3 Discrete Random Variables and Probability Distributions

CHAPTER OUTLINE

LEARNING OBJECTIVES

After careful study of this chapter you should be able to do the following:

1. Determine probabilities from probability mass functions and the reverse
2. Determine probabilities from cumulative distribution functions and cumulative distribution functions from probability mass functions, and the reverse
3. Calculate means and variances for discrete random variables
4. Understand the assumptions for each of the discrete probability distributions presented
5. Select an appropriate discrete probability distribution to calculate probabilities in specific applications
6. Calculate probabilities, determine means and variances for each of the discrete probability distributions presented

3-1 Discrete Random Variables

Many physical systems can be modeled by the same or similar random experiments and random variables. The distribution of the random variables involved in each of these common systems can be analyzed, and the results of that analysis can be used in different applications and examples. In this chapter, we present the analysis of several random experiments and **discrete random variables** that frequently arise in applications. We often omit a discussion of the underlying sample space of the random experiment and directly describe the distribution of a particular random variable.

3-1 Discrete Random Variables

Example 3-1

A voice communication system for a business contains 48 external lines. At a particular time, the system is observed, and some of the lines are being used. Let the random variable X denote the number of lines in use. Then, X can assume any of the integer values 0 through 48. When the system is observed, if 10 lines are in use, $x = 10$.

3-2 Probability Distributions and Probability Mass Functions

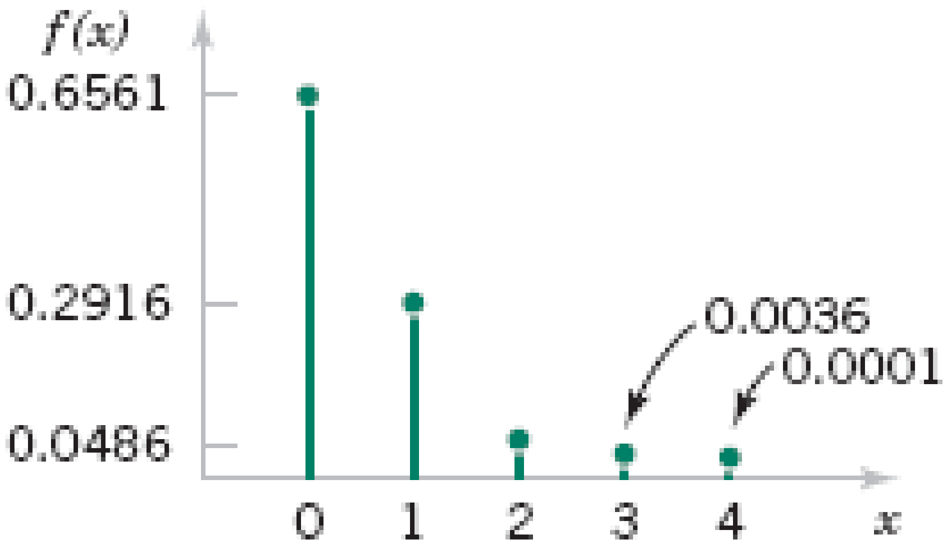


Figure 3-1 Probability distribution for bits in error.

3-2 Probability Distributions and Probability Mass Functions

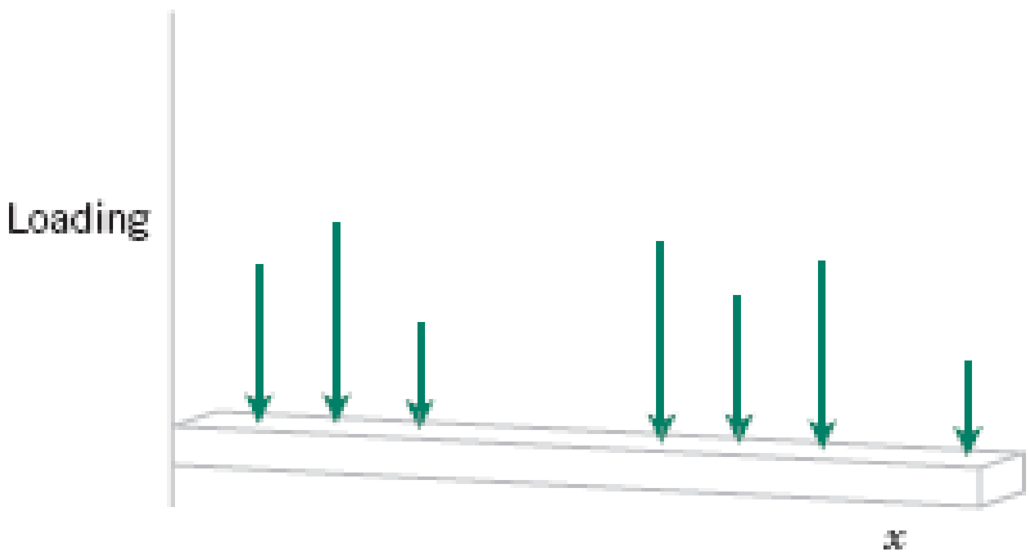


Figure 3-2 Loadings at discrete points on a long, thin beam.

