Common Discrete Distributions

5.1 Discrete Uniform Distribution

If the random variable X assumes the values $x_1, x_2, ..., x_k$, with equal probabilities, then the discrete uniform distribution is given by:

$$f(x;k) = \frac{1}{k}, \text{ for } x = x_1, x_2, ..., x_k.$$

Note: We have used the notation $f(x;k)$ instead of $f(x)$ to indicate that the uniform discrete distribution depends on the parameter k.
Example: When a die is tossed, each element of the sample space $S = \{1,2,3,4,5,6\}$ occurs with probability $\frac{1}{6}$. Therefore, we have a discrete uniform distribution with:

$$f(x;6) = \frac{1}{6}, \text{ for } x = 1,2,3,4,5,6.$$

Example: Suppose that a student is selected at random from a class of 25 to participate in a certain seminar. Each student has the same probability of $\frac{1}{25}$ of being selected. If we assume that the students have been numbered in some way from 1 to 25, the distribution is uniform with:

$$f(x;25) = \frac{1}{25}, \text{ for } x = 1,2,...,25.$$

5.2 The Bernoulli Distribution

Suppose that a trial, or an experiment, whose outcome can be classified as either a 'success' or as a 'failure' is performed. If we let $X = 1$ when the outcome is a success and $X = 0$ when it is a failure, then the probability mass function of X is given by:

$$P(X = 0) = 1 - p \quad \text{and} \quad P(X = 1) = p \tag{5.1}$$

where p is the probability that the trial is a "success".

A random variable X is said to be a **Bernoulli** random variable (after the Swiss mathematician James Bernoulli) if its probability mass function is given by Equation (5.1) for some $p \in (0,1)$. If we let $q = 1 - p$, the probability mass function of X can as well be written as:

$$f(x) = \begin{cases} p^x q^{1-x} & \text{for} \quad x = 0,1 \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

Its expected value is

$$E(X) = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = 1 \cdot p + 0 \cdot q = p$$

That is, the expectation of a Bernoulli random variable is the probability that the random variable assumes a value of 1 (success). Also,

$$E(X^2) = 1^2 \cdot p + 0^2 \cdot q = p$$

and

$$Var(X) = E(X^2) - [E(X)]^2 = pq$$

If the random variables in an infinite sequence $X_1, X_2, \ldots$ are independent and identically distributed (i.i.d), and if each random variable X_i has a Bernoulli distribution with parameter p, then it is said that the random variables $X_1, X_2, \ldots$ form an infinite sequence of **Bernoulli trials** with parameter p. Similarly, if n random variables $X_1, X_2, \ldots, X_n$ are i.i.d and each has a Bernoulli distribution with parameter p, then it is said that the variables $X_1, X_2, \ldots, X_n$ form n **Bernoulli trials with parameter** p.

For example, suppose that a fair coin is tossed repeatedly. Let $X_i = 1$ if a head is obtained on the ith toss, and let $X_i = 0$ if a tail is obtained ($i = 1, 2, \ldots$). Then the random variables $X_1, X_2, \ldots$ form an infinite sequence of Bernoulli trials with parameter $p = \frac{1}{2}$. Similarly, suppose that 10% of the items produced by a certain machine are defective, and that n items are selected at random and inspected. In this case, let $X_i = 1$ if the ith item is defective, and let $X_i = 0$ if it is nondefective (i=1,. . . , n). Then the variables $X_1, X_2, \ldots, X_n$ form n Bernoulli trials with parameter $p = \frac{1}{10}$.

5.3 The Binomial Distribution

Suppose now that n independent trials, each of which results in a 'success' with probability p and in a 'failure' with probability 1 - p, are to be performed. If X represents the number of successes that occur in the n trials, then X is said to be a binomial random variable with parameters (n, p). This result can be restated as follows:

If the random variables $X_1, X_2, \ldots, X_n$ form n Bernoulli trials with parameter p and if $X = X_1 + X_2 + \cdots + X_n$, then X has a binomial distribution with parameters n and p.

The probability mass function of a binomial random variable with parameters n and p is given by:

$$P(X = x) = \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1-p)^{n-x}$$

for $x = 0, 1, 2, ..., n$.

where $\left(\begin{array}{c} n \\ x \end{array} \right) = \dfrac{n!}{x!(n-x)!} =^n C_x$ with n being the number of trials, x the number of successes in the n trials and p the probability of success on a single trial.

If X is a random variable that is binomially distributed with parameters n and p, we write $X \sim binomial(n, p)$. Note that, by the binomial theorem, the probabilities sum to 1, that is,

$$\sum_{x=0}^{n} P(X = x) = \sum_{x=0}^{n} \left(\begin{array}{c} n \\ x \end{array} \right) p^x q^{n-x} = (p+q)^n = 1$$

Example 1 *A coin is tossed 5 times. Find the probability that exactly 3 heads will appear.*
Soln: Let X be the random variable representing the number of heads that will appear. $X \sim binomial(n, p)$, where $n = 5$ and $p = \frac{1}{2}$.

$$P(X = x) = \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1-p)^{n-x}, \quad x = 0, 1, 2, 3, 4, 5$$

Thus

$$P(exactly\ three\ heads) = P(X = 3) = \left(\begin{array}{c} 5 \\ 3 \end{array} \right) \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = \frac{5}{16}$$

5.3.1 Reading Binomial Tables

Frequently, you will be interested in problems where you need to find probabilities such as $P(X < a), P(a \leq X \leq b)$ *or* $P(X > b)$. To obtain desired probabilities from tables, use the following definitions.

1. $P(X \leq a) = \sum_{x=0}^{a} P(X = x)$

2. $P(X < a) = \sum_{x=0}^{a-1} P(X = x)$

3. $P(X > a) = 1 - P(X \leq a)$

4. $P(a \leq X \leq b) = P(X \leq b) - P(X < a)$

5. $P(a < X \leq b) = P(X \leq b) - P(X \leq a)$

Example 2 *The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted the disease, what is the probability that:*

(i) *Exactly 5 survive,*

(ii) *At least 10 survive the disease,*

(iii) *From 3 to 8 survive the disease.*

Soln:
Let X be the number of people that survive. $X \sim binomial(n, p)$, *where* $n = 15$ *and* $p = 0.4$ *and*

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, 15$$

(i) $P(X = 5) = ^{15}C_5 (0.4)^5 (0.6)^{10} = 0.1859$

(ii) $P(X \geq 10) = 1 - P(X < 10) = 1 - P(X \leq 9) =$
$1 - (P(X = 0) + P(X = 1) + P(X = 2) + \ldots + P(X = 9))$

(iii) $P(3 \leq X \leq 8) = P(X \leq 8) - P(X < 3) = P(X \leq 8) - P(X \leq 2)$

We can use tables to answer this question as well, look at the Cumulative Binomial table at the back.

(i) $P(X = 5) = \sum_{x=0}^{5} b(x; 15, 0.4) - \sum_{x=0}^{4} b(x; 15, 0.4) = 0.4032 - 0.2173 = 0.1859$

(ii) $P(X \geq 10) = 1 - P(X < 10) = 1 - \sum_{x=0}^{9} b(x; 15, 0.4) = 1 - 0.9662 = 0.0338$

(iii) $P(3 \leq X \leq 8) = P(X \leq 8) - P(X < 3) = P(X \leq 8) - P(X \leq 2) =$
$\sum_{x=0}^{8} b(x; 15, 0.4) - \sum_{x=0}^{2} b(x; 15, 0.4) = 0.9050 - 0.0271 = 0.8779$

Example 3 *A fair coin is tossed 14 times. What is the probability that the head will appear:*

(a) *at least 8 times,*

(b) *at most 6 times,*

(c) *exactly 5 times,*

(d) *less than 12 times,*

(e) *less than 9 but greater than 4 times.*

Example 4 *A new car dealer knows from past experience that on the average, she will make sale to about 20% of her customers. What is the probability that in five randomly selected presentations, she makes a sale to:*

(a) exactly three customers,

(b) at most one customer,

(c) at least 2 customers.

