UNIT 12

GEOMETRIC AND NEGATIVE BINOMIAL DISTRIBUTIONS

Structure

12.1 INTRODUCTION

We know that support of a discrete random variable either may be a finite set or may be a countably infinite set of values. In Units 9 and 10 we studied four probability distributions namely (i) discrete uniform (ii) Bernoulli (iii) Binomial and (iv) Multinomial distributions each of which has a finite support. In Unit 11, we studied Poisson distribution which has countably infinite support. Another two discrete probability distributions which have countably infinite support are geometric and negative binomial distributions which will be discussed in this unit. In Sec. 12.2, we will discuss PMF and CDF of geometric distribution while in Sec. 12.3, we will discuss MGF and some other summary measures of this probability distribution.

In Secs. 12.4 and 12.5, we will do similar studies about negative binomial distribution. Some applications of these distributions are discussed in Sec. 12.6.

What we have discussed in this unit is summarised in Sec. 12.7. Self-Assessment Questions (SAQs) have been given in some sections which are generally based on the content discussed in that section. But to give you a good practice of what we have discussed in this unit some more questions based on the entire unit are given in Sec. 12.8 under the heading Terminal Questions. Due to the reason mentioned in Sec. 1.1 of Unit 1 of this course, solutions of all the SAQs and Terminal Questions are given in Sec. 12.9.

In the next unit, you will do a similar study about two continuous probability distributions known as continuous uniform and exponential distributions.

Expected Learning Outcomes

After completing this unit, you should be able to:

- ❖ identify situations where you can apply geometric and negative binomial distributions;
- ❖ define PMF, CDF, MGF and some summary measures of geometric and negative binomial distributions; and
- ❖ apply geometric and negative binomial distributions to solve problems based on these two probability distributions.

12.2 STORY, DEFINITION, PMF AND CDF OF GEOMETRIC DISTRIBUTION

In Sec. 10.2 of the Unit 10, you have studied binomial distribution. Recall that in a binomial distribution, we have

- The number of trials ‘n’ is finite in numbers. … (12.1)
- All the ‘n’ trials are independent. … (12.2)
- Probability of success p remains constant in each trial. That is, p does not change from trial to trial. … (12.3)

Recall that in binomial distribution, we perform the experiment n times (**a fixed number**) and are interested in the probability of getting x (**variable**) successes out of n trials where x = 0, 1, 2, 3, …, n. But if we perform the experiment till we get the first success then, we can define two random variables as follows.

First Random Variable: Let X be the random variable which counts the number of failures before the first success.

Second Random Variable: Let Y be the random variable which counts the number of trials till the first success including the first success.

The probability distributions of both the random variables X and Y are known as geometric distribution. We call it geometric because successive probabilities form geometric progression refer to (12.7). We will proceed with the first random variable.

We are following the same notations as we used in the binomial distribution. So, probability of success and failure will be denoted by p and q = 1 – p. Suppose we get the first success in x^{th} trial then all the first x – 1 trials will be failure. Since trials are independent refer to (12.2) so, probability of getting the first success in x^{th} trial is given by

$$\mathcal{P}(X = x) = \underbrace{qqq\ldots q}_{x-\text{ times}}p = q^x p = (1-p)^x p \quad \ldots (12.4)$$

But the first success may occur in the first trial or second trial or third trial, and so on. If we get the first success in the first trial then the number of failures before the first success will be 0. If we get the first success in the second trial

then the number of failure before the first success will be 1. If we get the first success in the third trial then the number of failures before the first success will be 2, and so on. So, values of x in (12.4) may be 0 or 1 or 2 or 3 ….
Hence, finally (12.4) can be written as

$$\mathcal{P}\left(X=x\right)=\underbrace{qqq\ldots q}_{x-\text{times}}p=q^{x}p=(1-p)^{x}p,\ \ x=0,1,2,3,4,\ldots \qquad \ldots (12.5)$$

Now, we can discuss the story of the geometric distribution.

Story of Geometric Distribution: If we perform Bernoulli trials till, we get the first success then both the random variables which count the number of:

(i) failures before the first success; and

(ii) trials till the first success including the first success

forms a probability distribution known as a geometric distribution.

Definition and PMF of Geometric Distribution: If we perform Bernoulli trials till, we get the first success and the random variable X counts the number of failures before the first success then the PMF of X is given by

$$p_X(x)=\mathcal{P}\left(X=x\right)=\begin{cases}(1-p)^{x}p, & x=0,1,2,3,\ldots\\ 0, & \text{otherwise}\end{cases} \qquad \ldots (12.6)$$

Note that successive probabilities of the PMF given by (12.6) are $p,\ (1-p)p,\ (1-p)^2p,\ (1-p)^3p,\ \ldots$ which form a geometric progression (GP) with the first term p and the common ratio 1 – p because of this reason PMF given by (12.6) is known as PMF of a geometric distribution with single parameter p and is denoted by writing X ~ Geom(p). … (12.7)

Like other distributions, we read $X \sim \text{Geom}(p)$ as X is distributed as a geometric distribution with parameter p. Or we read it as X follows a geometric distribution with parameter p. … (12.8)

Since the statistical software used for hands-on practice for the lab courses of this programme is R programming language, so, you should also know its notation in R. In R notation that is used for geometric distribution is geom(prob) in the stats package, where prob represents the value of the parameter p of the geometric distribution. In fact, there are four functions for geometric distribution namely dgeom(x, prob, …), pgeom(q, prob, …), qgeom(p, prob, …), and rgeom(n, prob, …). We have already explained the meaning of these functions in Unit 9. … (12.9)

Let us check the **validity of the PMF of the Geometric distribution**.

We have to check two conditions non-negativity and normality refer to (4.62).

(1) **Non-negativity**: Since $0<p\le 1 \Rightarrow (1-p)^{x}p\ge 0\ \ \forall\ x=0,1,2,3,\ldots$ hence

$$p_X(x)=\mathcal{P}\left(X=x\right)\ge 0 \qquad \ldots (12.10)$$

(2) **Normality**: Let us obtain the sum of all probabilities of geometric distribution.

$$\sum_{x=0}^{\infty}(1-p)^x p = \sum_{x=0}^{\infty} q^x p = p\sum_{x=0}^{\infty} q^x = p\left(q^0 + q^1 + q^2 + q^3 + \ldots\right)$$

$$= p\left(1 + q + q^2 + q^3 + \ldots\right)$$

$$= p\frac{1}{1-q}\left[\begin{array}{l}\because \text{Sum of infinite GP } a + ar + ar^2 + ar^3 + \ldots \\ = \dfrac{a}{1-r}, \text{ where } |r| < 1\end{array}\right]$$

$$= \frac{p}{p} = 1 \quad [\because 1 - q = p] \qquad \ldots (12.11)$$

This proves that sum of all probabilities of geometric distribution is 1. So, we can say that PMF of the geometric random variable is a valid PMF.

Now, we define the CDF of a geometric distribution.

CDF of Geometric Distribution: If X ~ Geom(p) then PMF of X is given by

$$p_X(x) = \mathcal{P}\left(X = x\right) = \begin{cases}(1-p)^x p, & x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise}\end{cases}$$

Let x be any fixed real number then CDF of the random variable X is given by

$$F_X(x) = \mathcal{P}\left(X \le x\right) = \begin{cases}0, & \text{if } x < 0 \\ \displaystyle\sum_{k=0}^{[x]}(1-p)^x p = p\frac{1-(1-p)^{[x]+1}}{1-(1-p)} = 1 - q^{[x]+1}, & \text{if } x \ge 0\end{cases}$$

$$\left[\begin{array}{l}\because \text{Sum of n terms of a GP } a + ar + ar^2 + ar^3 + \ldots = \dfrac{a(1-r^n)}{1-r}, \text{ where } |r| < 1 \\ \text{Here, } a = p,\ r = 1 - p \text{ and } n = [x] + 1\end{array}\right]$$

or

$$F_X(x) = \mathcal{P}\left(X \le x\right) = \begin{cases}0, & \text{if } x < 0 \\ 1-(1-p)^{[x]+1}, & \text{if } x \ge 0\end{cases} \qquad \ldots (12.12)$$

Let us do one example.

Example 1: A biased coin which has the probability of getting a head as 1/4 and a tail as 3/4 is tossed till, we get the first head. If X counts the number of failures before the first head, then obtain the PMF and CDF of X and also plot the PMF and CDF of X.