3-2 Probability Distributions and Probability Mass Functions

Definition

For a discrete random variable X with possible values $x_1, x_2, \ldots, x_n$, a **probability mass function** is a function such that

(1) $f(x_i) \geq 0$

(2) $\sum_{i=1}^{n} f(x_i) = 1$

(3) $f(x_i) = P(X = x_i)$ (3-1)

Example 3-5

Let the random variable X denote the number of semiconductor wafers that need to be analyzed in order to detect a large particle of contamination. Assume that the probability that a wafer contains a large particle is 0.01 and that the wafers are independent. Determine the probability distribution of X.

Let p denote a wafer in which a large particle is present, and let a denote a wafer in which it is absent. The sample space of the experiment is infinite, and it can be represented as all possible sequences that start with a string of a's and end with p. That is,

$$s = \{p, ap, aap, aaap, aaaap, aaaaap, \text{and so forth}\}$$

Consider a few special cases. We have $P(X = 1) = P(p) = 0.01$. Also, using the independence assumption

$$P(X = 2) = P(ap) = 0.99(0.01) = 0.0099$$

Example 3-5 (continued)

A general formula is

$$P(X = x) = \underbrace{P(aa \ldots ap)}_{(x-1)a\text{'s}} = 0.99^{x-1}(0.01), \qquad \text{for } x = 1, 2, 3, \ldots$$

Describing the probabilities associated with X in terms of this formula is the simplest method of describing the distribution of X in this example. Clearly $f(x) \geq 0$. The fact that the sum of the probabilities is 1 is left as an exercise. This is an example of a geometric random variable, and details are provided later in this chapter.

3-3 Cumulative Distribution Functions

Definition

The **cumulative distribution function** of a discrete random variable X, denoted as $F(x)$, is

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

For a discrete random variable X, $F(x)$ satisfies the following properties.

(1) $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$

(2) $0 \leq F(x) \leq 1$

(3) If $x \leq y$, then $F(x) \leq F(y)$ (3-2)

Example 3-8

Suppose that a day's production of 850 manufactured parts contains 50 parts that do not conform to customer requirements. Two parts are selected at random, without replacement, from the batch. Let the random variable X equal the number of nonconforming parts in the sample. What is the cumulative distribution function of X?

The question can be answered by first finding the probability mass function of X.

$$P(X = 0) = \frac{800}{850} \cdot \frac{799}{849} = 0.886$$

$$P(X = 1) = 2 \cdot \frac{800}{850} \cdot \frac{50}{849} = 0.111$$

$$P(X = 2) = \frac{50}{850} \cdot \frac{49}{849} = 0.003$$

Therefore,

$$F(0) = P(X \leq 0) = 0.886$$

$$F(1) = P(X \leq 1) = 0.886 + 0.111 = 0.997$$

$$F(2) = P(X \leq 2) = 1$$

The cumulative distribution function for this example is graphed in Fig. 3-4. Note that $F(x)$ is defined for all x from $-\infty < x < \infty$ and not only for 0, 1, and 2.

Example 3-8

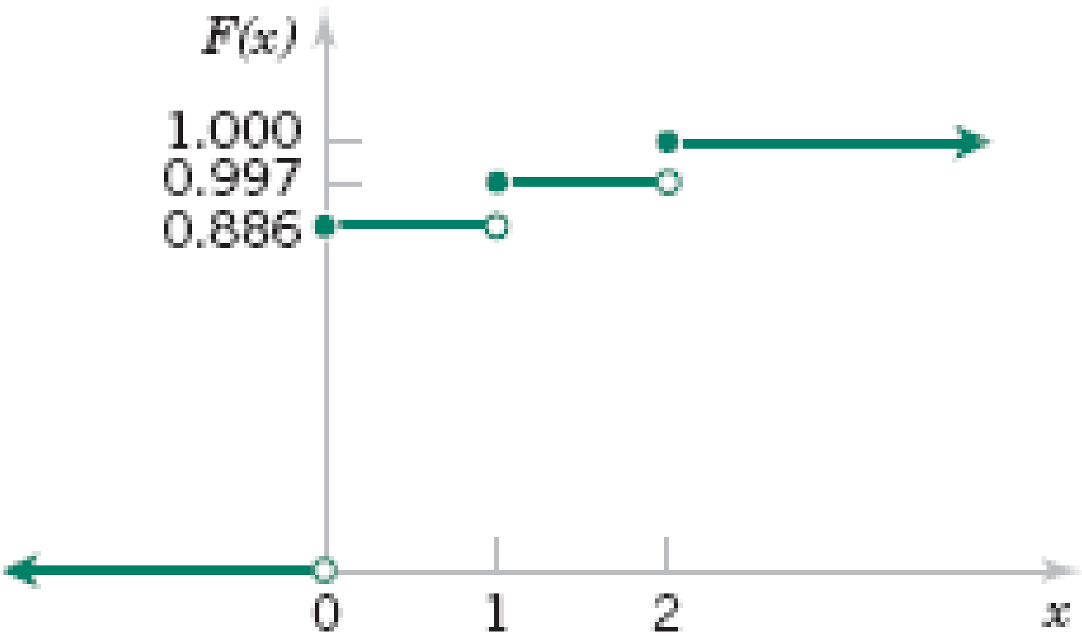


Figure 3-4 Cumulative distribution function for Example 3-8.

