MODULE 4

SOME SPECIAL DISCRETE DISTRIBUTIONS AND THEIR PROPERTIES

LECTURES 17-19

Topics

4.1 BERNOULLI EXPERIMENT AND RELATED DISTRIBUTIONS

4.1.1 Bernoulli Distribution

4.1.2 Binomial Distribution

4.1.3 Binomial Distribution and Sampling with Replacement

4.2 NEGATIVE BINOMIAL DISTRIBUTION

4.3 THE HYPERGEOMETRIC DISTRIBUTION

4.4 THE POISSON DISTRIBUTION

4.5 DISCRETE UNIFORM DISTRIBUTION

MODULE 4

SOME SPECIAL DISCRETE DISTRIBUTIONS AND THEIR PROPERTIES

LECTURE 17

Topics

4.1 BERNOULLI EXPERIMENT AND RELATED DISTRIBUTIONS

4.1.1 Bernoulli Distribution
4.1.2 Binomial Distribution
4.1.3 Binomial Distribution and Sampling with Replacement

4.2 NEGATIVE BINOMIAL DISTRIBUTION

The probability distribution of a random variable (r.v.) X defined on a probability space $(\Omega, \mathcal{F}, P)$ describes the probability law according to which X takes values in various Borel sets. Recall that the probability distribution of a r.v. X is completely determined by its distribution function (d.f.) or by its probability mass function/probability density function (p.m.f. /p.d.f.). Also recall that a r.v. X is of discrete type if there exists a non-empty countable set S_X such that $P(\{X = x\}) > 0, \forall x \in S_X$ and $\sum_{x \in S_X} P(\{X = x\}) = 1$. The set S_X is called the support of the distribution of X (or of X) and the function $f_X: \mathbb{R} \to \mathbb{R}$ defined by

$$f_X(x) = \begin{cases} P(\{X = x\}), & \text{if } x \in S_X \\ 0, & \text{otherwise} \end{cases}$$

is called the p.m.f. of X. In this module we will discuss some special discrete probability distributions and will study their properties.

4.1 BERNOULLI EXPERIMENT AND RELATED DISTRIBUTIONS

Let $(\Omega, \mathcal{F}, P)$ be a probability space corresponding to a random experiment $\mathcal{E}$. Each replication of the random experiment E will be called a *trial*. We say that a collection of trials forms a collection of *independent trials* if any collection of corresponding events forms a collection of independent events.

Definition 1.1

(i) The random experiment $\mathcal{E}$ is said to be a *Bernoulli experiment* if its each trial results in just two possible outcomes, labeled as success (S) and failure (F).

(ii) Each replication of a Bernoulli experiment is called a *Bernoulli trial*. ■

Note that, for a Bernoulli experiment $\mathcal{E}$, the sample space is $\Omega = \{S, F\}$, the event space (a sigma-field) is $\mathcal{F} = \mathcal{P}(\Omega) = \{\phi, \Omega, \{S\}, \{F\}\}$ and any function $P: \mathcal{F} \to [0, 1]$, defined by $P(\phi) = 0, P(\Omega) = 1, P(\{S\}) = p$ and $P(\{F\}) = 1 - p$ is a probability measure on $\mathcal{F}$; here $\mathcal{P}(\Omega)$ denotes the power set of Ω and $p \in (0, 1)$ is a fixed constant.

Now suppose that $\mathcal{E}$ is an arbitrary random experiment with corresponding probability space $(\Omega, \mathcal{F}, P)$. In many situations we may not be interested in the whole space $(\Omega, \mathcal{F}, P)$, rather we may be just interested in occurrence or non-occurrence of a given event $E \in \mathcal{F}$. For example consider a sequence of random rolls of a fair dice. In each roll of the dice a person bets on occurrence of upper face with six dots. Let the event of occurrence of upper face with six dots be denoted by E. Here, in each trial, one is only interested in the occurrence or non-occurrence of the event E. In such situations let us label the occurrence of event E by S (success) and its non-occurrence by F (failure). Then there is no need to study the whole space $(\Omega, \mathcal{F}, P)$, rather one may study the restricted space $(\Omega^*, \mathcal{F}^*, P^*)$, where $\Omega^* = \{S, F\}, \mathcal{F}^* = \{\phi, \Omega, \{S\}, \{F\}\}, P^*(\phi) = 0, P^*(\Omega) = 1,$ $P^*(\{S\}) = P(E) = p$ (say) and $P^*(\{F\}) = P(E^c) = 1 - p.$ This leads to the set-up of Bernoulli experiment. In the sequel we will study some of the probability distributions arising out of a sequence of independent Bernoulli trials.

