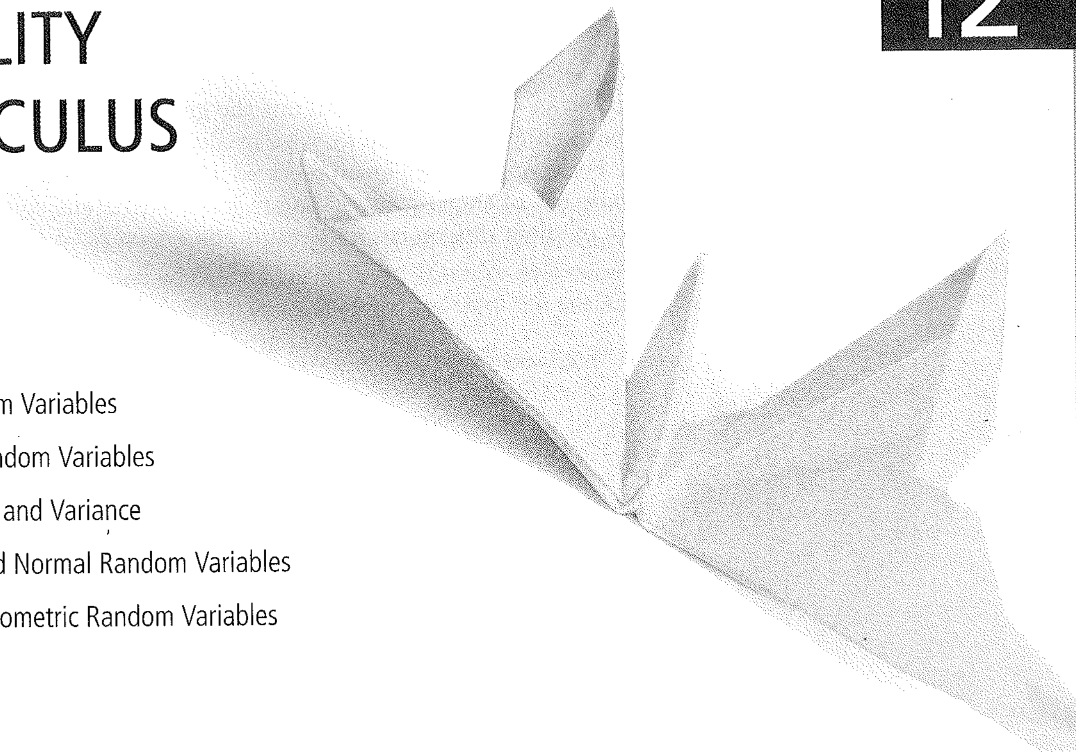
CHAPTER 12

PROBABILITY AND CALCULUS



In this chapter we shall survey a few applications of calculus to the theory of probability. Since we do not intend this chapter to be a self-contained course in probability, we shall select only a few salient ideas to present a taste of probability theory and provide a starting point for further study.

12.1 Discrete Random Variables

We will motivate the concepts of mean, variance, standard deviation, and random variable by analyzing examination grades.

Suppose that the grades on an exam taken by 10 people are 50, 60, 60, 70, 70, 90, 100, 100, 100, 100. This information is displayed in a frequency table in Fig. 1.

One of the first things we do when looking over the results of an exam is to compute the *mean* or *average* of the grades. We do this by totaling the grades and dividing by the number of people. This is the same as multiplying each distinct grade by the frequency with which it occurs, adding up those products, and dividing by the sum of the frequencies:

$$[\text{mean}] = \frac{50\cdot 1 + 60\cdot 2 + 70\cdot 2 + 90\cdot 1 + 100\cdot 4}{10} = \frac{800}{10} = 80.$$

Grade	50	60	70	90	100
Frequency	1	2	2	1	4

Figure 1

[Grade] − [Mean]	−30	−20	−10	10	20
Frequency	1	2	2	1	4

Figure 2

To get an idea of how spread out the grades are, we can compute the difference between each grade and the average grade. We have tabulated these differences in Fig. 2. For example, if a person received a 50, then [grade] − [mean] is $50 - 80 = -30$. As a measure of the spread of the grades, statisticians compute the average of the squares of these differences and call it the *variance* of the grade distribution. We have

$$[\text{variance}] = \frac{(-30)^2 \cdot 1 + (-20)^2 \cdot 2 + (-10)^2 \cdot 2 + (10)^2 \cdot 1 + (20)^2 \cdot 4}{10}$$
$$= \frac{900 + 800 + 200 + 100 + 1600}{10} = \frac{3600}{10} = 360.$$