Solution: If we define getting a head as a success and getting a tail as a failure then the random variable X which counts the number of failures before the first head follows a geometric distribution with parameter p = 1/4. So, the PMF of X is given by

$$p_X(x) = \mathcal{P}\left(X = x\right) = \begin{cases}\left(\dfrac{3}{4}\right)^x\left(\dfrac{1}{4}\right), & x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise}\end{cases} \qquad \ldots (12.13)$$

and the CDF of X is given by

$$F_X(x) = \mathcal{P}\left(X \le x\right) = \begin{cases}0, & \text{if } x < 0 \\ 1-\left(\dfrac{3}{4}\right)^{[x]+1}, & \text{if } x \ge 0\end{cases} \qquad \ldots (12.14)$$

Probabilities for X = 0, 1, 2, 3, …, 10 are given by

$$\mathcal{P}(X=0)=\left(\frac{3}{4}\right)^0\left(\frac{1}{4}\right)=0.25 \qquad \mathcal{P}(X=6)=\left(\frac{3}{4}\right)^6\left(\frac{1}{4}\right)=0.04449463$$

$$\mathcal{P}(X=1)=\left(\frac{3}{4}\right)^1\left(\frac{1}{4}\right)=0.1875 \qquad \mathcal{P}(X=7)=\left(\frac{3}{4}\right)^7\left(\frac{1}{4}\right)=0.03337097$$

$$\mathcal{P}(X=2)=\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)=0.140625 \qquad \mathcal{P}(X=8)=\left(\frac{3}{4}\right)^8\left(\frac{1}{4}\right)=0.02502823$$

$$\mathcal{P}(X=3)=\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)=0.1054688 \qquad \mathcal{P}(X=9)=\left(\frac{3}{4}\right)^9\left(\frac{1}{4}\right)=0.01877117$$

$$\mathcal{P}(X=4)=\left(\frac{3}{4}\right)^4\left(\frac{1}{4}\right)=0.07910156 \qquad \mathcal{P}(X=10)=\left(\frac{3}{4}\right)^{10}\left(\frac{1}{4}\right)=0.01407838$$

$$\mathcal{P}(X=5)=\left(\frac{3}{4}\right)^5\left(\frac{1}{4}\right)=0.05932617 \qquad \mathcal{P}(X=11)=\left(\frac{3}{4}\right)^{11}\left(\frac{1}{4}\right)=0.01055878$$

Now, using (12.14) CDF of X is given by

$$F_X(x)=\mathcal{P}(X\le x)=\begin{cases} 0, & \text{if } x<0 \\ 0.25, & \text{if } 0\le x<1 \\ 0.4375, & \text{if } 1\le x<2 \\ 0.578125, & \text{if } 2\le x<3 \\ 0.6835938, & \text{if } 3\le x<4 \\ 0.7626953, & \text{if } 4\le x<5 \\ 0.8220215, & \text{if } 5\le x<6 \\ 0.8665161, & \text{if } 6\le x<7 \\ 0.8998871, & \text{if } 7\le x<8 \\ 0.9249153, & \text{if } 8\le x<9 \\ 0.9436865, & \text{if } 9\le x<10 \\ 0.9577649, & \text{if } 10\le x<11 \\ \vdots & \end{cases} \qquad \ldots (12.15)$$

PMF and CDF are plotted in Fig. 12.1 (a) and (b) respectively as follows.

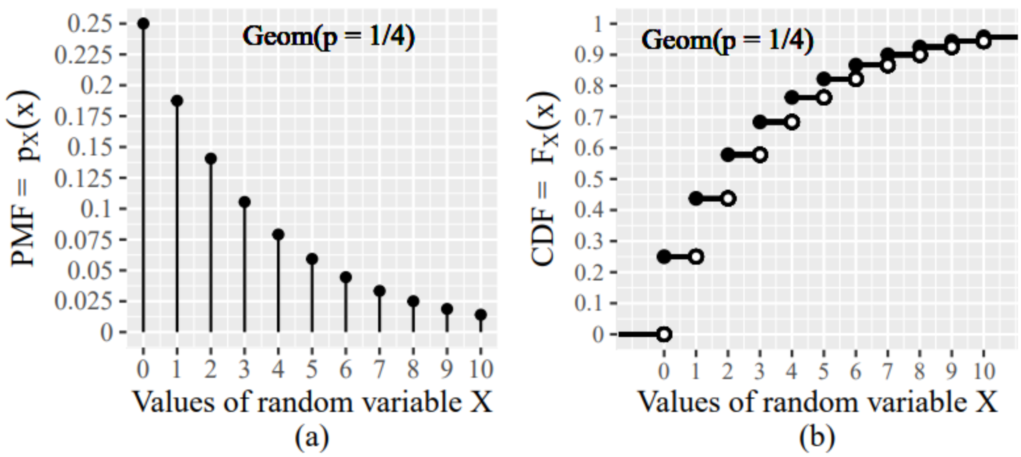


Fig. 12.1: Visualisation of (a) PMF (b) CDF of the Geom(1/4) discussed in Example 1

12.3 MGF AND OTHER SUMMARY MEASURES OF GEOMETRIC DISTRIBUTION

In the previous section, you have studied PMF and CDF of geometric distribution. In this section, we want to obtain MGF and some other summary measure of geometric distribution like mean and variance. Let us first obtain MGF of geometric distribution. We will obtain the MGF of geometric distribution using the definition of MGF, you may refer to (7.48).

Calculation of MGF

$$M_X(t) = E\left(e^{tX}\right),\ t \in \mathbb{R}$$

$$= \sum_{x=0}^{\infty} e^{tx} q^x p = p\sum_{x=0}^{\infty}\left(qe^t\right)^x = p\left[\left(qe^t\right)^0 + \left(qe^t\right)^1 + \left(qe^t\right)^2 + \left(qe^t\right)^3 + \ldots\right]$$

$$= p\left[1 + qe^t + \left(qe^t\right)^2 + \left(qe^t\right)^3 + \ldots\right]$$

$$= p\left(\frac{1}{1-qe^t}\right) \quad \left[\begin{array}{l}\because \text{ sum of infinite GP } a + ar + ar^2 + \ldots = \dfrac{a}{1-r} \\ \text{Provided } \left|qe^t\right| < 1 \Rightarrow \left|(1-p)e^t\right| < 1 \Rightarrow t < \ln\left(\dfrac{1}{1-p}\right)\end{array}\right]$$

$$= \frac{p}{1-(1-p)e^t},\quad t < \ln\left(\frac{1}{1-p}\right)$$

$$M_X(t) = \frac{p}{1-(1-p)e^t},\ t < \ln\left(\frac{1}{1-p}\right) \qquad \ldots (12.16)$$

Calculation for Mean

$$\text{Expected value} = E(X) = \sum_{x=0}^{\infty} x\,q^x\,p = p\sum_{x=1}^{\infty} x\,q^x \left[\begin{array}{l}\because \text{ when } x = 0, \text{ then we get} \\ \text{value of } x\,q^x\,p = 0\end{array}\right]$$

$$= p\sum_{x=1}^{\infty} x\,q\,q^{x-1} = pq\sum_{x=1}^{\infty} x\,q^{x-1} = pq\sum_{x=1}^{\infty}\frac{d}{dq}\left(q^x\right) \quad \left[\because \frac{d}{dx}\left(q^x\right) = x\,q^{x-1}\right]$$

$$= pq\frac{d}{dq}\left(\sum_{x=1}^{\infty} q^x\right) \quad [\because \text{ summation respect differentiation}]$$

$$= pq\frac{d}{dq}\left(q + q^2 + q^3 + q^4 + \ldots\right)$$

$$= pq\frac{d}{dq}\left(\frac{q}{1-q}\right) \quad \left[\begin{array}{l}\because \text{ sum of infinite GP,} \\ a + ar + ar^2 + ar^3 + \ldots = \dfrac{a\left(1-r^n\right)}{1-r}\end{array}\right]$$

$$= pq\left[\frac{(1-q)(1) - q(-1)}{(1-q)^2}\right] = pq\left[\frac{1}{p^2}\right] \quad [\because 1-q = p]$$

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

Hence, mean of X = expected value of $X = E(X) = \dfrac{1-p}{p}$. … (12.17)

Calculation for Variance and Standard Deviation

$$E\left(X(X-1)\right)=\sum_{x=0}^{\infty}x(x-1)q^x p = p\sum_{x=2}^{\infty}x(x-1)q^2q^{x-2}\left[\begin{array}{l}\because \text{ when } x=0,1 \text{ then}\\ \text{value of } x(x-1)q^x p=0\end{array}\right]$$

$$=pq^2\sum_{x=2}^{\infty}x(x-1)q^{x-2}=pq^2\sum_{x=2}^{\infty}\frac{d^2}{dq^2}\left(q^x\right)=pq^2\frac{d^2}{dq^2}\left(\sum_{x=2}^{\infty}q^x\right)$$

$$=pq^2\frac{d^2}{dq^2}\left(q^2+q^3+q^4+q^5+\ldots\right)=pq^2\frac{d^2}{dq^2}\left(\frac{q^2}{1-q}\right)$$

$$=pq^2\frac{d}{dq}\left(\frac{(1-q)(2q)-q^2(-1)}{(1-q)^2}\right)$$

$$=pq^2\frac{d}{dq}\left(\frac{2q-q^2}{(1-q)^2}\right)=pq^2\left(\frac{(1-q)^2(2-2q)-\left(2q-q^2\right)2(1-q)(-1)}{(1-q)^4}\right)$$

$$=pq^2\left(\frac{(1-q)(2-2q)+2\left(2q-q^2\right)}{(1-q)^3}\right)$$

$$=pq^2\left(\frac{2-2q-2q+2q^2+4q-2q^2}{p^3}\right)=q^2\left(\frac{2}{p^2}\right)=\frac{2q^2}{p^2}$$

$$E\left(X(X-1)\right)=\frac{2q^2}{p^2} \qquad \ldots (12.18)$$

$$E\left(X^2\right)=E\left(X(X-1)+X\right)=E\left(X(X-1)\right)+E\left(X\right)\quad\left[\begin{array}{l}\text{Using addition theorem of}\\ \text{expectation refer to (7.29)}\end{array}\right]$$

$$=\frac{2q^2}{p^2}+\frac{q}{p}\qquad\left[\text{Using (12.18) and (12.17)}\right]$$

$$\Rightarrow E\left(X^2\right)=\frac{2q^2}{p^2}+\frac{q}{p} \qquad \ldots (12.19)$$

