

§ 5.3 Negative binomial distribution

Example Consider the random experiment of repeatedly tossing a coin with heads probability $p \in (0,1)$ until getting r heads, for a given integer $r \geq 1$. Let $\bar{X}$ count the number of tosses made (up to and including the toss that gave the r th head). Find the ① range of $\bar{X}$ and ② pmf of $\bar{X}$.

① range is $\{r, r+1, r+2, \dots\}$

$$\textcircled{2} \quad P(\bar{X}=k) = \underbrace{\binom{k-1}{r-1}}_{\text{number of ways to distribute the first } r-1 \text{ heads in first } k-1 \text{ tosses.}} (1-p)^{k-r} p^r \quad \text{for } k \geq r.$$

Def A random variable $\bar{X}$ has the negative binomial distribution with parameters $r \geq 1$ and $p \in (0,1)$ if its range is $\{r, r+1, r+2, \dots\}$ and its probability mass function is given by $P(\bar{X}=k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$ for $k \geq r$.

Shorthand: $\bar{X} \sim \text{NegBin}(r, p)$

R commands: if $\bar{X} \sim \text{NegBin}(r, p)$

$$P(\bar{X}=k) \quad \text{dnbinom}(k-r, r, p)$$

$$P(\bar{X} \leq k) \quad \text{pnbinom}(k-r, r, p)$$

Remark: ① counts independent trials until r -th success

② if $\bar{X} \sim \text{Neg Bin}(r, p)$, then $\bar{X} = G_1 + \dots + G_r$
where $G_1, \dots, G_r \sim \text{Geom}(p)$ are i.i.d.

Example Let $\bar{X} \sim \text{Neg Bin}(r, p)$. Show $E[\bar{X}] = \frac{r}{p}$.

Note that $\bar{X} = G_1 + G_2 + \dots + G_r$ where

$G_1, G_2, \dots, G_r \sim \text{Geom}(p)$ are i.i.d., with G_i

counting the trials after the $(i-1)$ st success to the i th success. Thus

$$\begin{aligned} E[\bar{X}] &= E[G_1 + \dots + G_r] \\ &= E[G_1] + \dots + E[G_r] \\ &= \frac{r}{p}. \end{aligned}$$

Discussion Why the name "negative binomial"?

Let $\bar{X} \sim \text{Neg Bin}(r, p)$. Consider the event $\{\bar{X} > n\}$.

This event is equivalent to having fewer than r successes in the first n trials of the experiment. So if

$\bar{Y} \sim \text{Bin}(n, p)$, then $\{\bar{X} > n\} = \{\bar{Y} < r\}$. In this

sense the negative binomial and binomial distributions are "inverses" of each other.

Problem 1. Suppose applicants for a job are hired with probability $p = 0.15$ independently from person to person, one at a time as they come in.

- Find the probability of each of the following events by first expressing the event in terms of a negative binomially distributed random variable.
 - It took exactly 10 applicants until 3 people were hired.
 - It took at least 10 applicants until 3 people were hired.
 - It took exactly 15 applicants until 7 people were hired.
 - It took at least 15 applicants until 7 people were hired.
- Express each of the following events in terms of a binomially distributed random variable.
 - It took exactly 15 applicants until 7 people were hired.
 - It took at least 15 applicants until 7 people were hired.
- Find the expected value of the number of applicants seen until:
 - 3 hires are made.
 - 7 hires are made.

① $\bar{Y} \sim \text{Neg Bin}(3, 0.15),$
 $Z \sim \text{Neg Bin}(7, 0.15)$

① $P(\bar{Y} = 10)$

② $P(\bar{Y} \geq 10) = 1 - P(\bar{Y} \leq 9)$

③ $P(Z = 15)$

④ $P(Z \geq 15) = 1 - P(Z \leq 14)$

```

{r}
dnbinom(10-3, 3, 0.15) # part a1
1-pnbinom(9-3, 3, 0.15) # part a2
dnbinom(15-7, 7, 0.15) # part a3
1-pnbinom(14-7, 7, 0.15) # part a4
pbinom(2, 9, 0.15) # part b1
pbinom(6, 14, 0.15) # part b2

```

```

[1] 0.03895012
[1] 0.8591466
[1] 0.001398124
[1] 0.9977925
[1] 0.8591466
[1] 0.9977925

```

⑥ Let $B_1 \sim \text{Bin}(9, 0.15), B_2 \sim \text{Bin}(14, 0.15)$

① $\{B_1 \leq 2\}$

② $\{B_2 \leq 6\}$

⑦ ① $E[\bar{Y}] = \frac{3}{0.15}$

② $E[Z] = \frac{7}{0.15}$

Problem 2. Alice and Bob are playing a game. In each round of the game, one of them wins the round: Alice wins the round with probability $1/3$ and Bob wins with probability $2/3$. Whoever wins 50 rounds first wins the overall game. However, the game has to end prematurely before either of them gets to 50 wins. Alice needed to win 5 more rounds and Bob had to win 10 more rounds to reach 50.

- What is the maximum number of rounds they could have played before declaring an overall winner of the game?
- Find the probability that
 - Alice would have won 5 more rounds in the maximum number of possible rounds or fewer.
 - Bob would have won 10 more rounds in the maximum number of possible rounds or fewer.
- Who is more likely to have won if the game had not been ended prematurely?

Remark. This is an example of what is commonly called the *problem of points*. The original formulation of the problem was solved jointly by Blaise Pascal (1623-1662) and Pierre de Fermat (1601-1665) in a series of letters

$$(a) \quad 5 + 10 - 1 = 14$$

$$(b) \quad \bar{X} \sim \text{NegBin}(5, 1/3), \quad \bar{Y} \sim \text{NegBin}(10, 2/3)$$

$$(1) \quad P(\bar{X} \leq 14) = \sum_{k=5}^{14} \binom{k-1}{4} \left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^{k-5}$$

$$(2) \quad P(\bar{Y} \leq 14) = \sum_{k=10}^{14} \binom{k-1}{9} \left(\frac{2}{3}\right)^{10} \left(\frac{1}{3}\right)^{k-10}$$

$$(c) \quad \text{Alice}$$

```

{r}
pnbinom(14-5,5,1/3)
pnbinom(14-10,10,2/3)

```

```

[1] 0.5244995
[1] 0.4755005

```