You have $p = 20\% = 0.2, \quad n = 5$. *Therefore,*

(a) $P(X = 3) = P(X \leq 3) - P(X \leq 2) = 0.9933 - 0.9421 = 0.0512$

(b) $P(X \leq 1) = 0.7373$

(c) $P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1) = 1 - 0.7373 = 0.2627$

5.3.2 Expectation of Binomial Distribution

Definition 1 *The mean or expectation of a binomially distributed random variable X is*

$$E(X) = np$$

where n is the number of repeated trials and p, $(0 \leq p \leq 1)$ *is the probability of success.*

Proof: The first method is by using the meaning of a Bernoulli trial: If the random variables $X_1, X_2, \ldots, X_n$ form n Bernoulli trials with parameter p and if $X = X_1 + X_2 + \cdots + X_n$, then X has a binomial distribution with parameters n and p.
Recall that $E(X_i) = p$ for $i = 1, 2, \ldots, n$. Thus

$$\begin{aligned} E(X) =& E(X_1 + X_2 + \cdots + X_n) \\ =& E(X_1) + E(X_2) + \cdots + E(X_n) \\ =& \underbrace{p + p + \cdots + p}_{\text{n times}} = np \end{aligned}$$

The second method is by use of the definition of $E(X)$.

$$\begin{aligned}
E(X) &= \sum_{x=0}^{n} xP(X = x) \\
&= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} \\
&= \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x q^{n-x} \\
&= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x} \\
&= np \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x-1} q^{n-x} \\
&= np(p+q)^{n-1} = np
\end{aligned}$$

5.3.3 The Variance of Binomial Distribution

Definition 2 *The variance of a binomially distributed random variable X is given by $Var(X) = npq$.*

This can be proved in two different ways. First from the definition of a Bernoulli trial. Recall that for a Bernoulli random variable X_i, $Var(X_i) = pq$. Since a binomial random variable $X = X_1 + X_2 + \cdots + X_n$, then

$$\begin{aligned}
Var(X) &= Var(X_1 + X_2 + \cdots + X_n) \\
&= Var(X_1) + Var(X_2) + \cdots + Var(X_n) \\
&= \underbrace{pq + pq + \cdots + pq}_{\text{n times}} = npq
\end{aligned}$$

Thus $Var(X) = npq$ and the standard deviation of X is $\sqrt{npq}$.
In the second method, you use the definition of variance. Recall that $E(X) = np$ and from $Var(X) = E(X^2) - [E(X)]^2$, we seek to find $E(X^2)$. Also recall from the Binomial expansion that

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

Now

$$\begin{aligned}
E(X(X-1)) =& \sum_{x=0}^{n} x(x-1)P(X=x) \\
=& \sum_{x=0}^{n} x(x-1)\begin{pmatrix} n \\ x \end{pmatrix} p^x q^{n-x} \\
=& \sum_{x=0}^{n} x(x-1)\frac{n!}{x!(n-x)!} p^x q^{n-x} \\
=& \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x q^{n-x} \\
=& n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} q^{n-x} \\
=& n(n-1)p^2 \sum_{x=2}^{n} \begin{pmatrix} n-2 \\ x-2 \end{pmatrix} p^{x-2} q^{n-x} \\
=& n(n-1)p^2 (p+q)^{n-2} \\
=& n(n-1)p^2
\end{aligned}$$

But

$$E(X(X-1)) = E(X^2 - X) = E(X^2) - E(X)$$

Thus

$$E(X^2) = E(X(X-1)) + E(X) = n(n-1)p^2 + np$$

Therefore, 7

$$\begin{aligned}
Var(X) =& E(X^2) - [E(X)]^2 \\
=& n(n-1)p^2 + np - (np)^2 \\
=& n^2p^2 - np^2 + np - n^2p^2 \\
=& np - np^2 \\
=& np(1-p) = npq
\end{aligned}$$

Example 5 *In a family the probability of having a boy is 0.8. If there are seven children in the family determine the:*

(a) expected number of boys,

(b) expected number of girls,

(c) probability that at least four children are girls,

(d) probability that they are all boys.

***Solution**:*

(a) *let X be the random variable representing the number of boys.* *Then* $X \sim binomial(n, p)$ *with* $n = 7$ *and* $p = 0.8$. $E(X) = np = 7 * 0.8 = 5.6$

(b) *let Ybe the random variable representing the number of girls.* *Then* $Y \sim binomial(n, p)$ *with* $n = 7$ *and* $p = 0.2$. $E(Y) = 7 * 0.2 = 1.4$

(c) $P(Y \geq 4) = 1 - P(Y < 4) = 1 - P(Y \leq 3) = 1 - 0.9667 = 0.0333$

(d) *P(all are boys)*$=P(X = 7) = 0.2097$

Example 6 *A multiple choice quiz has ten questions each with four alternative answers of which one is correct. Determine the expected number of correct answers and the variance, if someone does the quiz by sheer guesswork.*
Solution:
Let X be the random variable representing the number of correct answers. Then $X \sim binomial(n, p)$ *with* $n = 10$ *and* $p = \frac{1}{4}$.

$$E(X) = np = 10 * \frac{1}{4} = 2.5$$

and

$$Var(X) = npq = 10 * 1/4 * 3/4 = 1.875$$

Example 7 *A multiple-choice exam consist of twenty questions each with five alternative answers of which only one is correct. Five marks are awarded for each answer and one mark is subtracted for each incorrect or un attempted question. If the exam is answered by sheer guesswork, find the:*

(i) *expected number of correct and incorrect answers,*

(ii) *expected overall mark,*

(iii) *variance of the mark in (ii)*

Solution:
Let X be the random variable representing the number of correct answers and Y the number of incorrect answers. Then, with $n = 20, p = \frac{1}{5} = 0.2, q = \frac{4}{5} = 0.8$

(i) $E(X) = np = 20 * \frac{1}{5} = 4$ *correct answers*
$E(Y) = nq = 20 * \frac{4}{5} = 16$ *wrong answers*

(ii) *The expected mark for correct answers* $=4*5 = 20$. *The expected mark for incorrect answers* $=16 * 1 = 16$. *Therefore, the expected overall mark is* $20 - 16 = 4$ *marks.*

(iii) $Var(X) = npq = 4 * \frac{1}{4} * \frac{3}{4} = 0.75$

Example 8 *If X is a binomially distributed with* $p = 0.45$ *and* $n = 10$, *find the most likely value of X to occur.*
Solution:
First compute the expected value of X, that is $E(X) = np = 10 * 0.45 = 4.5$. *When this*

is done, try to find the probabilities of values around the mean, in this case $3, 4, 5, 6$ *as below.*

$$\begin{aligned}
P(X=3) &= \binom{10}{3}(0.45)^3(0.55)^7 = 0.1664 \\
P(X=4) &= \binom{10}{4}(0.45)^4(0.55)^6 = 0.2383 \\
P(X=5) &= \binom{10}{5}(0.45)^5(0.55)^5 = 0.2340 \\
P(X=6) &= \binom{10}{6}(0.45)^6(0.55)^4 = 0.1595
\end{aligned}$$

From this you note that $X = 4$ *has the highest probability and as such the most occuring value, or the mode is 4.*

Example 9 *In Iganga, 80% of the inhabitants have a particular eye disorder. If 12 people are waiting to see the doctor, what is the most likely number of them to have the eye disorder?*
Solution
This is a binomial distribution with X=the number of people with eye disorder, $n = 12$ *and* $p = 0.8$*. First compute* $E(X) = np = 12 * 0.8 = 9.6$*. This implies that you find probabilities of values of X around the mean, say* $8, 9, 10, 11$*.*

$$\begin{aligned}
P(X=8) &= \binom{12}{8}(0.80)^8(0.20)^4 = 0.1328 \\
P(X=9) &= \binom{12}{9}(0.80)^9(0.20)^3 = 0.2362 \\
P(X=10) &= \binom{12}{10}(0.80)^{10}(0.20)^2 = 0.2834 \\
P(X=11) &= \binom{12}{11}(0.45)^{11}(0.20)^1 = 0.2061
\end{aligned}$$

Thus the most likely number of people with eye disorder is 10 since it has the highest probability.