3-4 Mean and Variance of a Discrete Random Variable

Definition

The **mean** or **expected value** of the discrete random variable X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \sum_x x f(x) \tag{3-3}$$

The **variance** of X, denoted as σ^2 or $V(X)$, is

$$\sigma^2 = V(X) = E(X-\mu)^2 = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2$$

The **standard deviation** of X is $\sigma = \sqrt{\sigma^2}$.

3-4 Mean and Variance of a Discrete Random Variable

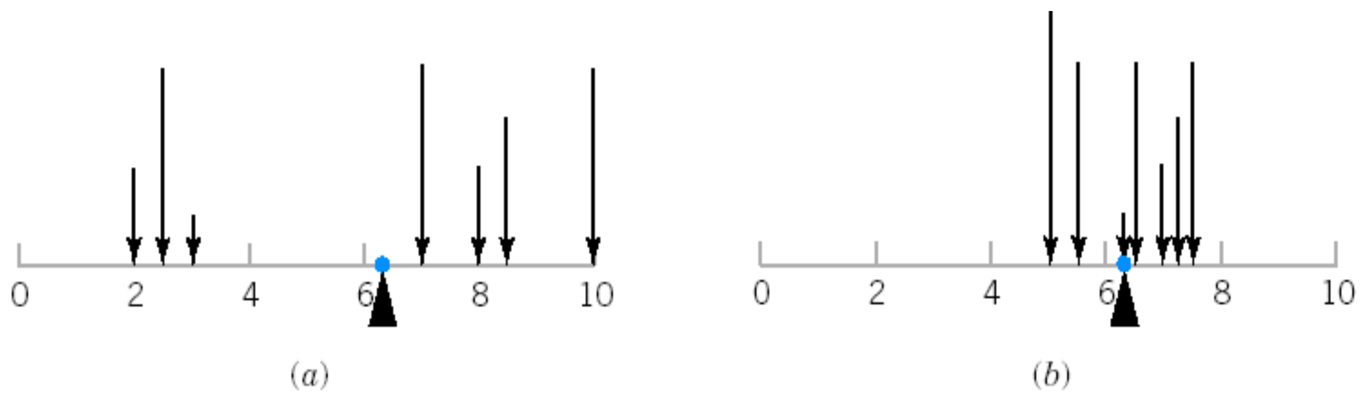


Figure 3-5 A probability distribution can be viewed as a loading with the mean equal to the balance point. Parts (a) and (b) illustrate equal means, but Part (a) illustrates a larger variance.

3-4 Mean and Variance of a Discrete Random Variable

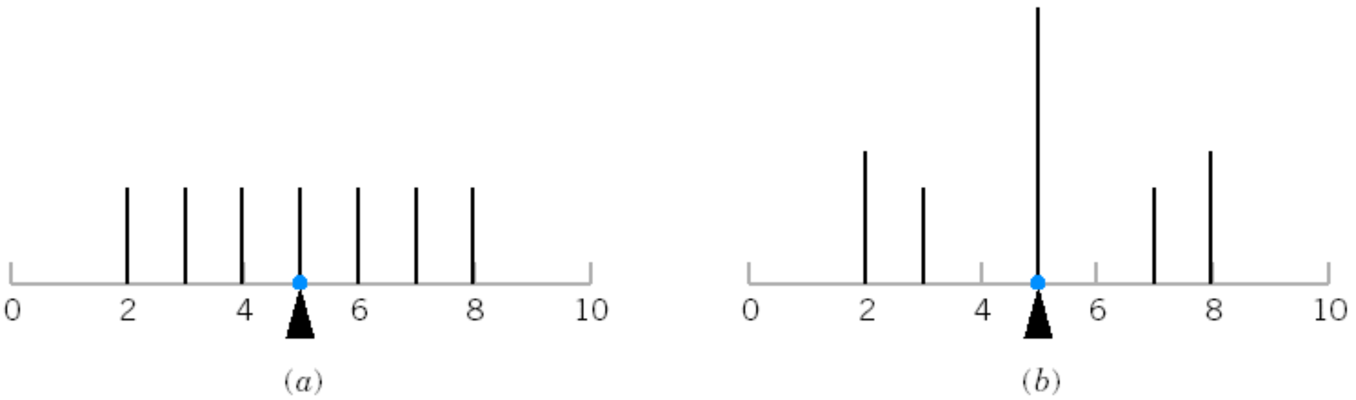


Figure 3-6 The probability distribution illustrated in Parts (a) and (b) differ even though they have equal means and equal variances.

Example 3-11

The number of messages sent per hour over a computer network has the following distribution:

x = number of messages	10	11	12	13	14	15
$f(x)$	0.08	0.15	0.30	0.20	0.20	0.07

Determine the mean and standard deviation of the number of messages sent per hour.

$$E(X) = 10(0.08) + 11(0.15) + \cdots + 15(0.07) = 12.5$$

$$V(X) = 10^2(0.08) + 11^2(0.15) + \cdots + 15^2(0.07) - 12.5^2 = 1.85$$

$$\sigma = \sqrt{V(X)} = \sqrt{1.85} = 1.36$$

3-4 Mean and Variance of a Discrete Random Variable

Expected Value of a Function of a Discrete Random Variable

If X is a discrete random variable with probability mass function $f(x)$,

$$E[h(X)] = \sum_{x} h(x)f(x) \tag{3-4}$$

3-5 Discrete Uniform Distribution

Definition

A random variable X has a **discrete uniform distribution** if each of the n values in its range, say, $x_1, x_2, \ldots, x_n$, has equal probability. Then,

$$f(x_i) = 1/n \tag{3-5}$$

3-5 Discrete Uniform Distribution

Example 3-13

The first digit of a part's serial number is equally likely to be any one of the digits 0 through 9. If one part is selected from a large batch and X is the first digit of the serial number, X has a discrete uniform distribution with probability 0.1 for each value in $R = \{0, 1, 2, \ldots, 9\}$. That is,

$$f(x) = 0.1$$

for each value in R. The probability mass function of X is shown in Fig. 3-7.