4.1.1 Bernoulli Distribution

Consider a Bernoulli trial with probability space $(\Omega, \mathcal{F}, P)$, where $\Omega = \{S, F\}, \mathcal{F} = \{\phi, \Omega, \{S\}, \{F\}\},$ $P(\{S\}) = p \in (0,1)$ and $P(\{F\}) = 1 - p = q$. Define the r.v. $X: \Omega \to \mathbb{R}$ by

$$X(\omega) = \begin{cases} 1, & \text{if } \omega = S \\ 0, & \text{if } \omega = F \end{cases}$$

$$= \text{ number of successes } (S) \text{ in a Bernoulli experiment.}$$

Then the r.v. X is of discrete type with support $S_X = \{0, 1\}$ and p.m.f.

$$f_X(x) = \mathrm{P}(\{X = x\}) = \begin{cases} q, & \text{if } x = 0 \\ p, & \text{if } x = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} p^x (1-p)^{1-x}, & \text{if } x \in \{0,1\} = S_X \\ 0 & \text{otherwise} \end{cases}. \qquad (1.1)$$

The d.f. of X is given by

$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ q, & \text{if } 0 \le x < 1. \\ 1, & \text{if } x \ge 1 \end{cases}$$

The distribution with p.m.f. (1.1) is called a *Bernoulli distribution* with success probability $p \in (0,1)$. Note that for each $p \in (0,1)$ we get a different Bernoulli distribution and in that sense we have a family of Bernoulli distributions. Various properties of Bernoulli distribution will be discussed in the next subsection where a generalization of Bernoulli distribution will be introduced.

4.1.2 Binomial Distribution

Consider a sequence of independent Bernoulli trials with probability of success (S) in each trial being$p \in (0,1)$. Here we may take the sample space$\Omega = \{(\omega_1, \dots, \omega_n): \ \omega_i \in \{S, F\}, \ i = 1, \dots, n\}$,where, in$(\omega_1, \omega_2, \dots, \omega_n) \in \Omega, \omega_i$ represents the outcome of the i-th Bernoulli trial. Since Ω is finite (has 2^n elements) we may take $\mathcal{F} = \mathcal{P}(\Omega)$. Define the r.v. $X: \Omega \to \mathbb{R}$ by

$$X\big((\omega_1, \dots, \omega_n)\big) = \text{number of S among } \omega_1, \omega_2, \dots, \omega_n$$

$$= \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{I}_{\{S\}}^{(\omega_\mathrm{i})}$$

The r.v. X describes the number of successes in n independent Bernoulli trials.

Clearly, $P(\{X = x\}) = 0$, if $x \notin \{0, 1, \dots, \mathrm{n}\}$. For $m \in \{0, 1, \dots, n\}$

$$P(\{X = m\}) = P(\{(\omega_1, \dots, \omega_n): \ X(\omega_1, \dots, \omega_n) = m\})$$

$$= \sum_{(\omega_1, \dots, \omega_n) \in S_m} P\big((\omega_1, \dots, \omega_n)\big),$$

where $S_m = \{(\omega_1, \dots, \omega_n): m \text{ of } \ \omega_i \text{ s are } S \text{ and remaining } n - m \text{ of } \omega_i \text{ s are } F\}, \ m = 0, 1, \dots, n$. Note that, for $m \in \{0, 1, \dots, n\}$ and $(\omega_1, \dots, \omega_n) \in S_m$,

$$P\big((\omega_1, \dots, \omega_n)\big) = p^m (1 - p)^{n-m},$$

since trials are independent. Moreover, for $m \in \{0, 1, \dots, n\}, S_m$ has $\binom{n}{m}$ elements. Therefore, for $m \in \{0, 1, \dots, n\}$,

$$P(\{X = m\}) = \sum_{(\omega_1, \dots, \omega_n) \in S_m} p^m (1 - p)^{n-m}$$

$$= \binom{n}{m} p^m (1 - p)^{n-m}.$$

It follows that the r.v. X is of discrete type with support $S_X = \{0, 1, \dots, n\}$ and p.m.f.

$$f_X(x) = P(\{X = x\}) = \begin{cases} \binom{n}{x} p^x (1 - p)^{n-x}, & \text{if } x \in S_X = \{0,1, \dots, n\} \\ 0, & \text{otherwise} \end{cases}, \qquad (1.2)$$

where $n \in \{1, 2, \dots, \}, p \in (0, 1)$ and $q = 1 - p$.The probability distribution with p.m.f. (1.2) is called a *Binomial distribution* with n $(\in \mathbb{N})$ trials and success probability $p \in (0, 1)$, and is denoted by $\mathrm{Bin}(n, p)$. We shall use the notation $X \sim \mathrm{Bin}(n, p)$ to indicate that the r.v. X has $\text{Bin}(n,p)$ distribution. Clearly we have a family $\{\text{Bin}(n,p): n \in \mathbb{N}, p \in (0,1)\}$ of binomial distributions corresponding to different choices of $(n,p) \in \mathbb{N} \times (0,1)$. Also, for $p \in (0,1)$, $\text{Bin}(1,p)$ distribution is nothing but a Bernoulli distribution with success probability p.