Using (7.63) variance of any random variable X is given by

$$\text{Variance of the geometric distribution}=\mu_2=E\left(X^2\right)-\left(E(X)\right)^2$$

$$=\frac{2q^2}{p^2}+\frac{q}{p}-\left(\frac{q}{p}\right)^2\quad\left[\begin{array}{l}\text{Using (12.19)}\\ \text{and (12.17)}\end{array}\right]$$

$$=\frac{q^2}{p^2}+\frac{q}{p}=\frac{q^2+pq}{p^2}=\frac{q(q+p)}{p^2}$$

$$=\frac{q}{p^2}=\frac{1-p}{p^2}\qquad\left[\because p+q=1\right]$$

Hence, variance of the geometric distribution $=\dfrac{1-p}{p^2}$. … (12.20)

We know that standard deviation of X is positive square root of variance of X.

Hence, $SD(X)=\sqrt{\text{Variance of } X}=\dfrac{\sqrt{1-p}}{p}$. … (12.21)

Similarly, we can obtain other measures of our interest. But keeping applied nature of the programme in view, we are not focusing on proof of each measure. Some commonly used summary measures of geometric distribution are shown in Table 12.1 given as follows.

Table 12.1: Summary measures of geometric distribution

Name of measure	Formula	Name of measure	Formula
Mean	$\frac{1-p}{p}$	MGF	$\frac{p}{1-(1-p)e^t}$
Variance	$\frac{1-p}{p^2}$	Skewness	$\frac{2-p}{\sqrt{1-p}}$
Standard deviation	$\frac{\sqrt{1-p}}{p}$	Kurtosis	$6+\frac{p^2}{1-p}$

12.4 STORY, DEFINITION, PMF AND CDF OF NEGATIVE BINOMIAL DISTRIBUTION

In geometric distribution, we perform the experiment till, we get the first success but instead of the first success if we perform the experiment till r^{th} success then distribution of the random variable which counts the number of failures before the r^{th} success is known as negative binomial distribution. So, we can say that negative binomial distribution is the generalisation of geometric distribution or geometric distribution is the particular case of negative binomial distribution when r = 1. … (12.22)

Now, we can write the story of negative binomial distribution as follows.

Story of Negative Binomial Distribution: If we perform Bernoulli trial where probability of success is p and probability of failure is q = 1 – p till, we get r^{th} success then distribution of the random variable X which counts the number of failures before the r^{th} success is known as negative binomial distribution. So, negative binomial distribution models the distribution of Bernoulli trials where we are interested in the number of failures before the r^{th} success. … (12.23)

Let us obtain the probability of getting r^{th} success in $(x + r)^{th}$ trial where x represents the number of failures before the r^{th} success. This can happen only when $(x+r)^{th}$ trial is a success, but the remaining (r – 1) successes occur in the $x+r-1$ trials before the $(x + r)^{th}$ trial. But each trial is a Bernoulli trial with the same probability of success p, so the probability of getting (r – 1) successes out of (x + r – 1) trials can be obtained using binomial distribution Bin(x + r – 1, p) and is given by

$$\mathcal{P}\begin{pmatrix}\text{getting } r-1 \text{ successes}\\ \text{out of } x+r-1 \text{ trials}\end{pmatrix}=\binom{x+r-1}{r-1}p^{r-1}(1-p)^{x+r-1-(r-1)}$$

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

Now, required probability is given by

$$\mathcal{P}\begin{pmatrix}\text{getting x failures}\\ \text{before } r^{th} \text{ success}\end{pmatrix} = \mathcal{P}\left(\begin{Bmatrix}\text{getting } x-1 \text{ failures in}\\ \text{the first } x+r-1 \text{ trials}\end{Bmatrix} \cap \begin{Bmatrix}\text{getting } r^{th} \text{ success}\\ \text{in } (x+r)^{th} \text{ trial}\end{Bmatrix}\right)$$

$$= \mathcal{P}\begin{pmatrix}\text{getting } x-1 \text{ failures in}\\ \text{the first } x+r-1 \text{ trials}\end{pmatrix} \mathcal{P}\begin{pmatrix}\text{getting } r^{th} \text{ success}\\ \text{in } (x+r)^{th} \text{ trial}\end{pmatrix}$$

$$[\because \text{ both the events are independent}]$$

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

$$= \binom{x+r-1}{r-1} p^r (1-p)^x \qquad \text{... (12.24)}$$

Now, we define negative binomial distribution as follows.

Definition and PMF of Negative Binomial Distribution: If we perform Bernoulli trial where probability of success is p and probability of failure is q = 1 – p till, we get r^{th} success then PMF of the random variable X which counts the number of failures before the r^{th} success is given by

$$p_X(x) = \mathcal{P}(X = x) = \begin{cases} \binom{x+r-1}{r-1} p^r (1-p)^x, & \text{if } x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases} \qquad \ldots \text{(12.25)}$$

If random variable X has PMF given by (12.25), then, we say that it follows negative binomial distribution with parameters r and p and is denoted by writing $X \sim \text{NBin}(r, p)$. … (12.26)

We read it as X follows negative binomial distribution with parameters r and p. … (12.27)

Since the statistical software used for hands-on practice for the lab courses of this programme is R programming language, so, you should also know its notation in R. In R notation that is used for negative binomial distribution is nbinom(size, prob, …) in the stats package, where size represents the number of success and p represents the probability of success. In fact, there are four functions for negative binomial distribution namely dnbinom(x, size, prob, …), pnbinom(q, size, prob, …), qnbinom(p, size, prob, …) and rnbinom(n, size, prob, …). We have already explained the meaning of these functions in Unit 9. ... (12.28)

Let us check the **validity of the PMF of the negative binomial distribution**.

In checking the validity of PMF of negative distribution, we will use Vandermonde's Identity. So, let first state and prove it.

Vandermonde's Identity: For positive integer n and non-negative integer x such that $x \le n$ prove that $\binom{n}{x} = \binom{n}{n-x}$. … (12.29)

Proof: Suppose a bag has n items. Each time, we select x items out of n items then automatically in the bag, we are left with n – x items. So, the number of ways of selecting x items out of n is equal to selecting n – x items out of n.

Thus, we have $\binom{n}{x} = \binom{n}{n-x}, \forall\ x, 0 \le x \le n.$

Now, we can check the **validity of the PMF of the negative binomial distribution**.

We have to check two conditions non-negativity and normality refer to (4.62).

(1) Non-negativity: Since $0 < p \le 1$ and $\binom{x+r-1}{r-1} > 0$, so

$$\binom{x+r-1}{r-1} p^r (1-p)^x \ge 0, \;\; \forall \; x = 0, 1, 2, 3, \ldots$$

$$\Rightarrow p_X(x) = \mathcal{P}\left(X = x\right) \ge 0 \;\; \forall \; x = 0, 1, 2, 3, \ldots \qquad \ldots (12.30)$$

(2) **Normality**: Required to prove

$$\sum_{x=0}^{\infty} p_X(x) = \sum_{x=0}^{\infty} \mathcal{P}\left(X = x\right) = \sum_{x=0}^{\infty} \binom{x+r-1}{r-1} p^r (1-p)^x = 1 \qquad \ldots (12.31)$$

Using (12.29), we have

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

$$\Rightarrow \binom{x+r-1}{r-1} = \binom{x+r-1}{x}$$

$$= \frac{\lfloor x+r-1}{\lfloor x \lfloor x+r-1-x} \quad \left[\because \binom{n}{x} = \frac{\lfloor n}{\lfloor x \lfloor n-x}\right]$$

$$= \frac{\lfloor r+x-1}{\lfloor x \lfloor r-1}$$

$$= \frac{\left(r+\left(x-1\right)\right)\left(r+\left(x-2\right)\right)\ldots\left(r+1\right)\left(r\right)\lfloor r-1}{\lfloor x \; \lfloor r-1}$$

$$= \frac{\left(r+x-1\right)\left(r+x-2\right)\ldots\left(r+1\right)r}{\lfloor x}$$

$$= \frac{r(r+1)(r+2)\ldots\left(r+x-2\right)\left(r+x-1\right)}{\lfloor x} \left[\begin{array}{l}\text{Writing terms in}\\ \text{reverse order}\end{array}\right]$$