3-5 Discrete Uniform Distribution

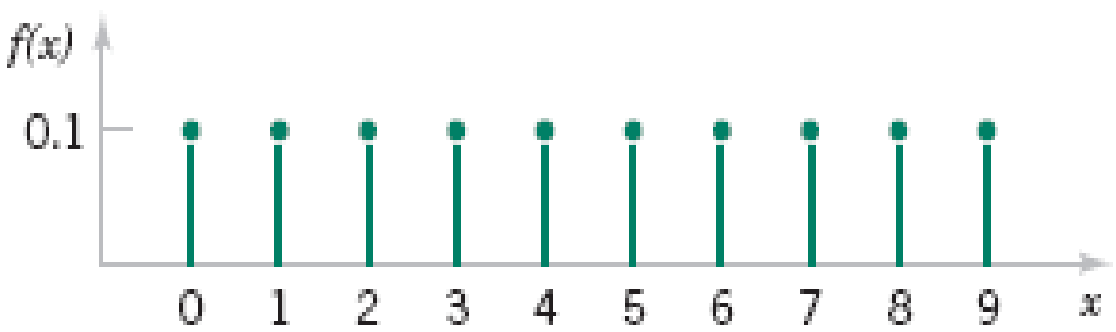


Figure 3-7 Probability mass function for a discrete uniform random variable.

3-5 Discrete Uniform Distribution

Mean and Variance

Suppose X is a discrete uniform random variable on the consecutive integers $a, a+1, a+2, \ldots, b$, for $a \le b$. The mean of X is

$$\mu = E(X) = \frac{b+a}{2}$$

The variance of X is

$$\sigma^2 = \frac{(b-a+1)^2 - 1}{12} \tag{3-6}$$

3-6 Binomial Distribution

Random experiments and random variables

1. Flip a coin 10 times. Let X = number of heads obtained.
2. A worn machine tool produces 1% defective parts. Let X = number of defective parts in the next 25 parts produced.
3. Each sample of air has a 10% chance of containing a particular rare molecule. Let X = the number of air samples that contain the rare molecule in the next 18 samples analyzed.
4. Of all bits transmitted through a digital transmission channel, 10% are received in error. Let X = the number of bits in error in the next five bits transmitted.

3-6 Binomial Distribution

Random experiments and random variables

5. A multiple choice test contains 10 questions, each with four choices, and you guess at each question. Let X = the number of questions answered correctly.
6. In the next 20 births at a hospital, let X = the number of female births.
7. Of all patients suffering a particular illness, 35% experience improvement from a particular medication. In the next 100 patients administered the medication, let X = the number of patients who experience improvement.

3-6 Binomial Distribution

Definition

A random experiment consists of n Bernoulli trials such that

(1) The trials are independent

(2) Each trial results in only two possible outcomes, labeled as "success" and "failure"

(3) The probability of a success in each trial, denoted as p, remains constant

The random variable X that equals the number of trials that result in a success has a **binomial random variable** with parameters $0 < p < 1$ and $n = 1, 2, \ldots$. The probability mass function of X is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \ldots, n \qquad (3\text{-}7)$$