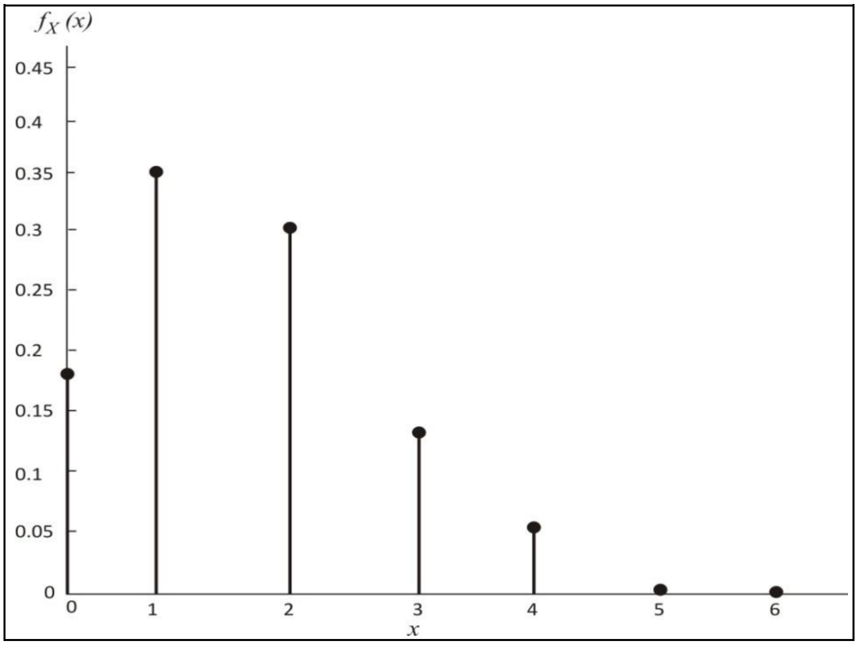

Figure 1.1. Plot of p.m.f. of Bin$(6, \frac{1}{4})$

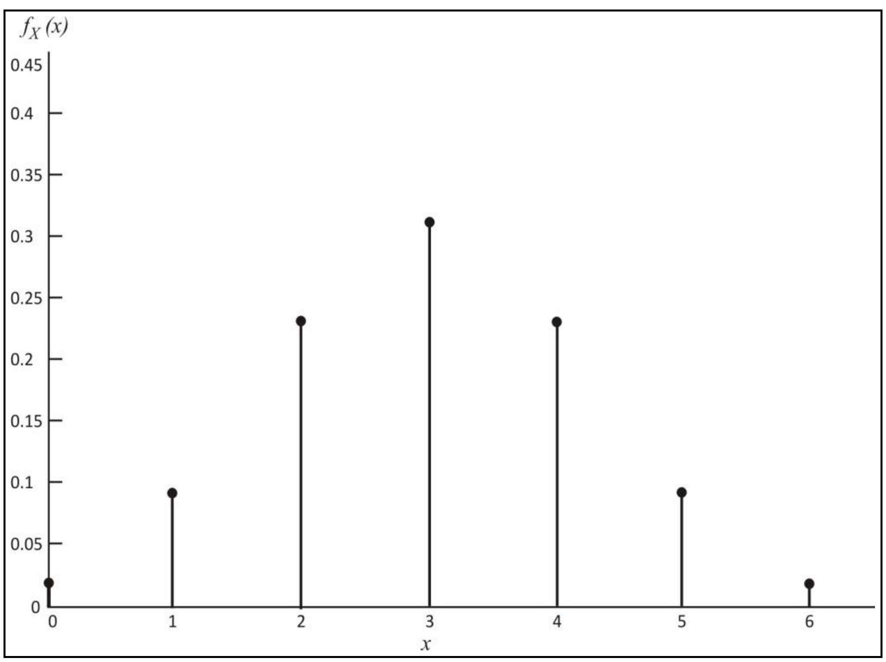

Figure 1.2. Plot of p.m.f. of Bin$(6, \frac{1}{2})$

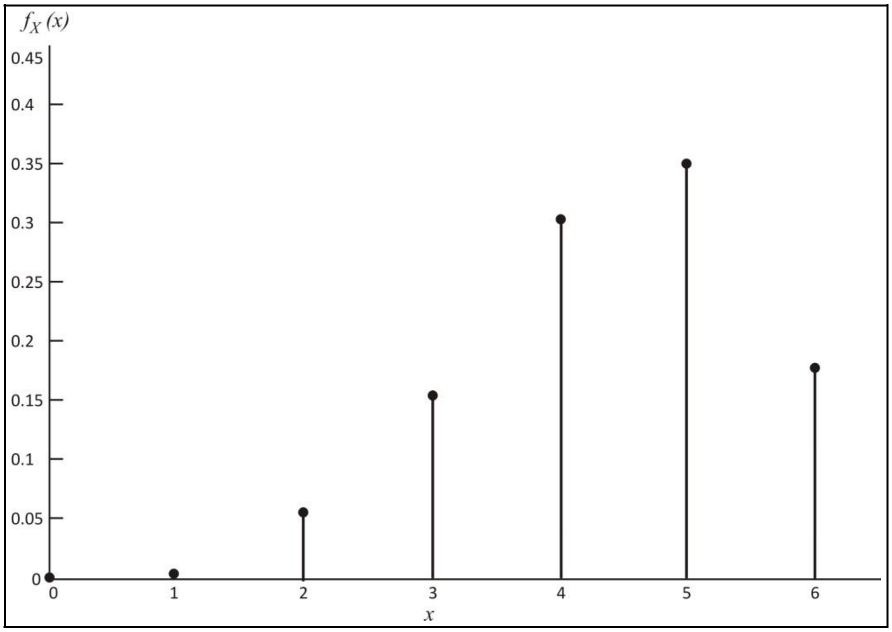

Figure 1.3. Plot of p.m.f. of $\text{Bin}(6, \frac{3}{4})$

Note that

$$\sum_{x \epsilon S_X} f_X(x) = \sum_{x=0}^{\mathrm{n}} \binom{n}{x} p^x (1-p)^{n-x} = (p+1-p)^n = 1.$$

For $r \in \{1,2, \dots, n\}$, define $X_{(r)} = X(X-1)(X-2)\cdots(X-r+1)$. Then, for $r \in \{1, 2, \dots, n\}$,

$$E\left(X_{(r)}\right) = E\big(X(X-1)(X-2)\cdots(X-r+1)\big)$$

$$= \sum_{x=0}^{n} x(x-1)(x-2)\cdots(x-r+1)\binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=r}^{n} \frac{n!}{(x-r)!\ (n-x)!} p^x (1-p)^{n-x}$$

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

$$= n(n-1)\cdots(n-r+1)p^r \sum_{x=0}^{n-r} \binom{n-r}{x} p^x (1-p)^{n-r-x}$$

$$= n(n-1)\cdots(n-r+1)p^r(p+1-p)^{n-r}$$

$$\Rightarrow E\big(X_{(r)}\big) = n(n-1)\cdots(n-r+1)p^r,\ r \in \{1,2,\dots\}.$$

The quantity $E\big(X_{(r)}\big)$ is called the r-th $(r = 1,2,\dots)$ factorial moment of X. We have