Taking **–** 1 common from each of the x terms in the numerator, we have

$$\binom{x+r-1}{r-1} = \frac{\left(-1\right)^x \left(-r\right)\left(-r-1\right)\left(-r-2\right)\ldots\left\{-r-\left(x-2\right)\right\}\left\{-r-\left(x-1\right)\right\}}{\lfloor x}$$

We know that the symbol $\binom{n}{x}$ stands for ${}^nC_x$ which represents the number of combinations of n things taken x at a time if n and x are positive integers and is equal to $\frac{n\left(n-1\right)\left(n-2\right)\ldots\left\{n-\left(x-1\right)\right\}}{\lfloor x}$. We may also use the symbol $\binom{n}{x}$ if n is any real number but, in that case, though it will not have the interpretation as mentioned in the case of positive integers. Yet it is equal to $\frac{n\left(n-1\right)\left(n-2\right)\ldots\left\{n-\left(n-x\right)\right\}}{\lfloor x}$. So, we have

$$\binom{x+r-1}{r-1} = (-1)^x \binom{-r}{x} \qquad \ldots (12.32)$$

Using (12.32) in (12.31), we have

$$\sum_{x=0}^{\infty} p_X(x) = \sum_{x=0}^{\infty} \mathcal{P}(X = x) = \sum_{x=0}^{\infty} (-1)^x \binom{-r}{x} p^r (1-p)^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} (-1+p)^x \qquad \left[\because (-1)^x (1-p)^x = (-1+p)^x\right]$$

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} (-1+p)^x (1)^{-r-x} \qquad \left[\because (1)^{-r-x} = 1\right]$$

$$= p^r (-1+p+1)^{-r} \qquad \left[\begin{array}{l}\because \sum_{x=0}^{\infty} \binom{-m}{x} a^x (1)^{-m-x} = (a+1)^{-m} \\ \text{where m is a positive integer}\end{array}\right]$$

$$= p^r (p)^{-r} = p^{r-r} = p^0 = 1 \qquad \ldots (12.33)$$

This proves (12.31). Hence, sum of all probabilities of negative binomial distribution is 1.

Hence, PMF defined by (12.25) of negative binomial distribution is a valid PMF.

Now, we define CDF of negative binomial distribution.

CDF of Negative Binomial Distribution: If $X \sim \text{NBin}(r, p)$, then PMF of X is given by

$$p_X(x) = \mathcal{P}(X = x) = \begin{cases} \binom{x+r-1}{r-1} p^r (1-p)^x, & \text{if } x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}$$

Let x be any fixed real number then CDF of the random variable X is given by

$$F_X(x) = \mathcal{P}(X \le x) = \begin{cases} 0, & \text{if } x < 0 \\ \sum_{x=0}^{[x]} \binom{x+r-1}{r-1} p^r (1-p)^x, & \text{if } x \ge 0 \end{cases} \qquad \ldots (12.34)$$

Let us do one example.

Example 2: A biased coin which has the probability of getting a head as 1/4 and a tail as 3/4 is tossed till, we get the third head. If X counts the number of failures before the third head, then obtain the PMF and CDF of X and also plot the PMF and CDF of X.

Solution: If we define getting a head as a success and getting a tail as a failure then the random variable X which counts the number of failures before the third head follows a negative binomial distribution with parameters r = 3, p = 1/4. So, the PMF of X is given by putting r = 3, p = 1/4 in (12.25)

$$p_X(x) = \mathcal{P}(X = x) = \begin{cases} \binom{x+3-1}{3-1} \left(\frac{1}{4}\right)^3 \left(1-\frac{1}{4}\right)^x, & x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}$$

Or

$$p_X(x) = \mathcal{P}(X = x) = \begin{cases} \binom{x+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^x, & x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases} \quad \ldots (12.35)$$

and the CDF of X is given by

$$F_X(x) = \mathcal{P}(X \le x) = \begin{cases} 0, & \text{if } x < 0 \\ \sum_{x=0}^{[x]} \binom{x+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^x, & \text{if } x \ge 0 \end{cases} \quad \ldots (12.36)$$

Probabilities for X = 0, 1, 2, 3, …, 10 are given by

$$\left.\begin{aligned}
\mathcal{P}(X=0) &= \binom{0+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^0 = 0.015625 \\
\mathcal{P}(X=1) &= \binom{1+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^1 = 0.03515625 \\
\mathcal{P}(X=2) &= \binom{2+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^2 = 0.05273438 \\
\mathcal{P}(X=3) &= \binom{3+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^3 = 0.06591797 \\
\mathcal{P}(X=4) &= \binom{4+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^4 = 0.07415771 \\
\mathcal{P}(X=5) &= \binom{5+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^5 = 0.0778656 \\
\mathcal{P}(X=6) &= \binom{6+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^6 = 0.0778656 \\
\mathcal{P}(X=7) &= \binom{7+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^7 = 0.07508469 \\
\mathcal{P}(X=8) &= \binom{8+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^8 = 0.07039189 \\
\mathcal{P}(X=9) &= \binom{9+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^9 = 0.06452590 \\
\mathcal{P}(X=10) &= \binom{10+2}{2}\left(\frac{1}{64}\right)\left(\frac{3}{4}\right)^{10} = 0.05807331
\end{aligned}\right\} \quad \ldots (12.37)$$

Now, using (12.36) CDF of X is given by

$$F_X(x) = \mathcal{P}(X \le x) = \begin{cases} 0, & \text{if } x < 0 \\ 0.015625, & \text{if } 0 \le x < 1 \\ 0.05078125, & \text{if } 1 \le x < 2 \\ 0.10351563, & \text{if } 2 \le x < 3 \\ 0.16943359, & \text{if } 3 \le x < 4 \\ 0.24359131, & \text{if } 4 \le x < 5 \\ 0.32145691, & \text{if } 5 \le x < 6 \\ 0.39932251, & \text{if } 6 \le x < 7 \\ 0.4744072, & \text{if } 7 \le x < 8 \\ 0.54479909, & \text{if } 8 \le x < 9 \\ 0.60932499, & \text{if } 9 \le x < 10 \\ 0.66739830, & \text{if } 10 \le x < 11 \\ \vdots & \end{cases} \quad \ldots (12.38)$$

PMF and CDF are plotted in Fig. 12.2 (a) and (b) respectively as follows.

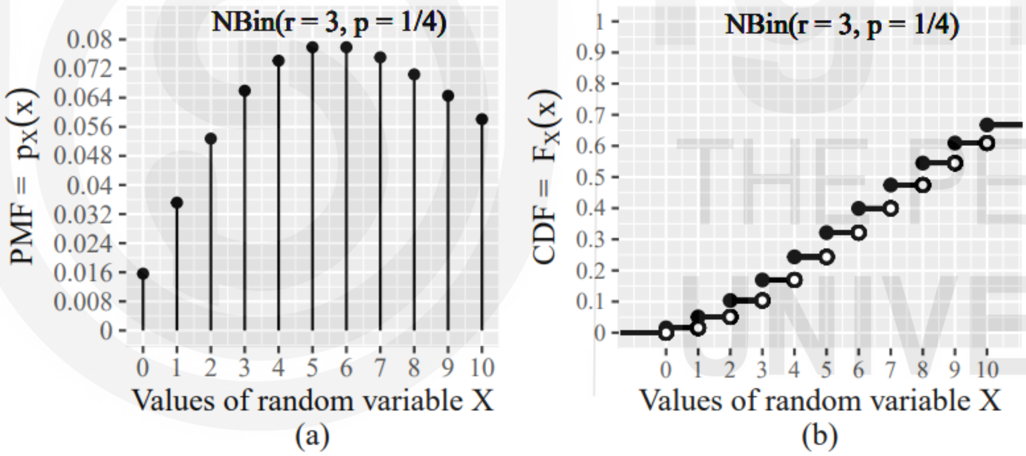


Fig. 12.2: Visualisation of (a) PMF (b) CDF of the NBin(3, 1/4) discussed in Example 2

12.5 MGF AND OTHER SUMMARY MEASURES OF NEGATIVE BINOMIAL DISTRIBUTION

In the previous section, you have studied PMF and CDF of negative binomial distribution. In this section, we want to obtain MGF, mean and variance of negative binomial distribution.