The square root of the variance is called the *standard deviation* of the grade distribution. In this case, we have

$$[\text{standard deviation}] = \sqrt{360} \approx 18.97.$$

There is another way of looking at the grade distribution and its mean and variance. This new point of view is useful because it can be generalized to other situations. We begin by converting the frequency table to a relative frequency table. (See Fig. 3.) Below each grade we list the fraction of the class receiving that grade. The grade of 50 occurred $\frac{1}{10}$ of the time, the grade of 60 occurred $\frac{2}{10}$ of the time, and so on. Note that the relative frequencies add up to 1, because they represent the various fractions of the class grouped by test scores.

Grade	50	60	70	90	100
Relative frequency	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{4}{10}$

Figure 3

It is sometimes helpful to display the data in the relative frequency table by constructing a *relative frequency histogram.* (See Fig. 4.) Over each grade we place a rectangle whose height equals the relative frequency of that grade.

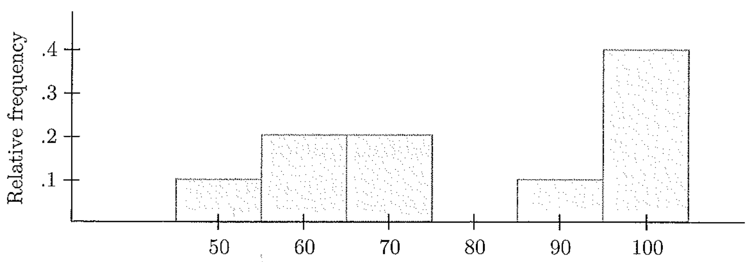


Figure 4. A relative frequency histogram.

An alternative way to compute the mean grade is

$$\begin{aligned}[\text{mean}] &= \frac{50\cdot 1 + 60\cdot 2 + 70\cdot 2 + 90\cdot 1 + 100\cdot 4}{10} \\ &= 50\cdot\frac{1}{10} + 60\cdot\frac{2}{10} + 70\cdot\frac{2}{10} + 90\cdot\frac{1}{10} + 100\cdot\frac{4}{10} \\ &= 5 + 12 + 14 + 9 + 40 = 80.\end{aligned}$$

Looking at the second line of this computation, we see that the mean is a sum of the various grades times their relative frequencies. We say that the mean is the *weighted sum* of the grades. (Grades are weighted by their relative frequencies.)

In a similar manner we see that the variance is also a weighted sum.

$$\begin{aligned}[\text{variance}] &= \left[(50-80)^2\cdot 1 + (60-80)^2\cdot 2 + (70-80)^2\cdot 2\right. \\ &\quad \left. + (90-80)^2\cdot 1 + (100-80)^2\cdot 4\right]\frac{1}{10} \\ &= (50-80)^2\cdot\frac{1}{10} + (60-80)^2\cdot\frac{2}{10} + (70-80)^2\cdot\frac{2}{10} \\ &\quad + (90-80)^2\cdot\frac{1}{10} + (100-80)^2\cdot\frac{4}{10} \\ &= 90 + 80 + 20 + 10 + 160 = 360.\end{aligned}$$

The relative frequency table shown in Fig. 3 is also called a *probability table.* The reason for this terminology is as follows. Suppose that we perform an *experiment* that consists of picking an exam paper at random from among the 10 papers. If the experiment is repeated many times, we expect the grade of 50 to occur about one-tenth of the time, the grade of 60 about two-tenths of the time, and so on. We say that the *probability* of the grade of 50 being chosen is $\frac{1}{10}$, the probability of the grade of 60 being chosen is $\frac{2}{10}$, and so on. In other words, the probability associated with a given grade measures the likelihood that an exam having that grade is chosen.

In this section we consider various experiments described by probability tables similar to the one in Fig. 3. The results of these experiments will be numbers (such as the preceding exam scores) called the *outcomes* of the experiment. We will also be given the probability of each outcome, indicating the relative frequency with which the given outcome is expected to occur if the experiment is repeated very often. If the outcomes of an experiment are $a_1, a_2, \ldots, a_n$, with respective probabilities $p_1, p_2, \ldots, p_n$, we describe the experiment by a probability table. (See Fig. 5.) Since the probabilities indicate relative frequencies, we see that

$$0 \le p_i \le 1$$

and

$$p_1 + p_2 + \cdots + p_n = 1.$$

Outcome	a_1	a_2	a_3	$\ldots$	a_n
Probability	p_1	p_2	p_3	$\ldots$	p_n

Figure 5

The last equation indicates that the outcomes $a_1, \ldots, a_n$ comprise all possible results of the experiment. We will usually list the outcomes of our experiments in ascending order, so $a_1 < a_2 < \cdots < a_n$.