$$\boxed{E(X) = E\big(X_{(1)}\big) = np;}$$

$$E(X^2) = E\big(X_{(2)} + X\big) = n(n-1)p^2 + np;$$

$$\boxed{\mathrm{Var}(X) = E(X^2) - \big(E(X)\big)^2 = np(1-p) = npq.}$$

Note that if $X \sim \mathrm{Bin}(n,p)$ then $\mathrm{Var}(X) = npq < np = E(X)$. Thus, for a binomial distribution, the variance is smaller than the mean.

The moment generating function (m.g.f.) of $X \sim \mathrm{Bin}(n,p)$ is given by

$$M_X(t) = E(e^{tX})$$

$$= \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x(1-p)^{n-x}$$

$$= \sum_{x=0}^{n}\binom{n}{x}(pe^t)^x(1-p)^{n-x}$$

$$= (pe^t + 1 - p)^n, \quad t \in \mathbb{R}.$$

Therefore,

$$M_X^{(1)}(t) = npe^t(pe^t + 1 - p)^{n-1}, \quad t \in \mathbb{R};$$

$$M_X^{(2)}(t) = n(n-1)p^2e^{2t}(pe^t + 1 - p)^{n-2} + npe^t(pe^t + 1 - p)^{n-1}, \quad t \in \mathbb{R};$$

$$E(X) = M_X^{(1)}(0) = np;$$

$$E(X^2) = M_X^{(2)}(0) = n(n-1)p^2 + np;$$

and $\mathrm{Var}(X) = E(X^2) - (E(X))^2 = np(1-p)$.

Example 1.1

A fair dice is rolled six times independently. Find the probability that on two occasions we get an upper face with 2 or 3 dots.