Calculation of MGF

Suppose we perform Bernoulli trials with probability of success p till, we get r^{th} success. Let $X_i,\ i = 1, 2, 3, \ldots, r$ be the random variable which counts the number of failures between i – 1 and the i^{th} success. So, each $X_i \sim \text{Geom}(p),\ i = 1, 2, 3, \ldots, r.$ Using (12.16), (12.17) and (12.20), we have

$$M_{X_i}(t) = \frac{p}{1-(1-p)e^t},\ t < \ln\left(\frac{1}{1-p}\right),\ i = 1, 2, 3, \ldots, r \qquad \ldots (12.39)$$

$$E(X_i) = \frac{1-p}{p},\ i = 1, 2, 3, \ldots, r \qquad \ldots (12.40)$$

$$V(X_i) = \frac{1-p}{p^2},\ i = 1, 2, 3, \ldots, r \qquad \ldots (12.41)$$

Let $X = X_1 + X_2 + X_3 + \ldots + X_r$ … (12.42)

$$M_X(t) = M_{X_1+X_2+X_3+\ldots+X_r}(t) = M_{X_1}(t)M_{X_2}(t)\ldots M_{X_r}(t)\ [\text{Using (7.94) of this course}]$$

$$= \prod_{i=1}^{r} M_{X_i}(t) \qquad [\because \text{ each } X_i \sim \text{Geom}(p),\ i = 1, 2, 3, \ldots, r]$$

$$= \prod_{i=1}^{r} \frac{p}{1-(1-p)e^t},\ t \in \mathbb{R} \quad [\text{Using (12.39)}]$$

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

$$M_X(t) = \left(\frac{p}{1-(1-p)e^t}\right)^r \qquad \ldots (12.43)$$

Calculation of Mean

Applying expectation on both sides of (12.42), we have

$$E(X) = E\left(X_1 + X_2 + X_3 + \ldots + X_r\right)$$

$$= E(X_1) + E(X_2) + E(X_3) + \ldots + E(X_r) \quad \left[\begin{array}{l}\text{Using addition theorem of}\\ \text{expectation refer to (7.29)}\end{array}\right]$$

$$= \underbrace{\frac{p}{1-p} + \frac{p}{1-p} + \frac{p}{1-p} + \ldots + \frac{p}{1-p}}_{r-\text{times}} \quad [\text{Using (12.40)}]$$

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

Hence, mean of X = expected value of $X = E(X) = \dfrac{rp}{1-p}$ … (12.44)

Calculation of Variance

Applying variance operator V(.) on both sides of (12.42), we have

$$V(X) = V\left(X_1 + X_2 + X_3 + \ldots + X_r\right)$$

$$= V(X_1) + V(X_2) + V(X_3) + \ldots + V(X_r) \quad \left[\begin{array}{l}\text{Here, } X_1, X_2, X_3, \ldots, X_r \text{ are}\\ \text{independent. So, using (7.74)}\end{array}\right]$$

$$= \underbrace{\frac{p}{(1-p)^2} + \frac{p}{(1-p)^2} + \frac{p}{(1-p)^2} + \ldots + \frac{p}{(1-p)^2}}_{r-\text{times}} \quad [\text{Using (12.41)}]$$

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

Hence, variance of $X = V(X) = \dfrac{rp}{(1-p)^2}$ … (12.45)

We know that standard deviation of X is positive square root of variance of X.

Hence, standard deviation $= \text{SD}(X) = \sqrt{\text{Variance of } X} = \dfrac{\sqrt{rp}}{(1-p)}$ … (12.46)

Similarly, we can obtain other measures of our interest. But keeping applied nature of the programme in view, we are not focusing on proof of each measure. Some commonly used summary measures of negative binomial distribution are shown in Table 12.2 given as follows.

Table 12.2: Summary measures of negative binomial distribution

Name of measure	Formula	Name of measure	Formula
Mean	$\dfrac{rp}{1-p}$	MGF	$\left(\dfrac{p}{1-(1-p)e^t}\right)^r$
Variance	$\dfrac{rp}{(1-p)^2}$	Skewness	$\dfrac{2-p}{\sqrt{r(1-p)}}$
Standard deviation	$\dfrac{\sqrt{rp}}{(1-p)}$	Kurtosis	$\dfrac{6}{r}+\dfrac{p^2}{r(1-p)}$

12.6 APPLICATIONS AND ANALYSIS OF GEOMETRIC AND NEGATIVE BINOMIAL DISTRIBUTIONS

In this section, we will apply geometric and negative binomial distributions to solve some problems where assumptions of these distributions make sense. We will also do some analysis of these two distributions.

Example 3: Probability that Kavita hits a target is 0.7. Find the probability that Kavita hits the target in her third attempt. Assume that all the trials are independent.

Solution: Here it is given that trials are independent which means probability of hitting the target does not change trial to trial and, we are given that constant probability of hitting the target in each trial is p = 0.7. We are interested in getting the probability of the first success. So, it is the situation of geometric distribution. Let X be the random variable which counts the number of failures before the first success. Hence, X ~ Geom(0.7). Thus, required probability is given by

$$\mathcal{P}(X=2)=(1-0.7)^2(0.7) \qquad \left[\because\ \mathcal{P}(X=x)=\begin{cases}(1-p)^x p, & x=0,1,2,3,\ldots\\ 0, & \text{otherwise}\end{cases}\right]$$

$$=0.063$$

Example 4: **Application in solving a game problem**: Arnav and Abhishek usually play table Tanis. On the basis of past experience, it is known that the probability that Arnav beats Abhishek is 0.55. Assume that each game is independent of the other. One day they decide to play till one of them wins five games. Find the probability that result of eighth game decide the winner.

Solution: Let X be the random variable which counts the number of failures before fifth success of Arnav while Y be the random variable which counts the number of failures before fifth success of Abhishek. Since each game is independent of the other so the probability that Arnav beats Abhishek is constant in each game and therefore, X ~ NBin(r = 5, p = 0.55). Similarly, Y ~ NBin(r = 5, p = 0.45). Now, result of the eighth game will decide winner of the game if either 8th game is the 5th success of Arnav or 8th game is the 5th success of Abhishek. So, required probability is given by

$$\mathcal{P}(X=3)+\mathcal{P}(Y=3)=\binom{3+5-1}{5-1}(0.55)^5(1-0.55)^3+\binom{3+5-1}{5-1}(0.45)^5(1-0.45)^3$$

$$,\left[\because \text{ if } Z\sim \text{NBin(r, p), then } \mathcal{P}(Z=z)=\begin{cases}\binom{x+r-1}{r-1}p^r(1-p)^z, & \text{if } x=0,1,2,3,\ldots\\ 0, & \text{otherwise}\end{cases}\right]$$

$$\Rightarrow \mathcal{P}(X=3)+\mathcal{P}(Y=3)=\binom{7}{4}(0.55)^5(0.45)^3+\binom{7}{4}(0.45)^5(0.55)^3$$

$$=0.2679693 \qquad \ldots\ (12.47)$$

Example 5: **Application in solving collector's problem**: There is a sequence of n toys. Suppose each packet of a particular food item contains one of these n toys. Assume that each toy is equally likely to be packed in the packet of this particular food item. Suppose you are interested in collecting complete sequence of toys. What is the expected value of the number of packets of food items you should buy.

Solution: Obviously the first packet that you will buy will contain one of the n toys. So, after purchasing one packet of this particular food you have one toy in your hand out of the total n toys. When you will buy the second food packet then it may contain the toy which you already have with you or you can get one among the remaining n – 1 toys. So, probability of getting additional toy is $\frac{n-1}{n}$ and probability of getting the existing toy is $\frac{1}{n}$. Let X_1 be the random variable which counts the number of failures before the first success when you buy second food item onward. So, $X_1\sim \text{Geom}\left(p_1=\frac{n-1}{n}\right)$. We know that expected value of geometric distribution with parameter p_1 is $1/p_1$.. So, in the present case expected value of X_1 is given by $E(X_1)=\frac{1}{p_1}=\frac{n}{n-1}$. … (12.48)

So, expected number of food packets that you should buy to get two different toys is $1+\frac{n}{n-1}$. … (12.49)

After having two different toys in your hand, when you will go for buying the more food items then either the new packet of food may contain one from the two toys which you already have with you or it may contain the one among the remaining n – 2 toys. So, probability of getting additional toy is $\frac{n-2}{n}$ and probability of getting the one from the two existing toys is $\frac{2}{n}$. Let X_2 be the random variable which counts the number of failures before the first success where counting starts after getting two toys of different types. So,

$X_2 \sim \text{Geom}\left(p_2 = \dfrac{n-2}{n}\right)$. Expected value of X_2 is given by $E(X_2) = \dfrac{1}{p_2} = \dfrac{n}{n-2}$. … (12.50)

So, expected number of food packets that you should buy to get three different toys is $1 + \dfrac{n}{n-1} + \dfrac{n}{n-2}$. … (12.51)

Continuing in this fashion, the expected number of food packets that you should buy to get all the toys of the sequence of n toys is given by

$$1 + \frac{n}{n-1} + \frac{n}{n-2} + \frac{n}{n-3} + \ldots + \frac{n}{n-(n-2)} + \frac{n}{n-(n-1)}$$