3-6 Binomial Distribution

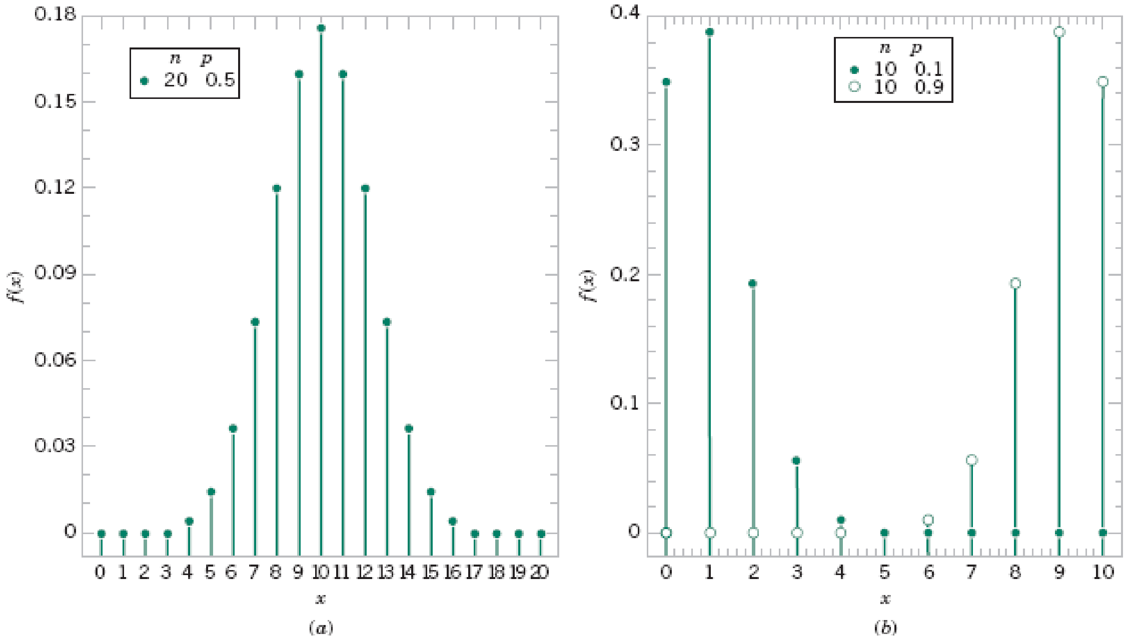


Figure 3-8 Binomial distributions for selected values of *n* and *p*.

3-6 Binomial Distribution

Example 3-18

Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. Find the probability that in the next 18 samples, exactly 2 contain the pollutant.

Let X = the number of samples that contain the pollutant in the next 18 samples analyzed. Then X is a binomial random variable with $p = 0.1$ and $n = 18$. Therefore,

$$P(X = 2) = \binom{18}{2}(0.1)^2(0.9)^{16}$$

Now $\binom{18}{2} = 18!/[2!\ 16!] = 18(17)/2 = 153$. Therefore,

$$P(X = 2) = 153(0.1)^2(0.9)^{16} = 0.284$$

3-6 Binomial Distribution

Example 3-18

Determine the probability that at least four samples contain the pollutant. The requested probability is

$$P(X \geq 4) = \sum_{x=4}^{18} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$

However, it is easier to use the complementary event,

$$P(X \geq 4) = 1 - P(X < 4) = 1 - \sum_{x=0}^{3} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$
$$= 1 - [0.150 + 0.300 + 0.284 + 0.168] = 0.098$$

Determine the probability that $3 \leq X < 7$. Now

$$P(3 \leq X < 7) = \sum_{x=3}^{6} \binom{18}{x} (0.1)^x (0.9)^{18-x}$$
$$= 0.168 + 0.070 + 0.022 + 0.005$$
$$= 0.265$$

3-6 Binomial Distribution

Mean and Variance

If X is a binomial random variable with parameters p and n,

$$\mu = E(X) = np \quad \text{and} \quad \sigma^2 = V(X) = np(1-p) \qquad (3\text{-}8)$$

3-6 Binomial Distribution

Example 3-19

For the number of transmitted bits received in error in Example 3-16, $n = 4$ and $p = 0.1$, so

$$E(X) = 4(0.1) = 0.4 \qquad \text{and} \qquad V(X) = 4(0.1)(0.9) = 0.36$$

and these results match those obtained from a direct calculation in Example 3-9.

3-7 Geometric and Negative Binomial Distributions

Example 3-20

The probability that a bit transmitted through a digital transmission channel is received in error is 0.1. Assume the transmissions are independent events, and let the random variable X denote the number of bits transmitted *until* the first error.

Then, $P(X = 5)$ is the probability that the first four bits are transmitted correctly and the fifth bit is in error. This event can be denoted as $\{OOOOE\}$, where O denotes an okay bit. Because the trials are independent and the probability of a correct transmission is 0.9,

$$P(X = 5) = P(OOOOE) = 0.9^4 0.1 = 0.066$$

Note that there is some probability that X will equal any integer value. Also, if the first trial is a success, $X = 1$. Therefore, the range of X is $\{1, 2, 3, \ldots\}$, that is, all positive integers.

3-7 Geometric and Negative Binomial Distributions

Definition

In a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote the number of trials until the first success. Then X is a **geometric random variable** with parameter $0 < p < 1$ and

$$f(x) = (1 - p)^{x-1}p \qquad x = 1, 2, \ldots \tag{3-9}$$

3-7 Geometric and Negative Binomial Distributions

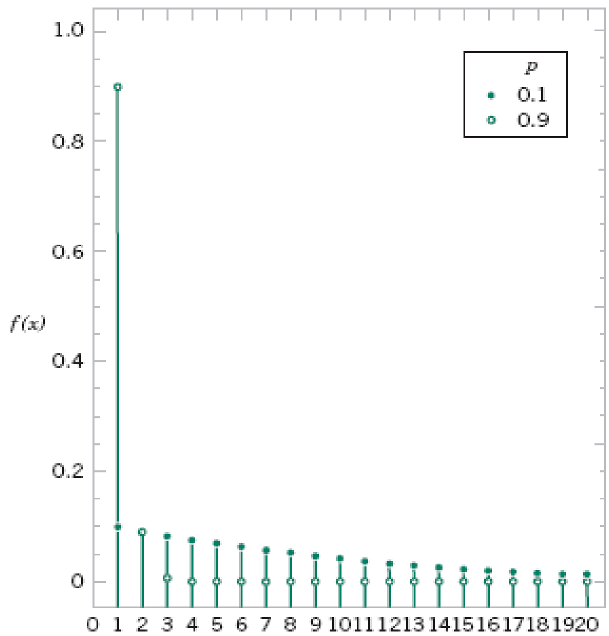


Figure 3-9. Geometric distributions for selected values of the parameter p.

