rightarrow (1-r)S = a - ar^N$$

named after Gauss

$$\Rightarrow S = \frac{a(1 - r^N)}{1 - r}$$

1st term

ratio

N = # terms

Moment generating function for negative binomial

Suppose X has a negative binomial distribution (with parameters r and p).

The mgf of X is given by

pmf of negative binomial

Recall that
$M(t) = E[e^{tx}]$

$$\begin{aligned} M(t) &= \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} p^r (1-p)^{x-r} \\ &= (pe^t)^r \sum_{x=r}^{\infty} \binom{x-1}{r-1} \left[(1-p)e^t\right]^{x-r} \\ &= \frac{(pe^t)^r}{\left[1-(1-p)e^t\right]^r}, \end{aligned}$$

$$e^{tx} = e^{t(x-r+r)} = e^{t(x-r)} \cdot e^{tr}$$

Binomial series

$$(1-w)^{-r} = \sum_{x=r}^{\infty} \binom{x-1}{r-1} w^{x-r}$$

with $w = (1-p)e^t$

where $(1-p)e^t < 1$ or, equivalently, $t < -\ln(1-p)$.

condition for convergence of binomial series

Negative binomial distribution mean and variance

Differentiating

$$M(t) = \frac{(pe^t)^r}{[1-(1-p)e^t]^r}$$

from previous slide

we obtain

use quotient rule + chain rule and simplify

$$M'(t) = r\left(pe^t\right)^r \left[\overbrace{1-(1-p)e^t}^{p}\right]^{-r-1}$$

product rule + chain rule and simplify

$$\begin{aligned} M''(t) &= r\left(pe^t\right)^r(-r-1)\left[1-(1-p)e^t\right]^{-r-2}\left[-(1-p)e^t\right] \\ &\quad + r^2\left(pe^t\right)^{r-1}\left(pe^t\right)\left[1-(1-p)e^t\right]^{-r-1} \end{aligned}$$

Evaluating at $t=0$, we find that

$$\mu = E(X) = M'(0) = \frac{r}{p}$$

$M'(0) = r \cdot p^r \cdot p^{-r-1} = \frac{r}{p}$

and

$$\sigma^2 = E\left(X^2\right) - [E(X)]^2 = M''(0) - [M'(0)]^2$$

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

$M''(0) = r\,p^r\,(r+1)\,p^{-r-2}\,(1-p) + r^2\,p^{r-1} \cdot p\,(p)^{-r-1}$

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

Conclusion: If X has a negative binomial distribution with parameters $r \in \mathbb{N}$ and $p \in (0, 1)$:

- $\mu = E[X] = \dfrac{r}{p}$ ✓
- $\sigma^2 = \mathrm{Var}(X) = \dfrac{rq}{p^2}$ ✓

Intuition: when p is the probability of success on each trial, the expected number of trials required to observe a single success is $1/p$.

obtained in earlier examples →

Special case: when $r = 1$, X has a geometric distribution and we have

- $M(t) = \dfrac{pe^t}{1 - (1 - p)e^t}$
- $\mu = E[X] = \dfrac{1}{p}$
- $\sigma^2 = \mathrm{Var}(X) = \dfrac{q}{p^2}$

← set $r = 1$ in the formulas for the negative binomial!

Example

Suppose that during practice a basketball player can make a free throw 80% of the time. Furthermore, assume that a sequence of free-throw shooting can be thought of as independent Bernoulli trials.

Let X equal the minimum number of free throws that this player must attempt to make a total of 10 shots.

Determine the mean and variance of X.

first 10 successes

X = r.v. counting the # of trials until making $\underbrace{\text{10 shots}}$

"free-throw shots"

success is a shot
and probability of success is $p=0.8$

We know that X has a negative binomial distribution with parameters:

$r = 10$ ⟵ X counts # of trials until 10 successes are observed

$p = 0.8$

mean is $\mu = E[X] = \frac{r}{p} = \frac{10}{0.8}$

Variance is $\sigma^2 = Var(X) = \frac{r\,q}{p^2} = \frac{10 \cdot (1-0.8)}{(0.8)^2} = \frac{2}{(0.8)^2}$

Example

A fair six-sided die is rolled until each face is observed at least once.

On the average, how many rolls of the die are needed?

↳ to observe each face at least once!

Roll die once ⟶ observe the first face (whatever it is) in 1 roll.

To observe the next distinct face : probability of success is $\frac{5}{6}$ ← one face already observed, 5 yet to be observed

times that we need to roll die until observing one of the 5 missing faces is a geometric r.v with $p = \frac{5}{6}$ $\Rightarrow$ mean is $\frac{1}{p} = \frac{6}{5}$

To observe the 3^{rd} distinct face : probability of success is $\frac{4}{6}$ } 2 faces were already observed

$\Rightarrow$ geometric r.v. with $p = \frac{4}{6}$ $\Rightarrow$ mean is $\frac{1}{p} = \frac{6}{4}$

4^{th} distinct face: prob of success is $\frac{3}{6}$

$\Rightarrow$ geometric r.v. with $p = \frac{3}{6}$ $\Rightarrow$ mean is $\frac{6}{3}$

5^{th} distinct face: prob. of success is $\frac{2}{6}$

$\Rightarrow$ geometric r.v. with $p = \frac{2}{6}$ $\Rightarrow$ mean is $\frac{6}{2}$

Finally, for the last distinct face: prob success is $\frac{1}{6}$

$\Rightarrow$ geometric r.v with $p = \frac{1}{6}$ $\Rightarrow$ mean is $\frac{6}{1}$

In total we need roll the die, on average, $1 + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + 6 = \frac{147}{10}$

(nearly 15 rolls to see all faces)