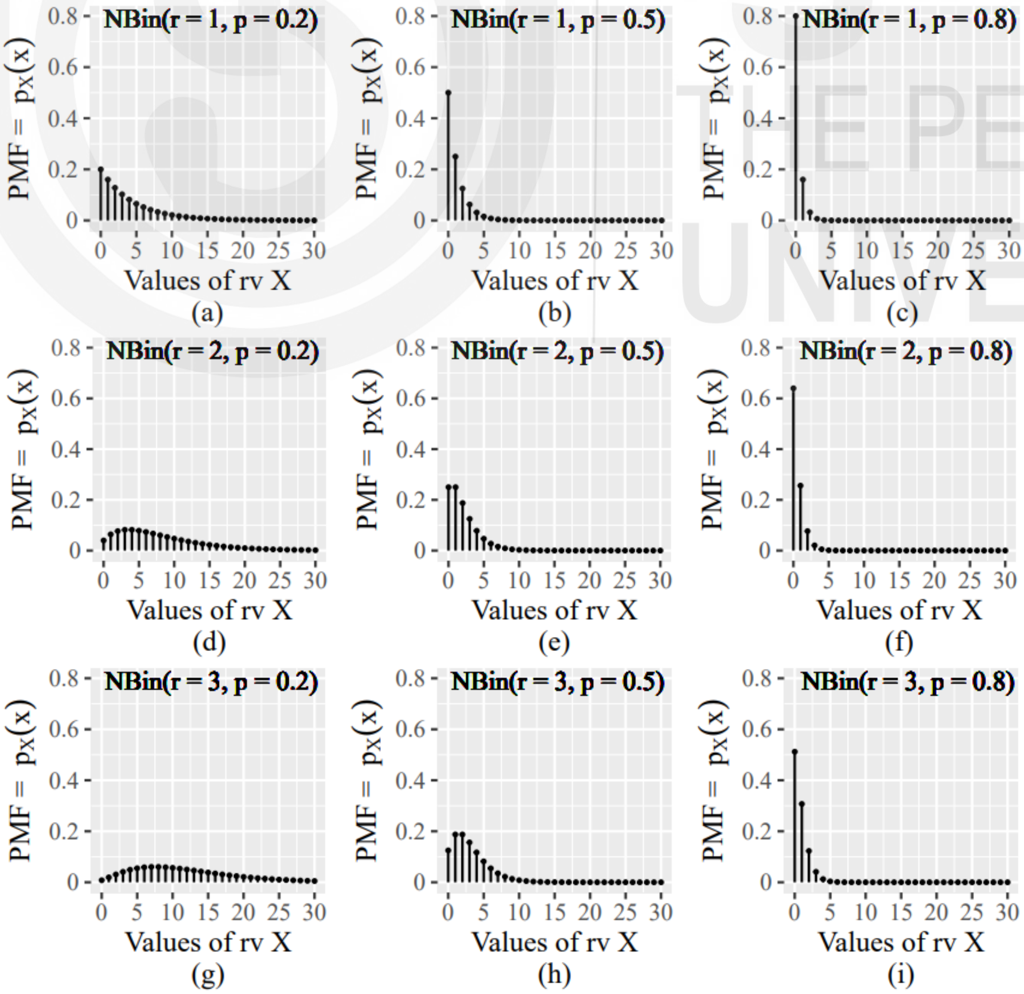
$$= 1 + \frac{n}{n-1} + \frac{n}{n-2} + \frac{n}{n-3} + \ldots + \frac{n}{2} + \frac{n}{1}$$

$$= \frac{n}{1} + \frac{n}{2} + \frac{n}{3} + \frac{n}{4} + \ldots + \frac{n}{n-2} + \frac{n}{n-1} + \frac{n}{n} \quad [\text{Writing terms in reverse order}]$$

$$= n\left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{n-2} + \frac{1}{n-1} + \frac{1}{n}\right) \quad \ldots (12.52)$$

Example 6: **Analysis of PMF of NBin(r, p) as r and p vary**: Plot PMF of negative binomial for different combinations of values of r and p where r = 1, 2, 3, 4, 5 and p = 0.2, 0.5, 0.8.

Solution: PMF of negative binomial distribution when p = 1 to 5 and p = 0.2, 0.5, 0.8 are shown in Fig. 12.3 (a) to (o) given as follows.



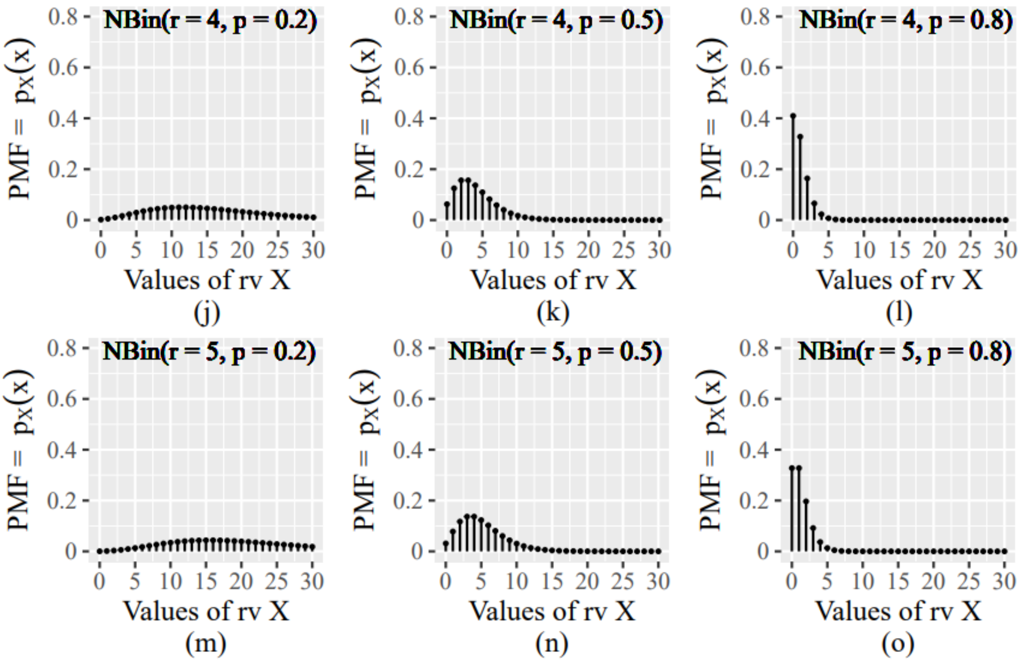


Fig. 12.3: Visualisation of PMF of negative binomial distributions when (a) r = 1, p = 0.2 (b) r = 1, p = 0.5 (c) r = 1, p = 0.8 (d) r = 2, p = 0.2 (e) r = 2, p = 0.5 (f) r = 2, p = 0.8 (g) r = 3, p = 0.2 (h) r = 3, p = 0.5 (i) r = 3, p = 0.8 (j) r = 4, p = 0.2 (k) r = 4, p = 0.5 (l) r = 4, p = 0.8 (m) r = 5, p = 0.2 (n) r = 5, p = 0.5 (o) r = 5, p = 0.8

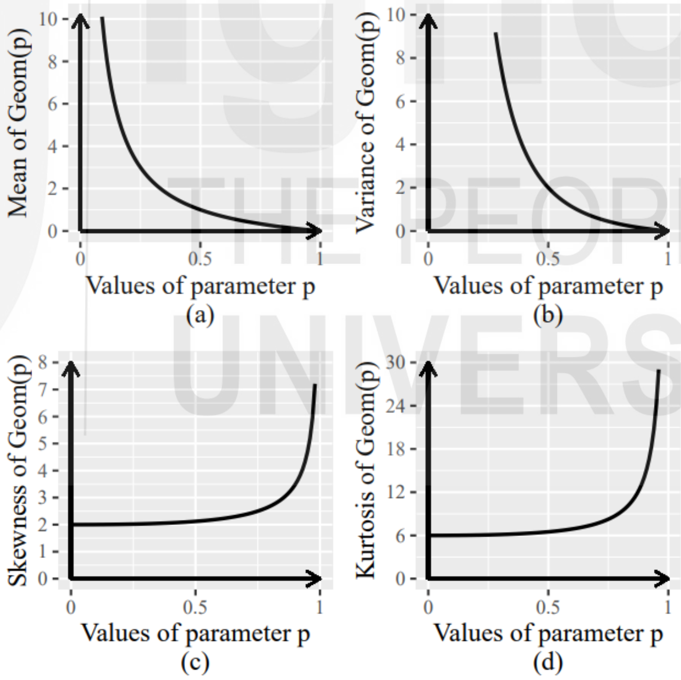


Fig. 12.4: Visualisation of summary measures (a) mean (b) variance (c) skewness (d) kurtosis of geometric distribution against values of p

Example 7: **Analysis of mean, variance, skewness and kurtosis of Geom(p) as p varies**: Plot mean, variance, skewness and kurtosis of geometric distribution as p varies. Also, specify some important observations noted in the graphical analysis of these summary measures.

Solution: If X ~ Geom(p), then, we know that mean, variance, skewness and kurtosis of geometric distribution are given by

$$\text{mean} = \frac{1-p}{p}, \ \text{variance} = \frac{1-p}{p^2}, \ \text{skewness} = \frac{2-p}{\sqrt{1-p}}, \ \text{and kurtosis} = 6 + \frac{p^2}{1-p}.$$

Mean, variance, skewness and kurtosis of geometric distribution as p varies are shown in Fig. 12.4 (a) to (d).

Some important observations from the graphical analysis of these summary measures are mentioned as follows.