Example 10 *Among companies doing highway or bridge construction, 80% test employees for substance abuse. A study involves the random selection of 10 such companies. What is the mean and standard deviation for the number of companies in this sample that test for substance abuse.*
Solution:
This is binomial, since there are 2 outcomes (test or not test), a fixed number of trials (10) and independence. So $\mu = 10*0.8 = 8, \sigma^2 = 10*0.8*0.2 = 1.6,$ *and* $\sigma = \sqrt{1.6} = 1.2649$

Example 11 *30% of college students own video cassette recorders. The Telektronic Company produced a videotape and sent pilot copies to 20 college students. What is the mean and standard deviation of this sample?*
Solution:
Once again -this is binomial, so $\mu = 20 * 0.3 = 6$, $\sigma^2 = 20 * 0.3 * 0.7 = 4.2$, *and* $\sigma = \sqrt{4.2} = 2.04939$.

Example 12 *In the testing of a product from a production line, the probability that the product survives the test is* $\frac{3}{4}$. *What is the probability that exactly two of the next four products tested will also survive the test?*
Solution
Assuming that the tests are independent, then $X \sim binomial(4, \frac{3}{4})$. *Therefore*

$$P(X = 2) = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 = \frac{27}{128}.$$

5.4 The Poisson Distribution

A random variable X, taking on one of the values $0, 1, 2, \ldots$, is said to be a Poisson random variable with mean λ, $\lambda > 0$, if its probability mass function is given by

$$f(x, \lambda) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots \tag{5.3}$$

The symbol e stands for a constant approximately equal to 2.7183. It is a famous constant in mathematics, named after the Swiss mathematician L. Euler, and it is also the base of the so-called natural logarithm.

It is clear that $f(x) \geq 0$ for each value of x. In order to verify that the function $f(x, \lambda)$ defined in equation (5.3) satisfies the requirements of every p.m.f , it must be shown that $\sum_{x=0}^{\infty} f(x, \lambda) = 1$. It is known from elementary calculus that for any real number λ,

$$e^{\lambda} = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$$

Therefore,

$$\sum_{x=0}^{\infty} f(x, \lambda) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1$$

Examples of Poisson random variables are:

- The number of misprints on a page (or a group of pages) of a book.
- The number of people in a community living to 100 years of age.

- The number of wrong telephone numbers that are dialed in a day.
- The number of transistors that fail on their first day of use.
- The number of customers entering a post office on a given day.
- The number of α-particles discharged in a fixed period of time from some radioactive particle.

Mean and Variance

We have stated that the distribution for which the p.m.f is given by Eq. (5.3) is defined to be the Poisson distribution with mean λ. In order to justify this definition, we must show that λ is, in fact, the mean of this distribution. The mean $E(X)$ is specified by the following infinite series:

$$E(X) = \sum_{x=0}^{\infty} x f(x, \lambda)$$

since the term corresponding to $x = 0$ in this series is 0, we can omit this term and can begin the summation with the term $x = 1$. Therefore,

$$E(X) = \sum_{x=1}^{\infty} x f(x, \lambda) = \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}$$

Now letting $y = x - 1$ in this summation, we obtain

$$E(X) = \lambda \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!} = \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda \cdot 1 = \lambda$$

The variance of a Poisson distribution can be found by a technique similar to the one that has just been given. We begin by considering the following expectation:

$$\begin{aligned} E[X(X-1)] &= \sum_{x=0}^{\infty} x(x-1) f(x, \lambda) = \sum_{x=2}^{\infty} x(x-1) f(x, \lambda) \\ &= \sum_{x=2}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{x!} = \lambda^2 \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!} \end{aligned}$$

If we let $y = x - 2$, we obtain

$$E[X(X-1)] = \lambda^2 \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!} = \lambda^2 \tag{5.4}$$

Since $E[X(X-1)] = E(X^2) - E(X) = E(X^2) - \lambda$, it follows from Eq. (5.4) that $E(X^2) = \lambda^2 + \lambda$. Therefore

$$Var(X) = E(X^2) - [E(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Thus, for a Poisson distribution for which the p.m.f is defined by Eq. (5.3), we have established the fact that both the mean and variance are equal to λ.

Example 13 *Suppose that the average number of accidents occurring weekly on a particular stretch of a highway equals 3. Calculate the probability that there is at least one accident this week.*
Soln: *Let X denote the number of accidents occurring on the stretch of highway in question during this week. Now, $X \sim Poisson(\lambda)$, where $\lambda = 3$.*

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{e^{-3}3^0}{0!} = 1 - e^{-3} = 0.9502$$

Example 14 *Consider an experiment that consists of counting the number of α particles given off in a one-second interval by one gram of radioactive material. If we know from past experience that, on the average, 3.2 such α-particles are given off, what is a good approximation to the probability that no more than 2 α-particles will appear?*
Soln: *The number of α-particles given off will be a Poisson random variable with parameter $\lambda = 3.2$. Hence the desired probability is*

$$P(X \leq 2) = e^{-3.2} + 3.2e^{-3.2} + \frac{(3.2)^2}{2}e^{-3.2} = 0.380$$

Example 15 *If the average number of claims handled daily by an insurance company is 5, what proportion of days have less than 3 claims?*
Soln: *Suppose that the number of claims handled daily, call it X, is a Poisson random variable. Since $E(X) = 5$, the probability that there will be fewer than 3 claims on any given day is*

$$P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2) = e^{-5} + e^{-5}\frac{5^1}{1!} + e^{-5}\frac{5^2}{2!} = 0.1247$$

Example 16 *The average number of field mice per acre of a field is estimated to be 10. Find the probability that a given acre will contain more than 15 mice.*
Soln: *Let X be the number of field mice, $X \sim Poisson(10)$.*

$$P(X > 15) = 1 - P(X \leq 15)$$

Note: *We can use cumulative Poisson table to get the answer easily. Look at the table at end of this Topic.*
$P(X > 15) = 1 - P(X \leq 15) = 1 - 0.9513 = 0.0487$.

Example 17 *Suppose that radioactive particles strike a certain target in accordance with a Poisson process at an average rate of 3 particles per minute. Determine the probability that 10 or more particles will strike the target in a particular 2-minute period.*
Soln: *In a Poisson process, the number of particles striking the target in any particular one-minute period has a Poisson distribution with mean λ. since the mean number of strikes in any one-minute period is 3, it follows that $\lambda = 3$. Therefore, the number of strikes X in any 2-minute period will have a Poisson distribution with mean 6. It can be found from the table of the Poisson distribution that $P(X \geq 10) = 0.0838$.*

Example 18 *Telephone calls enter a school switchboard on average of two every three minutes. What is the probability of at least 5 calls arriving in a 9-minute period?*
Soln: *Let X be the number of calls arriving in a 9-minute period. $X \sim Poisson(\lambda)$, where $\lambda = 6$.*

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.285 = 0.715$$

The Distribution of two Independent Poisson Random Variables

Theorem 1 *The sum of two independent Poisson variables with parameters λ_1 and λ_2 respectively is a Poisson random variable with parameter $\lambda_1 + \lambda_2$.*

Proof 1 *If $X \sim Poisson(\lambda_1)$, then $P(X = x) = \dfrac{e^{-\lambda_1}\lambda_1^x}{x!}$ so that*

$$\begin{aligned} P(X = 0) &= e^{-\lambda_1} \\ P(X = 1) &= e^{-\lambda_1} \cdot \lambda_1 \\ P(X = 2) &= \frac{e^{-\lambda_1} \cdot \lambda_1^2}{2!} \end{aligned}$$

Also if $Y \sim Poisson(\lambda_2)$, then $P(Y = y) = \dfrac{e^{-\lambda_2}\lambda_2^y}{y!}$ so that

$$\begin{aligned} P(Y = 0) &= e^{-\lambda_2} \\ P(Y = 1) &= e^{-\lambda_2} \cdot \lambda_2 \\ P(Y = 2) &= \frac{e^{-\lambda_2} \cdot \lambda_2^2}{2!} \end{aligned}$$

If X and Y are two independent Poisson random variables, then you will have

$$\begin{aligned} P(X + Y = 0) &= P(X = 0) \cdot P(Y = 0) = e^{-\lambda_1} \cdot e^{-\lambda_2} = e^{-(\lambda_1+\lambda_2)} \\ P(X + Y = 1) &= P(X = 0) \cdot P(Y = 1) + P(X = 1) \cdot P(Y = 0) \\ &= e^{-\lambda_1} \cdot e^{-\lambda_2} \cdot \lambda_2 + e^{-\lambda_1} \cdot e^{-\lambda_2} \cdot \lambda_1 = e^{-(\lambda_1+\lambda_2)} \cdot (\lambda_1 + \lambda_2) \\ P(X + Y = 2) &= P(X = 0) \cdot P(Y = 2) + P(X = 2) \cdot P(Y = 0) + P(X = 1) \cdot P(Y = 1) \\ &= e^{-(\lambda_1+\lambda_2)} \cdot \frac{(\lambda_1 + \lambda_2)^2}{2!} \end{aligned}$$

And the distribution of $X + Y$ is as in the table below which shows that $X + Y$ is a Poisson distribution with parameters $\lambda_1 + \lambda_2$.