3-7 Geometric and Negative Binomial Distributions

3-7.1 Geometric Distribution

Example 3-21

The probability that a wafer contains a large particle of contamination is 0.01. If it is assumed that the wafers are independent, what is the probability that exactly 125 wafers need to be analyzed before a large particle is detected?

Let X denote the number of samples analyzed until a large particle is detected. Then X is a geometric random variable with $p = 0.01$. The requested probability is

$$P(X = 125) = (0.99)^{124}0.01 = 0.0029$$

3-7 Geometric and Negative Binomial Distributions

Definition

If X is a geometric random variable with parameter p,

$$\mu = E(X) = 1/p \quad \text{and} \quad \sigma^2 = V(X) = (1-p)/p^2 \qquad (3\text{-}10)$$

3-7 Geometric and Negative Binomial Distributions

Lack of Memory Property

A geometric random variable has been defined as the number of trials until the first success. However, because the trials are independent, the count of the number of trials until the next success can be started at any trial without changing the probability distribution of the random variable. For example, in the transmission of bits, if 100 bits are transmitted, the probability that the first error, after bit 100, occurs on bit 106 is the probability that the next six outcomes are *OOOOOE*. This probability is $(0.9)^5(0.1) = 0.059$, which is identical to the probability that the initial error occurs on bit 6.

The implication of using a geometric model is that the system presumably will not wear out. The probability of an error remains constant for all transmissions. In this sense, the geometric distribution is said to lack any memory. The **lack of memory property** will be discussed again in the context of an exponential random variable in Chapter 4.

3-7 Geometric and Negative Binomial Distributions

3-7.2 Negative Binomial Distribution

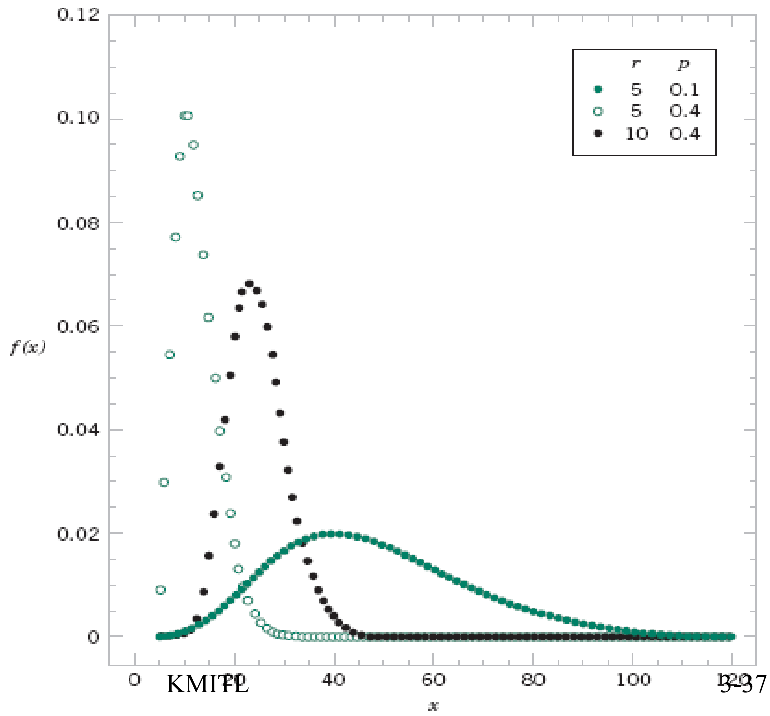
A generalization of a geometric distribution in which the random variable is the number of Bernoulli trials required to obtain r successes results in the **negative binomial distribution**.

In a series of Bernoulli trials (independent trials with constant probability p of a success), let the random variable X denote the number of trials until r successes occur. Then X is a **negative binomial random variable** with parameters $0 < p < 1$ and $r = 1, 2, 3, \ldots$, and

$$f(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r \qquad x = r, r+1, r+2, \ldots \qquad (3\text{-}11)$$

3-7 Geometric and Negative Binomial Distributions

Figure 3-10. Negative binomial distributions for selected values of the parameters *r* and *p*.



3-7 Geometric and Negative Binomial Distributions

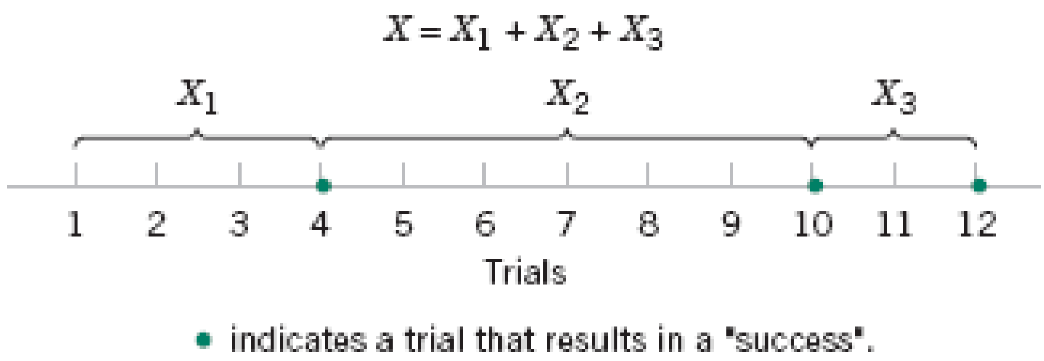


Figure 3-11. Negative binomial random variable represented as a sum of geometric random variables.