- As p increases mean and variance of the distribution both continuously decrease and tend to 0 if p tends to 1 while they tend to infinity as p tends to 0. At p = 1/2 value of mean is 1 while value of variance is 2. Variance increases more rapidly than mean.
- Minimum value of skewness is 2 when p = 0 and it remains almost 2 till $p \le 1/2$. But as p crosses 1/2 it starts to increase. But p always remains greater than equal to 2 so geometric distribution is always positively highly skewed distribution to know why you may refer to (9.78) to (9.81).
- Minimum value of kurtosis is 6 when p = 0 and it remains almost 6 till $p \le 1/2$. But as p crosses 1/2 it starts to increase. But p always remains greater than equal to 6 so geometric distribution is always leptokurtic refer (9.90).

Example 8: **Analysis of mean, variance, skewness and kurtosis of NBin(r, p) as r and p vary**: Plot mean, variance, skewness and kurtosis of negative binomial distribution as p varies for at least three different values of r. Also, specify some important observations noted in the graphical analysis of these summary measures. After getting the idea from the graphs comment on the skewness and kurtosis of the distribution.

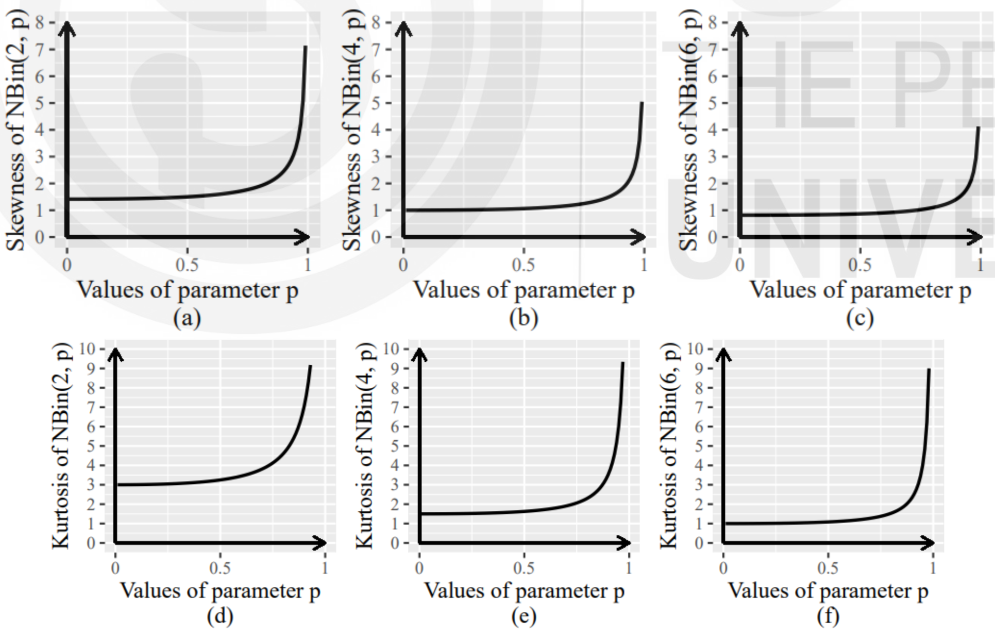


Fig. 12.5: Visualisation of summary measures (a) skewness when r = 2 (b) skewness when r = 4 (c) skewness when r = 6 (d) kurtosis when r = 2 (e) kurtosis when r = 4 (f) kurtosis when r = 6

Solution: If X ~ NBin(r, p), then, we know that mean, variance, skewness and kurtosis of negative binomial distribution are given by

$$\text{mean} = \frac{r(1-p)}{p}, \text{ variance} = \frac{r(1-p)}{p^2}, \text{ skewness} = \frac{2-p}{\sqrt{r(1-p)}}, \text{ kurtosis} = \frac{6}{r} + \frac{p^2}{r(1-p)}$$

We know that when r = 1 then negative binomial distribution becomes geometric distribution and, we have already done this type of analysis for geometric distribution in Example 7. So, let us consider three values of r as 2, 4 and 6. But mean and variance are just r times the mean and variance of the geometric distribution so there is no need of their graphical analysis. But r is sitting in denominator in the formulae of skewness and kurtosis so as r increases, they will decrease. So, let us analysis their graphical behaviour as p varies for r = 2, 4, 6 and are shown in Fig. 12.5 (a) to (f).

Some important observations from the graphical analysis of these summary measures are mentioned as follows.

- As p increases mean and variance of the distribution both continuously decrease and tend to 0 if p tends to 1 while they tend to infinity as p tends to 0. At p = 1/2 value of mean is 1 while value of variance is 2. Variance increases more rapidly then mean.
- Minimum value of skewness is 2 when p = 0 and it remains almost 2 till $p \leq 1/2.$ But as p crosses 1/2 it starts to increase. But p always remains greater than equal to 2 so geometric distribution is always positively highly skewed distribution to know why you may refer to (9.78) to (9.81).
- Minimum value of kurtosis is 6 when p = 0 and it remains almost 6 till $p \leq 1/2.$ But as p crosses 1/2 it starts to increase. But p always remains greater than equal to 6 so geometric distribution is always leptokurtic refer (9.90).

Observation: From Fig. 12.3 (a) to (o), we note that as value of r increases or value of p decreases PMF of negative binomial distribution spread on horizontal axis towards right.

Now, you can try the following two Self-Assessment Questions.

SAQ 1

Arnav and Abhishek usually play table Tanis. On the basis of past experience, it is known that the probability that Arnav beats Abhishek is 0.55. Assume that each game is independent of the other. One day they decide to play till one of them wins five games. If we are given that 8^{th} game has decided the winner then find the probability that Arnav is the winner.

SAQ 2

A person has 10 keys in his hand and he really does not know which one is the right key of the lock he wants to open. Out of these 10 keys there is only one key which can open the desired lock. Assume that he tries keys one by one randomly by selecting one key out of the 10 keys in his hand. Find the expected number of try he has to do.

12.7 SUMMARY

A brief summary of what we have covered in this unit is given as follows:

- **Definition and PMF of Geometric Distribution**: If we perform Bernoulli trials till, we get the first success and the random variable X counts the number of failures before the first success then the PMF of X is given by

$$p_X(x) = \mathcal{P}\left(X = x\right) = \begin{cases} (1-p)^x\, p, & x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}$$

This is the PMF of geometric distribution with parameter p and is denoted by writing $X \sim \text{Geom}(p)$.

- **CDF of Geometric Distribution**: Let x be any fixed real number then CDF of the random variable X is given by

$$F_X(x) = \mathcal{P}\left(X \le x\right) = \begin{cases} 0, & \text{if } x < 0 \\ 1-(1-p)^{[x]+1}, & \text{if } x \ge 0 \end{cases}$$

- **Summary measures of Poisson distribution**

Name of measure	Formula	Name of measure	Formula
Mean	$\frac{1-p}{p}$	MGF	$\frac{p}{1-(1-p)e^t}$
Variance	$\frac{1-p}{p^2}$	Skewness	$\frac{2-p}{\sqrt{1-p}}$
Standard deviation	$\frac{\sqrt{1-p}}{p}$	Kurtosis	$6+\frac{p^2}{1-p}$

- **Definition of Negative Binomial Distribution**: If we perform Bernoulli trial where probability of success is p and probability of failure is q = 1 – p till, we get r^{th} success then PMF of the random variable X which counts the number of failures before the r^{th} success is given by

$$p_X(x) = \mathcal{P}\left(X = x\right) = \begin{cases} \binom{x+r-1}{r-1} p^r (1-p)^x, & \text{if } x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}$$

- **CDF of Negative Binomial Distribution**: Let x be any fixed real number then CDF of the random variable X is given by

$$F_X(x) = \mathcal{P}\left(X \le x\right) = \begin{cases} 0, & \text{if } x < 0 \\ \sum_{x=0}^{[x]} \binom{x+r-1}{r-1} p^r (1-p)^x, & \text{if } x \ge 0 \end{cases}$$

- **Summary measures of Hypergeometric distribution**

Name of measure	Formula	Name of measure	Formula
Mean	$\frac{rp}{1-p}$	MGF	$\left(\frac{p}{1-(1-p)e^t}\right)^r$
Variance	$\frac{rp}{(1-p)^2}$	Skewness	$\frac{2-p}{\sqrt{r(1-p)}}$
Standard deviation	$\frac{\sqrt{rp}}{(1-p)}$	Kurtosis	$\frac{6}{r}+\frac{p^2}{r(1-p)}$

12.8 TERMINAL QUESTIONS

1. Arnav and Abhishek usually play table Tanis. On the basis of past experience, it is known that the probability that Arnav beats Abhishek is

0.55. Assume that each game is independent of the other. One day they decide to play till one of them wins five games. Find the probability that eventually Arnav is the winner on that day.

2. Suppose you have a 20 faces fair die. Using result of Example 5 or otherwise find the expected of throws to get every face of the die as the outcome.

3. **Lack of Memory Property or Memoryless Property of Geometric Distribution**: State and prove that geometric distribution has memoryless property. **Statement**: If X ~ Geom(p) then prove that $\mathcal{P}\left(X \geq j+k \mid X \geq j\right)=\mathcal{P}\left(X \geq j\right)$.

4. Why the name negative binomial distribution is given to the distribution of the random variable which counts the number of failures before the r^{th} success.