$x + y$	0	1	2	$\cdots$
$P(X + Y = x + y)$	$e^{-(\lambda_1+\lambda_2)}$	$e^{-(\lambda_1+\lambda_2)} \cdot (\lambda_1 + \lambda_2)$	$e^{-(\lambda_1+\lambda_2)} \cdot \frac{(\lambda_1+\lambda_2)^2}{2!}$	$\cdots$

Example 19 *A company has 2 machines A and B. On average there are 0.8 breakdowns per week on machine A and 1.2 breakdowns on machine B. What is the probability of*

there being a total of 2 breakdowns on these two machines in a given week?
Soln: *Let X= Breakdowns on A per week and Y= Breakdowns on B per week. Then* $X \sim Poisson(0.8)$ *and* $Y \sim Poisson(1.2)$. *Let T= Total breakdowns per week.* $T \sim Poisson(0.8 + 1.2)$.

$$P(T = 2) = \frac{2^2}{2!}e^{-2} = 0.2707$$

Example 20 *The center page of the Monitor magazine consists of 1 page of local news and 1 page of adverts. The number of misprints in the local news page has a Poisson distribution with mean* 2.3 *and the number of misprints in the adverts pages has a Poisson distribution with mean* 1.7.

(a) *Find the probability that on the center page, there will be*
   (i) *no misprints*
   (ii) *more than 5 misprints*

(b) *Find the smallest integer n such that the probability that there are more than n misprints on the center page is less than* 0.1.

Solution *Let X be the number of misprints on the local page and Y the number of misprints on the adverts page. Then both X and Y follow a Poisson distribution. So, you have* $X \sim Poisson(2.3)$ *and* $Y \sim Poisson(1.7)$. *If T is the number of misprints on the center page, then* $T = X + Y$ *is a Poisson distribution with parameter* $\lambda = 2.3 + 1.7 = 4$. *Thus,*

(a) (i) *P(no misprints)=* $P(T = 0) = e^{-4} = 0.018$
   (ii) $P(T > 5) = 1 - P(T \leq 5) = 1 - 0.7851 = 0.215$*(ans 0.7851 is from tables)*

(b) *To find the smallest integer n you compute the following probabilities*

$$\begin{aligned} P(T > 5) &= 0.21 > 0.1 \\ P(T > 6) &= 0.111 > 0.1 \\ P(T > 7) &= 0.00051 < 0.1 \end{aligned}$$

Thus the smallest integer n such that the probability that there are more than n misprints on the center page is less than 0.1 *is* $n = 7$.

Example 21 *An insurance company offers policies covering houses, the contents of houses and cars. The number of claims received per day of these policies may be assumed to have Poisson distributions with means 2, 3 and 5 respectively. Find the probability that in a day the company will receive more than 12 claims in total.*

Example 22 *On average a certain intersection result in 3 traffic accidents per month. What is the probability that in any given month at this intersection (a) exactly 5 accidents will occur (b) less than 3 accidents will occur (c) at least 2 accidents will occur.* **Solution:** *Let X be the number of accidents that occur at this intersection. Then* $X \sim Poisson(3)$ $P(X = 5) = \frac{e^{-\lambda}3^5}{5!}$

Example 23 *A certain kind of sheet metal has on average five defects per 10 square feet. If we assume a Poisson distribution, what is the probability that a 15 square foot sheet of metal will have at least six defects?* **Solution:** *Let X denote the number of defects in a 15 square foot sheet of the metal. Then, since the unit of area is 10 square feet, t=15/10=1.5 this implies that* $\lambda' = \lambda t = 5 * 1.5 = 7.5$. *Thus* $P(X \geq 6) = 1 - P(X \leq 5) = 1 - 0.2414 = 0.7586$.

The Poisson Approximation to the Binomial Distribution

We shall now show that when the value of n is large and the value of p is close to 0, the binomial distribution with parameters n and p can be approximated by a Poisson distribution with mean np. This approximation is most adequate for practical purposes and quite accurate if $n \geq 20$ and $p \leq 0.05$ and it is very good if $n \geq 100$ and $np \leq 10$.

Suppose that a random variable X has a binomial distribution with parameters n and p, and let $P(X = x) = f(x, n, p)$, for any given value of x. Then for $x = 1, 2, \ldots, n$,

$$f(x,n,p) = \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1-p)^{n-x} = \frac{n(n-1)\cdots(n-x+1)}{x!} p^x (1-p)^{n-x}$$

If we let $\lambda = np$, then $f(x, n, p)$ can be rewritten in the following form:

$$f(x,n,p) = \frac{\lambda^x}{x!} \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-x+1}{n} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$

We shall now let $n \to \infty$ and $p \to 0$ in such a way that the value of the product np remains equal to the fixed value λ throughout this limiting process. Since the values of λ and x are held as fixed as $n \to \infty$, then

$$\lim_{n\to\infty} \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-x+1}{n} \left(1 - \frac{\lambda}{n}\right)^{-x} = 1$$

Furthermore, it is known from elementary calculus that

$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$$

It now follows that for any nonnegative integer x,

$$f(x,n,p) \to \frac{e^{-\lambda}\lambda^x}{x!}$$

The expression on the right is the p.m.f $f(x, \lambda)$ of the Poisson distribution with mean λ. Therefore when n is large and p is close to zero, the value of the p.m.f $f(x, n, p)$ of the binomial distribution can be approximated, $x = 0, 1, 2, \ldots$, by the value of the p.m.f $f(x, \lambda)$ of the Poisson distribution for which $\lambda = np$.

Example 24 *A factory packs bolts in boxes of 500. The probability that a bolt is defective is 0.002. Find the probability that a box contains 2 defective bolts.* **Solution** *Let X be the number of defective bolts in a box. Then X is binomially distributed with* $n = 500$ *and* $p = 0.002$. *This problem solved as a pure binomial distribution has the following solution.*

$$P(X = 2) = \binom{500}{2}(0.002)^2(0.988)^{498} = 0.184$$

By Poisson approximation, note that $n = 500$ *is large and* $p = 0.002$ *is small. Also* $\lambda = np = 500 * 0.002 = 1$. *This*

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \Rightarrow P(X = 2) = \frac{e^{-1}}{2!} = 0.184$$

Both answers agree to 3 decimal places, but may seem much easier in Poisson approximation.

Example 25 *A manufacturer of Christmas tree light bulbs knows that 2% of his are defective. Find the probability that a box of 100 bulbs contains at most 3 defective bulbs.* **Solution** *Since* $n = 100$ *(large) and* $p = 2\% = 0.02$ *(small) then you approximate with Poisson distribution, with* $\lambda = np = 100 * 0.02 = 2$. *Thus*

$$P(X \leq 3) = \sum_{x=0}^{3} p(x; 2) = 0.857$$

**from tables.*
Using the binomial distribution we have

$$P(X \leq 3) = \sum_{x=0}^{3} b(x; 100, 0.02)$$

compute!!

Example 26 *Suppose that on average 1 person in 1000 is HIV positive. Find the probability that a random sample of 8000 people will yield less than 7 HIV-positive cases.* **Solution** *This is a binomial distribution problem where* $n = 8000$ *and* $p = 0.001$. *Since n is quite large and p is close to zero, you approximate it with a Poisson,* $\lambda = np = 8000 * 0.001 = 8$ *and so*

$$P(X < 7) = \sum_{x=0}^{6} b(x; 8000, 0.001) = \sum_{x=0}^{6} p(x; 8) = 0.3134$$

from tables.

