

§ 5.1, 5.3 Geometric and negative binomial distributions

Example Consider the random experiment of repeatedly tossing a coin with heads probability $p \in (0, 1)$ until the first time you get heads. Let \bar{X} count the number of tosses made (including the toss that gave the first heads). Find the ① range of \bar{X} and ② the pmf of \bar{X} .

① range is $\{1, 2, 3, \dots\}$

$$\textcircled{2} \quad P(\bar{X}=k) = (1-p)^{k-1} p$$

Def A random variable \bar{X} has the geometric distribution with parameter $p \in (0, 1)$ if its range is $\{1, 2, 3, \dots\}$ and its probability mass function is $P(\bar{X}=k) = (1-p)^{k-1} p$ for $k \geq 1$.

Shorthand: $\bar{X} \sim \text{Geom}(p)$.

Remark Geometric distribution models number of trials up to and including first success in repeated independent trials of a success/failure experiment.

R commands: if $\bar{X} \sim \text{Geom}(p)$

$$P(\bar{X}=k) \quad \text{dgeom}(k-1, p)$$

$$P(\bar{X} \leq k) \quad \text{pgem}(k-1, p)$$

Example Let $\bar{X} \sim \text{Geom}(p)$. Show $E[\bar{X}] = \frac{1}{p}$.

First, recall that $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$ when $|x| < 1$

Therefore, $\frac{1}{(1-x)^2} = \frac{d}{dx} \left(\frac{1}{1-x} \right) = \sum_{k=1}^{\infty} kx^{k-1}$ when $|x| < 1$.

Now, we see that

$$\begin{aligned} E[\bar{X}] &= \sum_{k=1}^{\infty} k \cdot P(\bar{X}=k) \\ &= \sum_{k=1}^{\infty} k (1-p)^{k-1} p \\ &= p \sum_{k=1}^{\infty} k (1-p)^{k-1} \\ &= p \frac{1}{(1-(1-p))^2} \\ &= \frac{1}{p} \end{aligned}$$

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 $\{r, r+1, r+2, \dots\}$

② $P(\bar{X}=k) = \underbrace{\binom{k-1}{r-1}}_{\substack{\text{number of ways to distribute} \\ \text{the first } r-1 \text{ heads in first } k-1 \text{ tosses.}}} (1-p)^{k-r} p^r$ 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: $\bar{X} \sim \text{NegBin}(r, p)$

Remarks: ① counts independent trials until r th success
 ② if $\bar{X} \sim \text{NegBin}(r, p)$, then $\bar{X} = G_1 + \dots + G_r$
 where $G_1, \dots, G_r \sim \text{Geom}(p)$ are i.i.d.

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

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

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

Example Let $\bar{X} \sim \text{NegBin}(r, p)$. 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

$$E[\bar{X}] = E[G_1 + \dots + G_r]$$

$$= E[G_1] + \dots + E[G_r]$$

$$= \frac{r}{p}.$$

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. Unless otherwise stated, use geometric or negative binomial random variables to do the following.

- Find the probability of each of the following events.
 - It took exactly 3 applicant until the first person was hired.
 - It took at least 3 applicants until the first person was hired.
 - 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.
 - Repeat parts 4 and 6. using a binomially distributed random variable.
- Find the expected value of the number of applicants seen until
 - a hire is made.
 - 3 hires are made.
 - 7 hires are made.

① Let $\bar{X} \sim \text{Geom}(0.15)$, $\bar{Y} \sim \text{NegBin}(3, 0.15)$,

$$Z \sim \text{NegBin}(7, 0.15)$$

$$\textcircled{1} \quad P(\bar{X} = 3)$$

$$\textcircled{2} \quad P(\bar{X} \geq 3) = 1 - P(\bar{X} \leq 2)$$

$$\textcircled{3} \quad P(\bar{Y} = 10)$$

$$\textcircled{4} \quad P(\bar{Y} \geq 10) = 1 - P(\bar{Y} \leq 9)$$

$$\textcircled{5} \quad P(Z = 15)$$

$$\textcircled{6} \quad P(Z \geq 15) = 1 - P(Z \leq 14)$$

$$\textcircled{7} \quad \text{Let } B_1 \sim \text{Bin}(9, 0.15), B_2 \sim \text{Bin}(14, 0.15)$$

$$P(\bar{Y} \geq 10) = P(B_1 \leq 2)$$

$$P(Z \geq 15) = P(B_2 \leq 6)$$

Problem 1

Part a:

```
```{r}
dgeom(3-1, 0.15) # part 1
1-pgeom(2-1, 0.15) # part 2
dnbinom(10-3, 3, 0.15) # part 3
1-pnbinom(9-3, 3, 0.15) # part 4
dnbinom(15-7, 7, 0.15) # part 5
1-pnbinom(14-7, 7, 0.15) # part 6
pbinom(2, 9, 0.15) # part 7
pbinom(6, 14, 0.15) # part 8
...```

```

```
[1] 0.108375
[1] 0.7225
[1] 0.03895012
[1] 0.8591466
[1] 0.001398124
[1] 0.9977925
[1] 0.8591466
[1] 0.9977925
```

$$\textcircled{8} \quad \textcircled{1} \quad E[\bar{X}] = \frac{1}{0.15}$$

$$\textcircled{2} \quad E[\bar{Y}] = \frac{3}{0.15}$$

$$\textcircled{3} \quad E[Z] = \frac{7}{0.15}$$