12.9 SOLUTIONS/ANSWERS

Self-Assessment Questions (SAQs)

1. Let the random variables X and Y be the same as explained in Example 4 of this unit. Now, 8^{th} game has decided the winner it means winner may be Arnav or Abhishek. So, using conditional probability required probability is given by

$$\mathcal{P}\left(X=3 \mid X=3 \text{ or } Y=3\right)=\frac{\mathcal{P}\left(\left(X=3\right)\cap\left(X=3 \text{ or } Y=3\right)\right)}{\mathcal{P}\left(X=3 \text{ or } Y=3\right)}$$

$$\left[\because \mathcal{P}\left(A \mid B\right)=\frac{\mathcal{P}\left(A\cap B\right)}{\mathcal{P}\left(B\right)}\right]$$

$$=\frac{\mathcal{P}\left(X=3\right)}{\mathcal{P}\left(X=3\right)+\mathcal{P}\left(X=3\right)} \quad \ldots (12.53)$$

$$\left[\begin{array}{l}\because \text{Events } X=3 \text{ and } Y=3\\ \text{are mutually exclusive}\end{array}\right]$$

we have already obtained value of denominator in Example 4 which is 0.2679693. So, let us first obtain the value of numerator.

$$\mathcal{P}\left(X=3\right)=\binom{3+5-1}{5-1}(0.55)^5(1-0.55)^3$$

$$\left[\because \mathcal{P}\left(X=x\right)=\begin{cases}\binom{x+r-1}{r-1}p^r(1-p)^x, & \text{if } x=0,1,2,3,\ldots\\ 0, & \text{otherwise}\end{cases}\right]$$

$$\Rightarrow \mathcal{P}\left(X=3\right)=\binom{7}{4}(0.55)^5(0.45)^3=0.1605163$$

Using values of numerator and denominator in (12.53), we get

$$\mathcal{P}\left(X=3 \mid X=3 \text{ or } Y=3\right)=\frac{0.1605163}{0.2679693}=0.59901.$$

2. Out of the 10 keys in his hand only one can open the lock so probability of opening the lock is 1/10. If we call opening the lock as success then probability of success is p = 1/10. Since each time he selects a key

randomly out of the 10 keys in his hand so probability of success in each trial is constant. Let the random variable X counts the number of failure before the first success. So, X ~ Geom(1/10). Hence, expected number of tries will be simply mean of the geometric distribution with parameter p = 1/10. But we know that expected value of mean of the geometric distribution is 1/p. Thus, expected number of tries he has to do is 1/(1/10) = 10.

Terminal Questions

1. As per the rules of the game to win, Arnav has to win 5 games before Abhishek wins 5 games. So, Arnav can be winner of that day if either his 5th game is his 5th success or 6th game is his 5th success or 7th game is his 5th or 8th game is his 5th success or 9th game is his 5th success. In other words, either Arnav has 0 failures before 5th success or 1 failure before 5th success or 2 failures before 5th success or 3 failures before 5th success or at the most 4 failures before 5th success. So, required probability is given by (recall that in this problem X counts the number of failures before the fifth success)

$$\mathcal{P}(X=0 \text{ or } 1 \text{ or } 2 \text{ or } 3 \text{ or } 4) = \mathcal{P}(X=0)+\mathcal{P}(X=1)+\mathcal{P}(X=2) + \mathcal{P}(X=3)+\mathcal{P}(X=4) \quad \ldots (12.54)$$

$$\left[\because \text{Events } X=0,\ X=1,\ X=2,\ X=3 \text{ and } X=4 \text{ are mutually exclusive}\right]$$

$$\Rightarrow \mathcal{P}(X=0 \text{ or } 1 \text{ or } 2 \text{ or } 3 \text{ or } 4) = \binom{0+5-1}{5-1}(0.55)^5(1-0.55)^0 + \binom{1+5-1}{5-1}(0.55)^5(1-0.55)^1 + \binom{2+5-1}{5-1}(0.55)^5(1-0.55)^2 + \binom{3+5-1}{5-1}(0.55)^5(1-0.55)^3 + \binom{4+5-1}{5-1}(0.55)^5(1-0.55)^4$$

$$\left[\because \mathcal{P}(X=x) = \begin{cases} \binom{x+r-1}{r-1} p^r (1-p)^x, & \text{if } x = 0, 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}\right]$$

$$\Rightarrow \mathcal{P}(X=0 \text{ or } 1 \text{ or } 2 \text{ or } 3 \text{ or } 4) = \binom{4}{4}(0.55)^5(0.45)^0 + \binom{5}{4}(0.55)^5(0.45)^1 + \binom{6}{4}(0.55)^5(0.45)^2 + \binom{7}{4}(0.55)^5(0.45)^3 + \binom{8}{4}(0.55)^5(0.45)^4$$

$$= 0.6214209 \qquad [\text{Using scientific calculator}]$$

In your lab exam you will obtain value of such probability using following R code. Screenshot with output is shown as follows.

```
> pnbinom(q = 4, size = 5, prob = 0.55)
[1] 0.6214209
```

2. Using result of Example 5 the expected number of throws to get every face of the die as the outcome is given by

$$20\left(\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots+\frac{1}{18}+\frac{1}{19}+\frac{1}{20}\right)=71.95479\simeq 72.$$

3. We know that PMF of geometric distribution with parameter p is given by

$$p_X(x)=\mathcal{P}\left(X=x\right)=\begin{cases}(1-p)^x\,p, & x=0,1,2,3,\ldots\\ 0, & \text{otherwise}\end{cases}$$

We claim that $\mathcal{P}\left(X\geq j\right)=(1-p)^j$.

$$\begin{aligned}\mathcal{P}\left(X\geq j\right)&=\mathcal{P}\left(X\geq j\right)+\mathcal{P}\left(X\geq j+1\right)+\mathcal{P}\left(X\geq j+2\right)+\ldots\\ &=(1-p)^j p+(1-p)^{j+1}p+(1-p)^{j+2}p+\ldots\\ &=(1-p)^j p\left[1+(1-p)+(1-p)^2+\ldots\right]\\ &=(1-p)^j p\left[\frac{1}{1-(1-p)}\right]\quad \left[\begin{array}{l}\because \text{ sum of infinite GP } a+ar+ar^2+\ldots\\ =\dfrac{a}{1-r},\ \text{provided } |r|<1\end{array}\right]\\ &=\frac{(1-p)^j p}{p}=(1-p)^j \qquad \ldots (12.55)\end{aligned}$$

Now,

$$\begin{aligned}\mathcal{P}\left(X\geq j+k\,|\,X\geq j\right)&=\frac{\mathcal{P}\left(\left(X\geq j+k\right)\cap\left(X\geq j\right)\right)}{\mathcal{P}\left(X\geq k\right)}\quad\left[\because\ \mathcal{P}\left(A\,|\,B\right)=\frac{\mathcal{P}\left(A\cap B\right)}{\mathcal{P}\left(B\right)}\right]\\ &=\frac{\mathcal{P}\left(X\geq j+k\right)}{\mathcal{P}\left(X\geq k\right)}\\ &=\frac{(1-p)^{j+k}}{(1-p)^k}\qquad[\text{Using } (12.55)]\\ &=(1-p)^k\end{aligned}$$

The above result reveals that the conditional probability of at least first $j+k$ trials are unsuccessful before the first success given that at least first j trials were unsuccessful, is the same as the probability that the first k trials were unsuccessful. So, the probability to get first success remains same if we start counting of k unsuccessful trials from anywhere provided all the trials preceding to it are unsuccessful, i.e., the future does not depend on past, it depends only on the present. So, the geometric distribution forgets the preceding trials and hence this property is given the name “forgetfulness property” or “Memoryless property” or “lack of memory” property.

4. Using (12.32) in (12.25) PMF of negative binomial distribution is given by

$$p_X(x)=\mathcal{P}\left(X=x\right)=\begin{cases}(-1)^x\dbinom{-r}{x}p^r(1-p)^x, & \text{if } x=0,1,2,3,\ldots\\ 0, & \text{otherwise}\end{cases}$$

Let us consider

$$p_X(x) = \mathcal{P}(X = x) = (-1)^x \binom{-r}{x} p^r (q)^x, \text{ if } x = 0, 1, 2, 3, \ldots \; [\because 1 - p = q]$$

$$= \binom{-r}{x} p^r (-q)^x, \text{ if } x = 0, 1, 2, 3, \ldots$$

$$= \binom{-r}{x} (-q)^x (1)^{-r-x} p^r, \text{ if } x = 0, 1, 2, 3, \ldots \qquad [\because (1)^{-r-x} = 1]$$

So, probabilities for X = 0, 1, 2, 3, 4, … are successive terms of the binomial expansion given as follows.

$$(1 + (-q))^{-r} p^r \qquad \left[\because (a+b)^n = \binom{n}{0} a^n b^0 + \binom{n}{1} a^{x-1} b + \binom{n}{2} a^x b^{n-x} + \ldots\right]$$

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

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

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

Which is a binomial expansion with negative index. That is why it is known as negative binomial distribution.