3-7 Geometric and Negative Binomial Distributions

3-7.2 Negative Binomial Distribution

If X is a negative binomial random variable with parameters p and r,

$$\mu = E(X) = r/p \qquad \text{and} \qquad \sigma^2 = V(X) = r(1-p)/p^2 \qquad (3\text{-}12)$$

3-7 Geometric and Negative Binomial Distributions

Example 3-25

A Web site contains three identical computer servers. Only one is used to operate the site, and the other two are spares that can be activated in case the primary system fails. The probability of a failure in the primary computer (or any activated spare system) from a request for service is 0.0005. Assuming that each request represents an independent trial, what is the mean number of requests until failure of all three servers?

Let X denote the number of requests until all three servers fail, and let X_1, X_2, and X_3 denote the number of requests before a failure of the first, second, and third servers used, respectively. Now, $X = X_1 + X_2 + X_3$. Also, the requests are assumed to comprise independent trials with constant probability of failure $p = 0.0005$. Furthermore, a spare server is not affected by the number of requests before it is activated. Therefore, X has a negative binomial distribution with $p = 0.0005$ and $r = 3$. Consequently,

$$E(X) = 3/0.0005 = 6000 \text{ requests}$$

3-7 Geometric and Negative Binomial Distributions

Example 3-25

What is the probability that all three servers fail within five requests? The probability is $P(X \le 5)$ and

$$\begin{aligned} P(X \le 5) &= P(X = 3) + P(X = 4) + P(X = 5) \\ &= 0.0005^3 + \binom{3}{2} 0.0005^3(0.9995) + \binom{4}{2} 0.0005^3(0.9995)^2 \\ &= 1.25 \times 10^{-10} + 3.75 \times 10^{-10} + 7.49 \times 10^{-10} \\ &= 1.249 \times 10^{-9} \end{aligned}$$

3-8 Hypergeometric Distribution

Definition

A set of N objects contains

K objects classified as successes

$N - K$ objects classified as failures

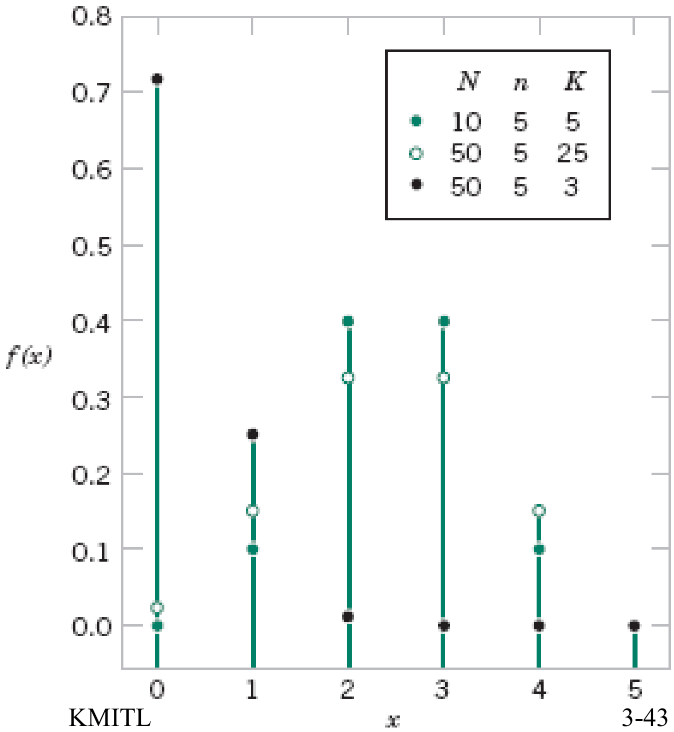
A sample of size n objects is selected randomly (without replacement) from the N objects, where $K \le N$ and $n \le N$.

Let the random variable X denote the number of successes in the sample. Then X is a **hypergeometric random variable** and

$$f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \qquad x = \max\{0, n + K - N\} \text{ to } \min\{K, n\} \tag{3-13}$$

3-8 Hypergeometric Distribution

Figure 3-12. Hypergeometric distributions for selected values of parameters *N*, *K*, and *n*.



3-8 Hypergeometric Distribution

Example 3-27

A batch of parts contains 100 parts from a local supplier of tubing and 200 parts from a supplier of tubing in the next state. If four parts are selected randomly and without replacement, what is the probability they are all from the local supplier?

Let X equal the number of parts in the sample from the local supplier. Then, X has a hypergeometric distribution and the requested probability is $P(X = 4)$. Consequently,

$$P(X = 4) = \frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}} = 0.0119$$

3-8 Hypergeometric Distribution

Example 3-27

What is the probability that two or more parts in the sample are from the local supplier?

$$P(X \geq 2) = \frac{\binom{100}{2}\binom{200}{2}}{\binom{300}{4}} + \frac{\binom{100}{3}\binom{200}{1}}{\binom{300}{4}} + \frac{\binom{100}{4}\binom{200}{0}}{\binom{300}{4}}$$

$$= 0.298 + 0.098 + 0.0119 = 0.408$$

What is the probability that at least one part in the sample is from the local supplier?