Solution. In each roll of the dice, let us label the occurrence of an upper face having 2 or 3 dots as success (S) and occurrence of any other upper face as failure (F). Then we have a sequence of six independent Bernoulli trials with probability of success in each trial as $\frac{1}{3}$. If X denotes the number of occasions on which we get S (i.e., an upper face having 2 or 3 dots) then $X \sim \text{Bin}\left(6, \frac{1}{3}\right)$. Thus the required probability is

$$P(\{X = 2\}) = \binom{6}{2}\left(\frac{1}{3}\right)^2\left(1 - \frac{1}{3}\right)^4 = \frac{80}{243}. \blacksquare$$

Example 1.2

Let $n\ (\geq 2)$ and $r \in \{1, 2, \ldots, n-1\}$ be fixed integers and let $p \in (0, 1)$ be a fixed real number. Using probabilistic arguments show that

$$\sum_{j=r}^{n}\binom{n}{j}p^j(1-p)^{n-j} - \sum_{j=r}^{n-1}\binom{n-1}{j}p^j(1-p)^{n-1-j} = \binom{n-1}{r}p^r(1-p)^{n-r}.$$

Solution. Consider a sequence of independent Bernoulli trials with probability of success in each trial as p. Let X_{n-1} denote the number of successes in the first $n-1$ trials and let X_n denote the number of successes in the first n trials, so that $X_{n-1} \sim \text{Bin}(n-1, p)$, $X_n \sim \text{Bin}(n, p)$ and

$$\sum_{j=r}^{n}\binom{n}{j}p^j(1-p)^{n-j} - \sum_{j=r}^{n-1}\binom{n-1}{j}p^j(1-p)^{n-1-j} = P(\{X_n \geq r\}) - P(\{X_{n-1} \geq r\}). \quad (1.3)$$

Let A_n denote the event that the n-th trial is success so that $P(A_n) = p$. Since the trails are independent, it is evident that the events A_n (an event concerning the n-th trial) and $\{X_{n-1} = r-1\}$ (an event concerning first $n-1$ trials) are independent. Moreover

$$\{X_n \geq r\} = \{X_{n-1} \geq r\} \cup \left\{\{X_{n-1} = r-1\} \cap A_n\right\}.$$

Therefore

$$\begin{aligned} P(\{X_n \geq r\}) &= P(\{X_{n-1} \geq r\}) + P\left(\left\{\{X_{n-1} = r-1\} \cap A_n\right\}\right) \\ &= P(\{X_{n-1} \geq r\}) + P(\{X_{n-1} = r-1\})P(A_n) \\ &= P(\{X_{n-1} \geq r\}) + \left\{\binom{n-1}{r-1}p^{r-1}(1-p)^{n-r}\right\}p, \end{aligned}$$

and the assertion follows on using (1.3). $\blacksquare$

4.1.3 Binomial Distribution and Sampling with Replacement

Suppose that we have a population comprising of N (≥ 2) units out of which a $(\in \{1, 2, \ldots, N-1\})$ are labeled as S (success) and remaining $N-a$ units are labeled as F (failure). Suppose that it is desired to draw a sample of n $(\in \{1, 2, \ldots, N-1\})$ units from this population drawing one unit at a time. Then the probability distribution of X, the number of successes in the drawn sample, may be of interest. Suppose that sampling is done in a manner that the draws are independent (i.e., corresponding events are independent) and after each draw the drawn unit is replaced back into the population. Such a sampling is called *simple random sampling with replacement*. Then we have a sequence of n independent Bernoulii trials with probability of success in each trial as $p = \frac{a}{N}$ and therefore $X \sim \text{Bin}\left(n, \frac{a}{N}\right)$. ■

4.2 NEGATIVE BINOMIAL DISTRIBUTION

Let r be a given positive integer. Suppose that we keep performing independent Bernoulli trials until the r-th success is observed. Further suppose that the probability of success in each trial is $p \in (0,1)$. In this case we may take the sample space $\Omega = \{(\omega_1, \omega_2, \ldots, \omega_n): n \in \{r, r+1, \ldots\}, \omega_n = S, \omega_i \in \{S, F\}, i = 1, \ldots, n-1;\ r-1$ of $\omega_1, \omega_2, \ldots, \omega_{n-1}$ are S and remaining $n-r$ of $\omega_1, \omega_2, \ldots, \omega_{n-1}$ are $F\}$, where an outcome $(\omega_1, \omega_2, \ldots, \omega_n) \in \Omega$ corresponds to one of $\binom{n-1}{r-1}$ ways in which the r-th success is obtained in the n-th Bernoulli trials $(\omega_n = S)$ and the first $n-1$ Bernoulli trials result in $r-1$ successes and $n-r$ failures $(r-1$ of $\omega_1, \omega_2 \ldots, \omega_{n-1}$ are S and remaining $n-r$ of $\omega_1, \omega_2 \ldots, \omega_{n-1}$ are $F)$. Since Ω is countably infinite we may take $\mathcal{F} = \mathcal{P}(\Omega)$. Define the r.v. $X: \Omega \to \mathbb{R}$ by

$$X\big((\omega_1, \ldots, \omega_n)\big) = n - r, \quad (\omega_1, \ldots, \omega_n) \in \Omega$$