5.5 The Geometric Distribution

Imagine a situation where an "experiment" is repeated until a "success" occurs for the first time. If X counts the number of repetitions needed until the first success, then X has a geometric distribution, e.g. the number of tosses of a coin until a first tail is obtained or a process where components from a production line are tested, in turn until the first defective item is found. The geometric distribution is another important discrete distribution. It is important to note that if a success occurs at the nth trial then the first $n - 1$ trials are failures.

A random experiment is said to be a geometric experiment if:

1. the outcome of each trial can be classified into one of the two categories, success or failure,
2. each trial has the same probability p of success,
3. the trials are mutually independent,
4. the trials are repeated until one success is obtained.

If X is the number of trials required until a success is obtained, then X is said to be a geometric random variable. Consider performing a series of independent trials each with a constant probability p of success and q of failure, where $q = 1 - p$. Let X be a random variable "the number of trials up to and including the first success". Then:

- $P(X = 1) = P(\text{success on first trial}) = p$
- $P(X = 2) = P(\text{success on second trial}) = qp$
- $P(X = 3) = P(\text{success on third trial}) = q^2p$
- $P(X = 4) = P(\text{success on fourth trial}) = q^3p$

Thus if an experiment is performed x times and success occurs on the $x - th$ time, then the probability distribution of the random variable X, the number of the trials on which the first success occurs is given by:

$$P(X = x) = (1 - p)^{x-1}p = q^{x-1}p, \quad x = 1, 2, \ldots$$

Thus a random variable X is said to have a geometric distribution with parameter p, $0 < p < 1$, if its p.m.f is given by:

$$P(X = x) = g(x, p) = pq^{x-1}, \quad x = 1, 2, \ldots$$

where $q = 1 - p$ and we write $X \sim Geo(p)$.

It is always desirable to show that such a distribution is indeed a probability mass function. That is $g(x, p) \geq 0$ for all x and $\sum_{x=1}^{\infty} g(x, p) = 1$.

Now

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

The expression under summation is a sum of geometric series and is equal to $\frac{1}{1-q}$ Thus,

$$\sum_{x=1}^{\infty} pq^{x-1} = \frac{p}{1-q} = 1$$

Hence $g(x,p) = pq^{x-1}, \quad x = 1,2,3,\ldots$, is a probability mass function.

Note: If we define X to be the number of failures before we get a success, then

$$P(X = x) = (1-p)^x p, \quad x = 0,1,2,\ldots$$

If X is a geometrically distributed random variable with parameter p and probability distribution $P(X = x) = pq^{x-1}, \quad x = 1,2,3,\ldots$; then:

1. $P(X = r) = pq^{r-1}$
2. $P(X \leq r) = P(\text{success at some trial in the first } r \text{ trials})$
   $=1 - p(\text{no success in first } r \text{ trials})=1 - q^r$
3. $P(X > r) = 1 - P(X \leq r) = 1 - (1 - q^r) = q^r$
4. $P[(X > a + b)|(X > a)] = P(X > b)$ because
   $P[(X > a + b)|X > a] = \frac{q^{a+b}}{q^a} = q^b = P(X > b).$

e.g. $P[(X > 10)|X > 7)] = P(X > 3) = q^3$

The geometric distribution is related to the binomial distribution in that both are based on independent trials in which the probability of success is constant and equal to p. However, a geometric random variable is the number of trials until the first success (or up to a first success), whereas a Binomial random variable is the number of successes in n trials.

Mean and Variance of a Geometric Distribution

If X is a geometric random variable with probability of success p, then the mean or expectation of X is given by

$$E(X) = \frac{1}{p}$$

Proof:

$$
\begin{aligned}
E(X) &= \sum_{allx} xP(X = x) \\
&= \sum_{x=1}^{\infty} xpq^{x-1} \\
&= p\sum_{x=1}^{\infty} xq^{x-1} \\
&= p\sum_{x=1}^{\infty} \frac{d}{dq}(q^x) = p\frac{d}{dq}\sum_{x=1}^{\infty} q^x \\
&= p\frac{d}{dq}\left(\frac{q}{1-q}\right) = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}
\end{aligned}
$$

Thus, if independent trials, having a common probability p of being successful, are performed until the first success occurs, then the expected number of required trials is equal to $\frac{1}{p}$. For example, we could expect to roll a fair die 6 times in order to obtain the value 1.

Definition 3 *If X is a geometric random variable with probability of success p, then the variance of X is given by*

$$Var(X) = \frac{q}{p^2}$$

.

Proof: First derive an expression for $E[X(X-1)]$.

$$
\begin{aligned}
E[X(X-1)] &= \sum_{allx} x(x-1)P(X = x) \\
&= \sum_{x=1}^{\infty} x(x-1)pq^{x-1} = pq\sum_{x=2}^{\infty} x(x-1)q^{x-2} \\
&= pq\sum_{x=2}^{\infty} \frac{d^2}{dq^2}(q^x) = pq\frac{d^2}{dq^2}\sum_{x=2}^{\infty}(q^x) = pq\frac{d^2}{dq^2}\left(\frac{q^2}{1-q}\right) \\
&= pq\frac{d}{dq}\left(\frac{2q-q^2}{(1-q)^2}\right) = pq\left(\frac{2}{(1-q)^3}\right) \\
&= \frac{2pq}{(1-q)^3} = \frac{2pq}{p^3} = \frac{2q}{p^2}
\end{aligned}
$$

Thus,

$$E[X(X-1)] = E(X^2) - E(X) = \frac{2q}{p^2} \Rightarrow E(X^2) = \frac{2q}{p^2} + \frac{1}{p} = \frac{2}{p^2} - \frac{1}{p}$$

Using the definition of variance, you get

$$\begin{aligned} Var(X) =& E(X^2) - [E(X)]^2 \\ =& \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} \\ =& \frac{1}{p^2} - \frac{1}{p} = \frac{1-p}{p^2} \\ =& \frac{q}{p^2} \end{aligned}$$

Example 27 *Each time a player plays a certain card game called "patience", there is a probability of 0.15 that the game will work out successfully. The probability of success in any game is independent of the outcome of any other game. A player plays games of patience until a game works out successfully.*

a) Find the probability that she plays 4 games altogether.

b) Find the probability that at most 6 games are played.

c) Write down the expected number of games that are played.

Solution:
The number of games played, X, has a geometric distribution.
$X \sim Geo(0.15)$.

a) If she plays 4 games altogether she must lose the first 3 games and be successful on the 4^{th} game. So

$$P(X = 4) = 0.85^3 \times 0.15 = 0.0921$$

b) $P(\text{at most 6 games}) = P(X = 1) + P(X = 2) + \cdots + P(X = 6)$.
This would give the correct answer but involves quite a lot of working out.
Instead, $P(\text{at most 6 games}) = 1 - P(\text{more than 6 games})$.
But if more than 6 games are needed then the she must have been unsuccessful at the first 6 goes.
So $P(\text{at most 6 games}) = 1 - P(\text{more than 6 games}) = 1 - 0.85^6 = 0.623$

c) $E(X) = \frac{1}{p} = \frac{1}{0.15} = 6.67$

Example 28 *If 40% of people in a certain population have blood type A, find the probability that in a sequential testing of randomly selected people, the first person with type A blood is found on the fourth test.*
Solution:
Let X be the number of tests up to the first person of type A blood. Then $X \sim Ge0(0.4)$.
We wish to find $P(X = 4) = (0.6)^3(0.4)$

Example 29 *In a certain manufacturing process it is known that, on average, 1 in every 100 items is defective. What is the probability that 5 items are inspected before a defective item is found?*

Solution

Using the geometric distribution with p=1/100=0.01, and $P(X = 5) = (0.01)(0.99)^4 = 0.0096$.

Example 30 *A coin is biased so that the probability of obtaining a head is 0.6. If X is the random variable "X=the number of tosses up to and including the first head", find:*

(a) $P(X \leq 4)$,

(b) $P(X > 5)$,

(c) $P[(X > 8)|(X > 5)$.