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{100}{0}\binom{200}{4}}{\binom{300}{4}} = 0.804$$

3-8 Hypergeometric Distribution

Mean and Variance

If X is a hypergeometric random variable with parameters N, K, and n, then

$$\mu = E(X) = np \quad \text{and} \quad \sigma^2 = V(X) = np(1-p)\left(\frac{N-n}{N-1}\right) \quad (3\text{-}14)$$

where $p = K/N$.

Here p is interpreted as the proportion of successes in the set of N objects.

3-8 Hypergeometric Distribution

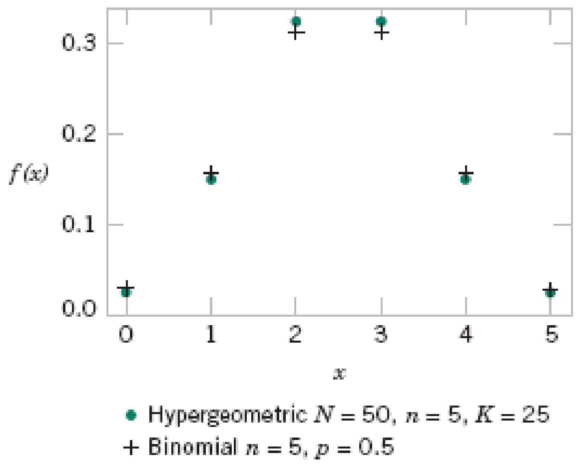
Finite Population Correction Factor

The term in the variance of a hypergeometric random variable

$$\frac{N-n}{N-1} \tag{3-15}$$

is called the finite population correction factor.

3-8 Hypergeometric Distribution



	0	1	2	3	4	5
Hypergeometric probability	0.025	0.149	0.326	0.326	0.149	0.025
Binomial probability	0.031	0.156	0.321	0.312	0.156	0.031

Figure 3-13. Comparison of hypergeometric and binomial distributions.

3-9 Poisson Distribution

Example 3-30

Consider the transmission of n bits over a digital communication channel. Let the random variable X equal the number of bits in error. When the probability that a bit is in error is constant and the transmissions are independent, X has a binomial distribution. Let p denote the probability that a bit is in error. Let $\lambda = pn$. Then, $E(x) = pn = \lambda$ and

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} = \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Now, suppose that the number of bits transmitted increases and the probability of an error decreases exactly enough that pn remains equal to a constant. That is, n increases and p decreases accordingly, such that $E(X) = \lambda$ remains constant. Then, with some work, it can be shown that

$$\binom{n}{x} \left(\frac{1}{n}\right)^x \to 1 \qquad \left(1 - \frac{\lambda}{n}\right)^{-x} \to 1 \qquad \left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda}$$

so that

$$\lim_{n \to \infty} P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

Also, because the number of bits transmitted tends to infinity, the number of errors can equal any nonnegative integer. Therefore, the range of X is the integers from zero to infinity.

3-9 Poisson Distribution

Definition

Given an interval of real numbers, assume events occur at random throughout the interval. If the interval can be partitioned into subintervals of small enough length such that

(1) the probability of more than one event in a subinterval is zero,

(2) the probability of one event in a subinterval is the same for all subintervals and proportional to the length of the subinterval, and

(3) the event in each subinterval is independent of other subintervals, the random experiment is called a **Poisson process.**

The random variable X that equals the number of events in the interval is a **Poisson random variable** with parameter $0 < \lambda$, and the probability mass function of X is

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \qquad x = 0, 1, 2, \ldots \qquad (3\text{-}16)$$

3-9 Poisson Distribution

Consistent Units

It is important to **use consistent units** in the calculation of probabilities, means, and variances involving Poisson random variables. The following example illustrates unit conversions. For example, if the

average number of flaws per millimeter of wire is 3.4, then the
average number of flaws in 10 millimeters of wire is 34, and the
average number of flaws in 100 millimeters of wire is 340.

3-9 Poisson Distribution

Example 3-33

Contamination is a problem in the manufacture of optical storage disks. The number of particles of contamination that occur on an optical disk has a Poisson distribution, and the average number of particles per centimeter squared of media surface is 0.1. The area of a disk under study is 100 squared centimeters. Find the probability that 12 particles occur in the area of a disk under study.

Let X denote the number of particles in the area of a disk under study. Because the mean number of particles is 0.1 particles per cm^2

$$E(X) = 100 \text{ cm}^2 \times 0.1 \text{ particles/cm}^2 = 10 \text{ particles}$$

Therefore,

$$P(X = 12) = \frac{e^{-10}10^{12}}{12!} = 0.095$$

3-9 Poisson Distribution

Example 3-33

The probability that zero particles occur in the area of the disk under study is

$$P(X = 0) = e^{-10} = 4.54 \times 10^{-5}$$

Determine the probability that 12 or fewer particles occur in the area of the disk under study. The probability is

$$P(X \le 12) = P(X = 0) + P(X = 1) + \cdots + P(X = 12) = \sum_{i=0}^{12} \frac{e^{-10}10^i}{i!}$$

3-9 Poisson Distribution

Mean and Variance

If X is a Poisson random variable with parameter λ, then

$$\mu = E(X) = \lambda \quad \text{and} \quad \sigma^2 = V(X) = \lambda \qquad (3\text{-}17)$$