$$= \text{number of failures proceeding the } r-\text{th success.}$$

Clearly, for $x \notin \{0, 1, 2, \ldots\}, P(\{X = x\}) = 0$. Also, for $k \in \{0,1,2, \ldots\}$, event $\{X = k\}$ occurs if, and only if, the $(r+k)$ - th trial results in success and, in the first $(r+k-1)$ trials, $(r-1)$ successes and k failures are observed. Since the trials are independent, for $k \in \{0,1,2, \ldots\}$, we have

$$P(\{X = k\}) = p_1 p_2,$$

where p_1 is the probability of observing $(r-1)$ successes in the first $(r+k-1)$ independent Bernoulli trials and p_2 is the probability of getting the success on the $(r+k)$ - th trial. Clearly $p_2 = p$, and using the property of binomial distribution

$$p_1 = \binom{r+k-1}{r-1} p^{r-1} (1-p)^k.$$

Therefore, for $k \in \{0,1,2,\dots\}$,

$$P(\{X=k\}) = \binom{r+k-1}{r-1} p^r (1-p)^k.$$

Thus the r.v. X is of discrete type with support $S_X = \{0,1,2,\dots\}$ and p.m.f.

$$f_X(x) = P(\{X=x\}) = \begin{cases} \binom{r+x-1}{r-1} p^r q^x, & \text{if } x \in \{0,1,2,\dots\} \\ 0, & \text{otherwise} \end{cases}, \tag{1.4}$$

where $q = 1-p$.