Solution

Using the definition, $P(X = x) = g(x; p) = pq^{x-1}$, *x=1,2,... together with* $p = 0.6, q = 0.4$ *and* $P(X > a) = q^a$ *you get:*

(a) $P(X \leq 4) = 1 - q^4 = 1 - (0.4)^4 = 0.9744$

(b) $P(X > 5) = q^5 = (0.4)^5 = 0.01024$

(c) $P[(X > 8)|(X > 5) = P(X > 3) = q^3 = (0.4)^3 = 0.064$

Example 31 *In a particular game, a player can get out of check only if she gets two heads when she tosses two coins.*

(a) Find the probability that more than 6 attempts are needed to get out of check.

(b) What is the smallest value of n if there is to be at least a 90% chance of getting out of check on or before the $n-th$ *attempt.*

Solution:

P(2heads when two coins are tossed)=$\frac{1}{4}$. *So* $P(success) = \frac{1}{4}$ *and* $q = \frac{3}{4}$. *If X= the number of attempts required to get out check, then*

(a) $P(X > 6) = q^6 = (\frac{3}{4})^6 = 0.17$

(b) $P(X > n) = (\frac{3}{4})^n$. *So* $P(X \leq n) = 1 - (\frac{3}{4})^n$. *To be at least 90% sure that one is out of check,* $\Rightarrow 1 - (\frac{3}{4})^n \geq 0.9 \Rightarrow (\frac{3}{4})^n \leq 0.1$. *If you take logs to base 10 on both sides you get*

$$n \log_{10} 0.75 \geq \log_{10} 0.1 \Rightarrow n \geq \frac{\log_{10} 0.1}{\log_{10} 0.75} = 8.0039$$

Therefore the smallest n is 9.

Example 32 *The probability that a marksman hits a target eye is 0.4 for each shot,and each shot is independent of the other. Find the:*

(a) probability that he hits the target for the first time on his fourth attempt,

(b) mean number of shots needed to hit the target and standard deviation,

(c) most likely number of shots until he hits the target.

Solution: *Let X be the number of attempts up to and including the first target. Then* $P(X = x) = pq^{x-1}, \ \ x = 1, 2, 3, ...;$ *with* $p = 0.4$ *and* $q = 0.6$.

(a) $P(X = 4) = q^3 p = (0.6)^3(0.4) = 0.0864$

(b) $E(X) = \frac{1}{p} = \frac{1}{0.4} = 2.5$ *and* $Var(X) = \frac{q}{p^2} = \frac{0.6}{(0.4)^2} = 3.75$. *Therefore the standard deviation* $= \sqrt{3.75} = 1.94$.

(c) $P(X = 1) = (0.4)(0.6)^0 = 0.4, P(X = 2) = (0.4)(0.6) = 0.24$ *and* $P(X = 3) = (0.4)(0.6)^2 = 0.144$.

Therefore, the most likely number of shots is 1, since the probabilities are decreasing and it is the one with the highest probability. Generally it is noted that $p > qp > q^2 p > q^3 p > ...$ *so the mode of the geometric distribution (or the most likely value) occurs at* $X = 1$ *with probability* $P(X = 1) = p$.

5.6 The Negative Binomial

This is a generalization of the geometric distribution. It gives the probability distribution of the k^{th} success at the x^{th} trial in a series of independent Bernoulli trials. Let X denote the trial at which the k^{th} success occurs, where k is a fixed integer. Then

$$P(X = x|k, p) = \begin{pmatrix} x-1 \\ k-1 \end{pmatrix} p^k q^{x-k}, \quad x = k, k+1, \ldots$$

and we say that X has a negative binomial(k, p) distribution. In most of the literature, it is denoted as $b^*(x; k, p)$.

The negative binomial distribution is sometimes defined in terms of the random variable Y =number of failures before kth success. This formulation is statistically equivalent to the one given above in terms of X =trial at which the kth success occurs, since Y = X -k. The alternative form of the negative binomial distribution is:

$$P(Y = y) = \begin{pmatrix} k+y-1 \\ y \end{pmatrix} p^k q^y, \quad y = 0, 1, \ldots$$

Example: Find the probability that a person tossing 3 coins will get either all heads or all tails for the second time on the fifth toss.
Solution
Using the Negative Binomial with $x = 5, k = 2$ and $p = \frac{1}{4}$, we have:

$b^*(5; 2, 0.25) = \begin{pmatrix} 5-1 \\ 2-1 \end{pmatrix} 0.25^2 0.75^3 = 0.1445$

Example: An oil company conducts a geological study that indicates that an exploratory oil well should have a 20% chance of striking oil. What is the probability that the first strike comes on the third well drilled?
Solution:

This a Negative Binomial problem with $k = 1$, $p = 0.2$, $q = 1 - p = 0.8$, $x = 3$. This gives:
$b^*(3; 1, 0.2) = \begin{pmatrix} 3-1 \\ 1-1 \end{pmatrix} 0.2^1 0.8^2 = 0.128$
Note: This problem can as well be solved using a Geometric distribution. It should be observed that when $k = 1$, the Negative Binomial distribution reduces to a Geometric distribution.

5.7 The Hypergeometric Distribution

Suppose we are sampling without replacement from a batch of items containing a variable number of defectives. We are essentially assuming that we know the probability p that a given item is defective but not the actual number of defective items contained in the batch. The number of defective items in the batch is a random variable in this case. When we sample from the batch, we are left with:

1. a smaller batch:
2. a (possibly) smaller (but still variable) number of defective items. The number of defective items is still a random variable.

While the probability of finding a given number of defectives in a sample drawn from the second batch will, (in general) be different from the probability of finding a given number of defectives in a sample drawn from the first batch, sampling from both batches may be described by the Binomial distribution for which:

$$P(X = r) = ^n C_r p^r (1-p)^{n-r}$$

Sampling in this case varies the values of n and p in general but not the underlying distribution describing the sampling process.

Example 33 *A batch of 100 piston rings is known to contain 10 defective rings. If two piston rings are drawn from the batch, write down the probabilities that:*

1. *the first ring is defective;*
2. *the second ring is defective given that the first one is defective.*

Soln:

1. *The probability that the first ring is defective is clearly* $\dfrac{10}{100} = \dfrac{1}{10}$.
2. *Assuming that the first ring selected is defective and we do not replace it, the probability that the second ring is defective is equally clearly* $\dfrac{9}{99} = \frac{1}{11}$

The **Hypergeometric distribution** may be thought of as arising from sampling from a batch of items where the number of defective items contained in the batch is known.
Essentially the number of defectives contained in the batch is not a random variable, it is fixed.
The calculations involved when using the Hypergeometric distribution are usually more complex than their Binomial counterparts. If we sample without replacement we may proceed in general as follows:

- we may select n items from a population of N items in ${}^{N}C_n$ ways;
- we may select r defective items from M defective items in ${}^{M}C_r$ ways;
- we may select $n - r$ non-defective items from $N - M$ non-defective items in ${}^{N-M}C_{n-r}$ ways;
- hence we may select n items containing r defectives in ${}^{M}C_r \times {}^{N-M} C_{n-r}$ ways.
- hence the probability that we select a sample of size n containing r defective items from a population of N items known to contain M defective items is

$$\frac{{}^{M}C_r \times {}^{N-M} C_{n-r}}{{}^{N}C_n}$$

Definition 1 *The distribution given by*

$$P(X = r) = \frac{{}^{M}C_r \times {}^{N-M} C_{n-r}}{{}^{N}C_n}, \quad r = 0, 1, \ldots, n$$

*which describes the probability of obtaining a sample of size n containing r defective items from a population of size N known to contain M defective items is known as the **Hypergeometric distribution**.*

Example 34 *A batch of 10 rocker cover gaskets contains 4 defective gaskets. If we draw samples of size 3 without replacement, from the batch of 10, find the probability that a sample contains 2 defective gaskets.*
Soln: *Using* $P(X = r) = \dfrac{{}^{M}C_r \times {}^{N-M} C_{n-r}}{{}^{N}C_n}$, *we know that* $N = 10, M = 4, n = 3$ *and* $r = 2$.
Hence $P(X = 2) = \dfrac{{}^{4}C_2 \times {}^{6} C_1}{{}^{10}C_3} = \dfrac{6 \times 6}{120} = 0.3$

Example 35 *A committee of size 5 is to be selected at random from three men and five women. Find the probability distribution for the number of:*

(a) men on the committee,

(b) women on the committee.