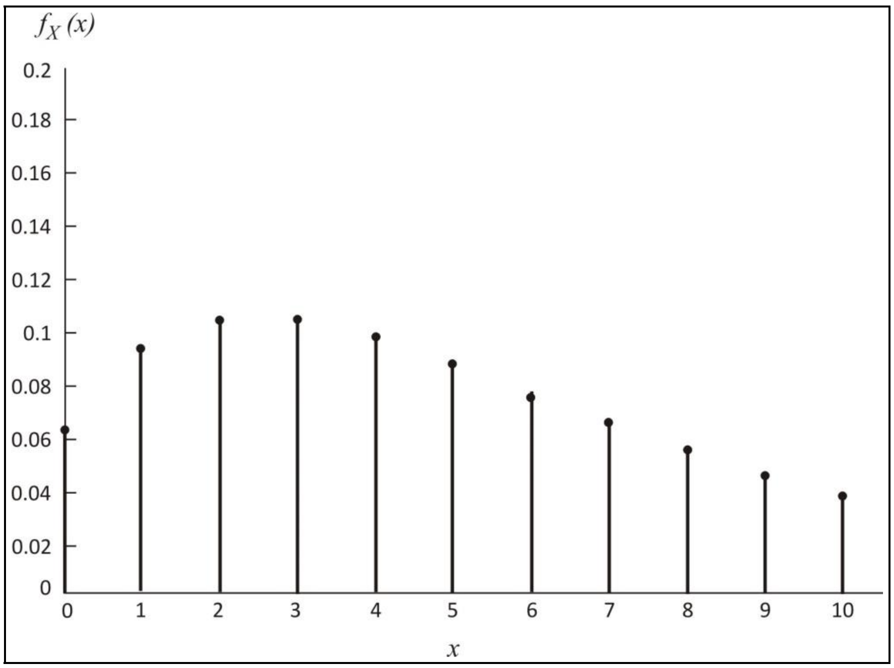

Figure 1.4. Plot of p.m.f. of NB $(2, \frac{1}{4})$

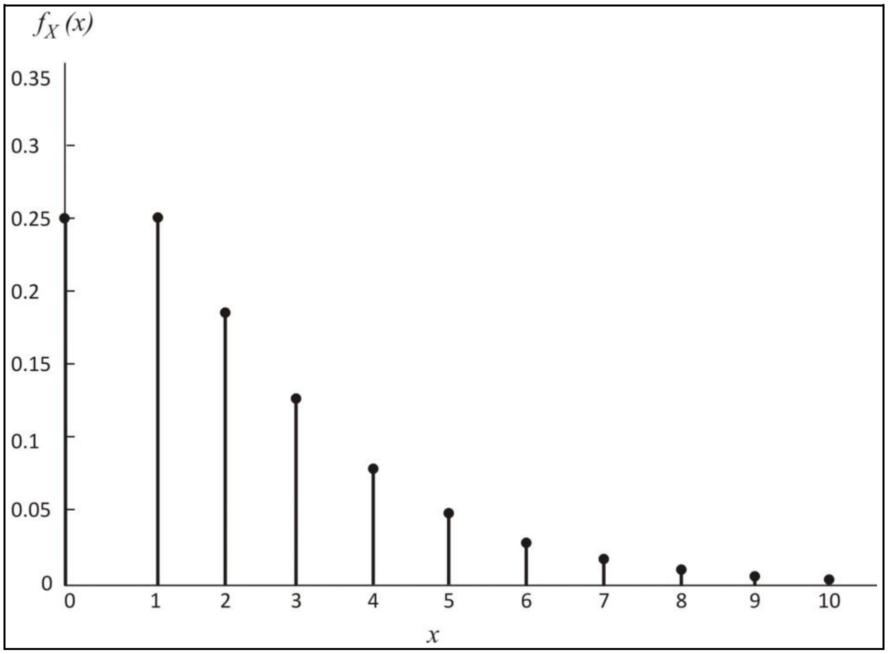

Figure 1.5. Plot of p.m.f. of NB $(2, \frac{1}{2})$

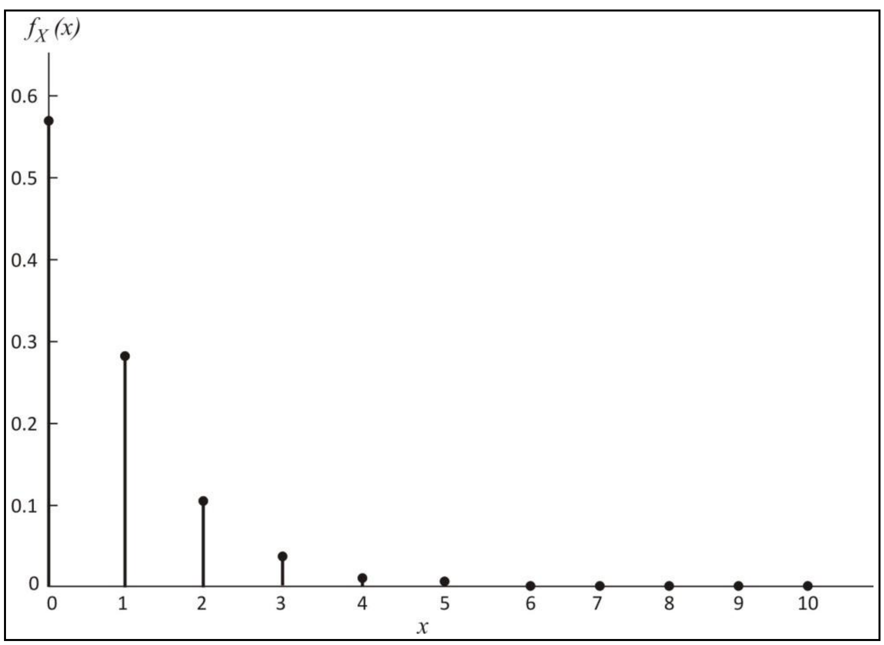

Figure 1.6. Plot of p.m.f. of NB $(2, \frac{3}{4})$

The probability distribution with p.m.f. (1.4) is called a *Negative Binomial distribution* with $r \in \{1, 2, \dots\}$ successes and success probability $p \in (0,1)$, and is denoted by $\mathrm{NB}(r,p)$. Notation $X \sim \mathrm{NB}(r,p)$ will be used to indicate that the r.v. X follows a negative binomial distribution with r successes and success probability p. Using the ratio test it is easy to verify that the series $\sum_{x=0}^{\infty}\binom{r+x-1}{r-1}t^x$ is absolutely convergent for $t\in(-1,1)$. For $t\in(-1,1)$

$$\sum_{x=0}^{\infty}\binom{r+x-1}{r-1}t^x = 1+\sum_{x=1}^{\infty}\frac{(r+x-1)(r+x-2)\cdots(r+1)r}{x!}t^x$$

$$= 1+rt+\frac{(r+1)r}{2!}t^2+\frac{(r+2)(r+1)r}{3!}t^3+\cdots$$

$$= (1-t)^{-r}. \qquad (1.5)$$

It follows that, for each $r\in\{1,2,\ldots,\}$ and $p\in(0,1)$,

$$\sum_{x\in S_X} f_X(x) = p^r\sum_{\mathrm{x}=0}^{\infty}\binom{r+x-1}{r-1}(1-p)^x = p^r(1-(1-p))^{-r} = 1.$$

Clearly we have a family $\{\mathrm{NB}(r,p): r\in\mathbb{N}, p\in(0,1)\}$ of negative binomial distributions corresponding to different choices of $(r,p)\in\mathbb{N}\times(0,1)$.