Soln:
(i) Let the variable X be the number of men on the committee. X has a hypergeometric distribution in which $N = 8, n = 5, M = 3$ *and* $r = 0, 1, 2, 3$. *Thus,*

$$P(X = 0) = \frac{{}^3C_0 \times^5 C_5}{{}^8C_5} = \frac{1}{56}$$
$$P(X = 1) = \frac{{}^3C_1 \times^5 C_4}{{}^8C_5} = \frac{15}{56}$$
$$P(X = 2) = \frac{{}^3C_2 \times^5 C_3}{{}^8C_5} = \frac{30}{56}$$
$$P(X = 3) = \frac{{}^3C_3 \times^5 C_2}{{}^8C_5} = \frac{10}{56}$$

In tabular form we have a hypergeometric distribution of X as

x	*0*	*1*	*2*	*3*
P(X=x)	$\frac{1}{56}$	$\frac{15}{56}$	$\frac{30}{56}$	$\frac{10}{56}$

(ii) If we let Y to be the number of women on the committee, the distribution can easily be seen from (i) as:

y	2	3	4	5
P(Y=y)	$\frac{10}{56}$	$\frac{30}{56}$	$\frac{15}{56}$	$\frac{1}{56}$

Note*: The committee cannot have one or no woman as this will require to have more than three men given that the committee must be composed of five individuals.*

Example 36 *Boxes of 40 light bulbs are called acceptable if they contain no more than 3 defectives. The procedure of sampling the box is to select 5 bulbs at random and reject a box if a defective bulb is found. What is the probability that exactly one defective bulb will be found in a sample if there are three defective bulbs in the entire box?*

Soln: *Using* $P(X = r) = \frac{{}^MC_r \times^{N-M} C_{n-r}}{{}^NC_n}$, *we know that* $N = 40, M = 3, n = 5$ *and* $r = 1$.
Hence $P(X = 1) = \frac{{}^3C_1 \times^{37} C_4}{{}^{40}C_5} = 0.3011$

It is possible to derive formulae for the mean and variance of the Hypergeometric distribution. However, the calculations are more difficult than their Binomial counterparts.

5.7.1 The Mean and Variance of a Hypergeometric Distribution

Let X be a hyper geometric random variable with p.m.f

$$P(X=x)=\frac{{}^{M}C_x \times^{N-M} C_{n-x}}{{}^{N}C_n}, \quad x=0,1,\ldots,n$$

The expectation (mean) and variance of of X are given by

$$E(X)=np \quad \text{and} \quad Var(X)=np(1-p)\frac{N-n}{N-1} \quad \text{where } p=\frac{M}{N}$$

Proof: From the definition of expectation of a discrete random variable

$$\begin{aligned}
E(X)&=\sum_{allx} xP(X=x)=\sum_{x=0}^{n} x\frac{{}^{M}C_x \times^{N-M} C_{n-x}}{{}^{N}C_n}\\
&=\sum_{x=1}^{n} x\frac{M!}{x!(M-x)!}\frac{{}^{N-M}C_{n-x}}{{}^{N}C_n}\\
&=M\sum_{x=1}^{n} \frac{(M-1)!}{(x-1)!(M-x)!}\frac{{}^{N-M}C_{n-x}}{{}^{N}C_n}\\
&=M\sum_{x=1}^{n} \frac{{}^{M-1}C_{x-1} \times^{N-M} C_{n-x}}{{}^{N}C_n}
\end{aligned}$$

Letting $y=x-1$, we have

$$E(X)=M\sum_{y=0}^{n-1} \frac{{}^{M-1}C_y \times^{N-M} C_{n-1-y}}{{}^{N}C_n}$$

Writing

$${}^{N-M}C_{n-1-y} =^{(N-1)-(M-1)} C_{n-1-y} \quad \text{and } {}^{N}C_n = \frac{N}{n}\times^{N-1} C_{n-1}$$

we have

$$E(X)=\frac{nM}{N}\sum_{y=0}^{n-1} \frac{{}^{M-1}C_y \times^{(N-1)-(M-1)} C_{n-1-y}}{{}^{N-1}C_{n-1}}$$

The expression inside the summation represents the total of all probabilities of a hypergeometric experiment when a sample of $n-1$ items is selected at random from $N-1$ items, of which $M-1$ are defective, and hence equals to 1. Therefore, the expectation is

$$E(X)=\frac{nM}{N}$$

To compute the variance, we first derive the expression for $E(X(X-1))$. Thus from the definition of the expectation, we have

$$
\begin{aligned}
E[X(X-1)] = \sum_{allx} x(x-1)P(X=x) &= \sum_{x=0}^{n} x(x-1)\frac{{}^{M}C_x \times {}^{N-M} C_{n-x}}{{}^{N}C_n} \\
&= M(M-1)\sum_{x=2}^{n} \frac{(M-2)!}{(x-2)!(M-x)!}\frac{{}^{N-M}C_{n-x}}{{}^{N}C_n} \\
&= \frac{n(n-1)M(M-1)}{N(N-1)}\sum_{y=0}^{n-2} \frac{{}^{M-2}C_y \times {}^{(N-2)-(M-2)} C_{n-2-y}}{{}^{N-2}C_{n-2}}
\end{aligned}
$$

The expression inside the summation is the sum of $(n-2)$ terms of the hypergeometric distribution and is equal to 1. Hence

$$
E[X(X-1)] = E(X^2) - E(X) = n(n-1)\frac{M(M-1)}{N(N-1)}
$$

Thus

$$
E(X^2) = n(n-1)\frac{M(M-1)}{N(N-1)} + \frac{nM}{N}
$$

Using the definition of variance we obtain

$$
\begin{aligned}
Var(X) &= E(X^2) - [E(X)]^2 \\
&= n(n-1)\frac{M(M-1)}{N(N-1)} + \frac{nM}{N} - \frac{n^2M^2}{N^2} \\
&= \frac{nM}{N}\left[(n-1)\frac{(M-1)}{(N-1)} + 1 - \frac{nM}{N}\right] \\
&= \frac{nM}{N}\left[\frac{(N-M)(N-n)}{N(N-1)}\right] \\
&= \frac{nM}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}
\end{aligned}
$$

If we set $p = \frac{M}{N}$, then the mean of the hypergeometric distribution coincides with the mean of the binomial distribution, and the variance of the hyper geometric distribution is $\frac{N-n}{N-1}$ times the variance of the binomial distribution. Therefore, we can approximate the hypergeometric distribution by using binomial distribution with mean $\mu = np = \frac{nM}{N}$ and variance $\sigma^2 = npq = \frac{nM}{N}(1-\frac{M}{N})$. This is so when n is small relative to N.

We note

- the mean of a hypergeometric distribution is the same as that of a binomial.

- The variance of a hypergeometric distribution is equal to the variance of a binomial distribution multiplied by a factor $f = \dfrac{N-n}{N-1}$. Since $f < 1$, the variance of a hypergeometric distribution is smaller than that of a binomial distribution. The factor f is called the finite population correction.

5.7.2 Applications

The hypergeometric distribution is used for calculating probabilities for samples drawn from relatively small populations and without replication, which means that an item's choice of being selected increases on each trial.
For sampling with replication, in which case an item has an equal chance of being selected, the binomial distribution should be used. If $N >> n$, in which case the population size of interest is too large compared to the sample size, $f \approx 1$ (the finite population correct can be neglected). The probabilities from a hypergeometric distribution and those of a binomial distributions are no different.