For $m\in\{1,2,\ldots\}$, the m-th factorial moment of X is given by

$$E\big(X_{(m)}\big) = E(X(X-1)\ldots(X-m+1))$$

$$= \sum_{\mathrm{x}=0}^{\infty}x(x-1)\cdots(x-m+1)\binom{r+x-1}{r-1}p^r(1-p)^x$$

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

$$= p^r\sum_{x=0}^{\infty}\frac{(r+x+m-1)!}{(r-1)!\,x!}(1-p)^{x+m}$$

$$= r(r+1)\cdots(r+m-1)p^r(1-p)^m\sum_{x=0}^{\infty}\binom{r+m+x-1}{r+m-1}(1-p)^x$$

$$= r(r+1)\cdots(r+m-1)p^r(1-p)^m\big(1-(1-p)\big)^{-(r+m)}$$

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

Therefore,

$$\boxed{E(X) = E\big(X_{(1)}\big) = \frac{r(1-p)}{p} = \frac{rq}{p};}$$

$$E(X^2) = E\big(X_{(2)} + X\big) = E\big(X_{(2)}\big) + E(X) = r(r+1)\left(\frac{q}{p}\right)^2 + \frac{rq}{p} = \frac{rq(rq+1)}{p^2};$$

$$\boxed{\text{Var}\,(X) = E(X^2) - \big(E(X)\big)^2 = \frac{r\text{q}}{p^2}.}$$

Note that if $X \sim \text{NB}(r, p)$ then

$$\text{Mean} = E(X) = \frac{rq}{p} < \frac{rq}{p^2} = \text{Var}\,(X),$$

i.e., for negative binomial distribution the mean is smaller than the variance.

The m.g.f. of $X \sim \text{NB}(r, p)$ is given by

$$M_X(\text{t}) = E(e^{tX})$$

$$= p^{\text{r}} \sum_{\text{x}=0}^{\infty} \binom{r+x-1}{r-1} ((1-p)e^t)^x$$

$$= p^{\text{r}}(1-(1-p)e^t)^{-\text{r}}, \quad |(1-p)e^t| < 1 \ \ (\text{using } (1.5))$$

$$= \left(\frac{p}{1-qe^t}\right)^{\text{r}}, \qquad t < -\ln q.$$

An $\text{NB}(1, p)$ distribution is called a *geometric distribution* with success probability p and is denoted by $\text{Ge}(p)$. Clearly, if $Y \sim \text{Ge}(p)$ then Y denotes the number of failures preceding the first success in a sequence of independent Bernoulli trials. The p.m.f. of $Y \sim \text{Ge}(p)$ is given by

$$f_Y(\text{y}) = P(\{Y = y\}) = \begin{cases} pq^y, & \text{if } y \in \{0, 1, 2, \ldots\} \\ 0, & \text{otherwise} \end{cases},$$

where $q = 1 - p$.

Since, for $k \in \{0, 1, 2, \ldots\}$, $\sum_{y=0}^{k} f_Y(y) = p \sum_{y=0}^{k} q^y = 1 - q^{k+1}$, the d.f. of $Y \sim \text{Ge}(p)$ is given by

$$F_Y(y) = P(\{Y \le y\}) = \begin{cases} 0, & \text{if } \ y < 0 \\ 1 - q^{k+1}, & \text{if } \ k \le y < k+1, \ \ k = 0, 1, 2, \ldots \end{cases}.$$

Note that if $Y \sim \mathrm{Ge}(p)$ then, for $m, n \in \{0,1,2,\dots\}$,

$$P(\{Y \geq m\}) = p \sum_{x=m}^{\infty} q^x = q^m,$$

and

$$\begin{aligned} P(\{Y \geq m+n\} | \{Y \geq m\}) &= \frac{P(\{Y \geq m+n,\ Y \geq m\})}{P(\{Y \geq m\})} \\ &= \frac{P(\{Y \geq m+n\})}{P(\{Y \geq m\})} \\ &= \frac{q^{m+n}}{q^m} \\ &= q^n \\ &= P(\{Y \geq n\}). \end{aligned}$$

It follows that if $Y \sim \mathrm{Ge}(p)$ then, for $m, n \in \{0,1,2,\dots\}$,

$$P(\{Y \geq m+n\} | \{Y \geq m\}) = P(\{Y \geq n\}) \qquad (1.6)$$

or equivalently

$$P(\{Y \geq m+n\}) \geq P(\{Y \geq m\}) P(\{Y \geq n\}).$$

Remark 1.2

The property (1.6) possessed by a geometric distribution has an interesting interpretation. Suppose that a system can fail only at discrete time points $0, 1, 2, \dots$ and let its lifetime be denoted by a discrete type r.v. T, having the support $S_T = \{0, 1, 2, \dots\}$.Then, for $m, n \in \{0,1,2,\dots\}$, $P(\{T \geq m+n\} | \{T \geq m\})$ represents the conditional probability that a system of age m or more will survive at least n additional units of time, and $P(\{T \geq n\}) = P(\{T \geq n\} | \{T \geq 0\})$ represents the probability that a fresh system (of age 0) will survive at least n units of time. Thus if the probability distribution of a r.v. T (representing the lifetime of a system) satisfies property (1.6) then the age of the system has no effect on the residual (remaining) life of the system (implying that an used system is as good as a new system). This property of a probability distribution (or random variable) is known as the *lack of memory property*. ■