Example 37 *In the manufacture of car tyres, a particular production process is known to yield 10 tyres with defective walls in every batch of 100 tyres produced. From a production batch of 100 tyres, a sample of 4 is selected for testing to destruction. Find:*

1. *the probability that the sample contains 1 defective tyre*
2. *the expectation of the number of defectives in samples of size 4*
3. *the variance of the number of defectives in samples of size 4.*

Soln:*Sampling is clearly without replacement and we use the Hypergeometric distribution with* $N = 100, M = 10, n = 4, x = 1$ *and* $p = 0.1$*. Hence:*

1. $P(X = x) = \dfrac{{}^{M}C_x \times^{N-M} C_{n-x}}{{}^{N}C_n}$ *gives*
   $P(X = 1) = \dfrac{{}^{10}C_1 \times^{100-10} C_{4-1}}{{}^{100}C_4} = \dfrac{10 \times 117480}{3921225} \approx 0.3$
2. $E(X) = np = 4 \times 0.1 = 0.4$
3. $Var(X) = np(1-p)\dfrac{N-n}{N-1} = 0.4 \times 0.9 \times \dfrac{96}{99} = 0.349$

Table A.1 (continued) Binomial Probability Sums $\sum_{x=0}^{r} b(x; n, p)$

		p									
n	***r***	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**	**0.90**
15	**0**	0.2059	0.0352	0.0134	0.0047	0.0005	0.0000				
	1	0.5490	0.1671	0.0802	0.0353	0.0052	0.0005	0.0000			
	2	0.8159	0.3980	0.2361	0.1268	0.0271	0.0037	0.0003	0.0000		
	3	0.9444	0.6482	0.4613	0.2969	0.0905	0.0176	0.0019	0.0001		
	4	0.9873	0.8358	0.6865	0.5155	0.2173	0.0592	0.0093	0.0007	0.0000	
	5	0.9978	0.9389	0.8516	0.7216	0.4032	0.1509	0.0338	0.0037	0.0001	
	6	0.9997	0.9819	0.9434	0.8689	0.6098	0.3036	0.0950	0.0152	0.0008	
	7	1.0000	0.9958	0.9827	0.9500	0.7869	0.5000	0.2131	0.0500	0.0042	0.0000
	8		0.9992	0.9958	0.9848	0.9050	0.6964	0.3902	0.1311	0.0181	0.0003
	9		0.9999	0.9992	0.9963	0.9662	0.8491	0.5968	0.2784	0.0611	0.0022
	10		1.0000	0.9999	0.9993	0.9907	0.9408	0.7827	0.4845	0.1642	0.0127
	11			1.0000	0.9999	0.9981	0.9824	0.9095	0.7031	0.3518	0.0556
	12				1.0000	0.9997	0.9963	0.9729	0.8732	0.6020	0.1841
	13					1.0000	0.9995	0.9948	0.9647	0.8329	0.4510
	14						1.0000	0.9995	0.9953	0.9648	0.7941
	15							1.0000	1.0000	1.0000	1.0000
16	**0**	0.1853	0.0281	0.0100	0.0033	0.0003	0.0000				
	1	0.5147	0.1407	0.0635	0.0261	0.0033	0.0003	0.0000			
	2	0.7892	0.3518	0.1971	0.0994	0.0183	0.0021	0.0001			
	3	0.9316	0.5981	0.4050	0.2459	0.0651	0.0106	0.0009	0.0000		
	4	0.9830	0.7982	0.6302	0.4499	0.1666	0.0384	0.0049	0.0003		
	5	0.9967	0.9183	0.8103	0.6598	0.3288	0.1051	0.0191	0.0016	0.0000	
	6	0.9995	0.9733	0.9204	0.8247	0.5272	0.2272	0.0583	0.0071	0.0002	
	7	0.9999	0.9930	0.9729	0.9256	0.7161	0.4018	0.1423	0.0257	0.0015	0.0000
	8	1.0000	0.9985	0.9925	0.9743	0.8577	0.5982	0.2839	0.0744	0.0070	0.0001
	9		0.9998	0.9984	0.9929	0.9417	0.7728	0.4728	0.1753	0.0267	0.0005
	10		1.0000	0.9997	0.9984	0.9809	0.8949	0.6712	0.3402	0.0817	0.0033
	11			1.0000	0.9997	0.9951	0.9616	0.8334	0.5501	0.2018	0.0170
	12				1.0000	0.9991	0.9894	0.9349	0.7541	0.4019	0.0684
	13					0.9999	0.9979	0.9817	0.9006	0.6482	0.2108
	14					1.0000	0.9997	0.9967	0.9739	0.8593	0.4853
	15						1.0000	0.9997	0.9967	0.9719	0.8147
	16							1.0000	1.0000	1.0000	1.0000

Table A.2 (continued) Poisson Probability Sums $\sum_{x=0}^{r} p(x;\mu)$

	μ								
r	**10.0**	**11.0**	**12.0**	**13.0**	**14.0**	**15.0**	**16.0**	**17.0**	**18.0**
0	0.0000	0.0000	0.0000						
1	0.0005	0.0002	0.0001	0.0000	0.0000				
2	0.0028	0.0012	0.0005	0.0002	0.0001	0.0000	0.0000		
3	0.0103	0.0049	0.0023	0.0011	0.0005	0.0002	0.0001	0.0000	0.0000
4	0.0293	0.0151	0.0076	0.0037	0.0018	0.0009	0.0004	0.0002	0.0001
5	0.0671	0.0375	0.0203	0.0107	0.0055	0.0028	0.0014	0.0007	0.0003
6	0.1301	0.0786	0.0458	0.0259	0.0142	0.0076	0.0040	0.0021	0.0010
7	0.2202	0.1432	0.0895	0.0540	0.0316	0.0180	0.0100	0.0054	0.0029
8	0.3328	0.2320	0.1550	0.0998	0.0621	0.0374	0.0220	0.0126	0.0071
9	0.4579	0.3405	0.2424	0.1658	0.1094	0.0699	0.0433	0.0261	0.0154
10	0.5830	0.4599	0.3472	0.2517	0.1757	0.1185	0.0774	0.0491	0.0304
11	0.6968	0.5793	0.4616	0.3532	0.2600	0.1848	0.1270	0.0847	0.0549
12	0.7916	0.6887	0.5760	0.4631	0.3585	0.2676	0.1931	0.1350	0.0917
13	0.8645	0.7813	0.6815	0.5730	0.4644	0.3632	0.2745	0.2009	0.1426
14	0.9165	0.8540	0.7720	0.6751	0.5704	0.4657	0.3675	0.2808	0.2081
15	0.9513	0.9074	0.8444	0.7636	0.6694	0.5681	0.4667	0.3715	0.2867
16	0.9730	0.9441	0.8987	0.8355	0.7559	0.6641	0.5660	0.4677	0.3751
17	0.9857	0.9678	0.9370	0.8905	0.8272	0.7489	0.6593	0.5640	0.4686
18	0.9928	0.9823	0.9626	0.9302	0.8826	0.8195	0.7423	0.6550	0.5622
19	0.9965	0.9907	0.9787	0.9573	0.9235	0.8752	0.8122	0.7363	0.6509
20	0.9984	0.9953	0.9884	0.9750	0.9521	0.9170	0.8682	0.8055	0.7307
21	0.9993	0.9977	0.9939	0.9859	0.9712	0.9469	0.9108	0.8615	0.7991
22	0.9997	0.9990	0.9970	0.9924	0.9833	0.9673	0.9418	0.9047	0.8551
23	0.9999	0.9995	0.9985	0.9960	0.9907	0.9805	0.9633	0.9367	0.8989
24	1.0000	0.9998	0.9993	0.9980	0.9950	0.9888	0.9777	0.9594	0.9317
25		0.9999	0.9997	0.9990	0.9974	0.9938	0.9869	0.9748	0.9554
26		1.0000	0.9999	0.9995	0.9987	0.9967	0.9925	0.9848	0.9718
27			0.9999	0.9998	0.9994	0.9983	0.9959	0.9912	0.9827
28			1.0000	0.9999	0.9997	0.9991	0.9978	0.9950	0.9897
29				1.0000	0.9999	0.9996	0.9989	0.9973	0.9941
30					0.9999	0.9998	0.9994	0.9986	0.9967
31					1.0000	0.9999	0.9997	0.9993	0.9982
32						1.0000	0.9999	0.9996	0.9990
33							0.9999	0.9998	0.9995
34							1.0000	0.9999	0.9998
35								1.0000	0.9999
36									0.9999
37									1.0000