We may display the data of a probability table in a histogram that has a rectangle of height p_i over the outcome a_i. (See Fig. 6.)

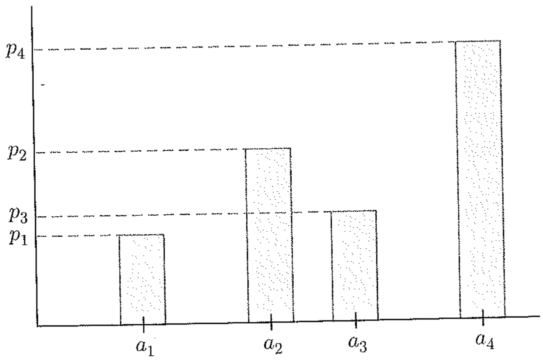


Figure 6

Let us define the *expected value* (or *mean*) of the probability table of Fig. 5 to be the weighted sum of the outcomes $a_1, \ldots, a_n$, each outcome weighted by the probability of its occurrence. That is,

$$[\text{expected value}] = a_1p_1 + a_2p_2 + \cdots + a_np_n.$$

Similarly, let us define the *variance* of the probability table to be the weighted sum of the squares of the differences between each outcome and the expected value. That is, if m denotes the expected value, then

$$[\text{variance}] = (a_1 - m)^2p_1 + (a_2 - m)^2p_2 + \cdots + (a_n - m)^2p_n.$$

To keep from writing the "outcome" so many times, we shall abbreviate by X the outcome of our experiment. That is, X is a variable that takes on the values $a_1, a_2, \ldots, a_n$ with respective probabilities $p_1, p_2, \ldots, p_n$. We will assume that our experiment is performed many times, being repeated in an unbiased (or random) way. Then X is a variable whose value depends on chance, and for this reason we say that X is a *random variable*. Instead of speaking of the expected value (mean) and the variance of a probability table, let us speak of the *expected value* and the *variance of the random variable* X that is associated with the probability table. We shall denote the expected value of X by $\mathrm{E}(X)$ and the variance of X by $\mathrm{Var}(X)$. The *standard deviation* of X is defined to be $\sqrt{\mathrm{Var}(X)}$.

EXAMPLE 1 One possible bet in roulette is to wager \$1 on red. The two possible outcomes are: lose \$1 and win \$1. These outcomes and their probabilities are given in Fig. 7. (*Note*: A roulette wheel in Las Vegas has 18 red numbers, 18 black numbers, and two green numbers.) Compute the expected value and the variance of the amount won.

Solution Let X be the random variable "amount won." Then

$$E(X) = -1 \cdot \frac{20}{38} + 1 \cdot \frac{18}{38} = -\frac{2}{38} \approx -.0526,$$

$$\mathrm{Var}(X) = \left[-1 - \left(-\frac{2}{38}\right)\right]^2 \cdot \frac{20}{38} + \left[1 - \left(-\frac{2}{38}\right)\right]^2 \cdot \frac{18}{38}$$

$$= \left(-\frac{36}{38}\right)^2 \cdot \frac{20}{38} + \left(\frac{40}{38}\right)^2 \cdot \frac{18}{38} \approx .997.$$

Amount won	−1	1
Probability	$\frac{20}{38}$	$\frac{18}{38}$

Figure 7. Las Vegas roulette.

The expected value of the amount won is approximately $-5\frac{1}{4}$ cents. In other words, sometimes we will win \$1 and sometimes we will lose \$1, but in the long run we can expect to lose an average of about $5\frac{1}{4}$ cents for each time we bet. ■

EXAMPLE 2 An experiment consists of selecting a number at random from the set of integers $\{1, 2, 3\}$. The probabilities are given by the table in Fig. 8. Let X designate the outcome. Find the expected value and the variance of X.

Number	1	2	3
Probability	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Figure 8

Solution

$$E(X) = 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{3} = 2,$$

$$\mathrm{Var}(X) = (1-2)^2 \cdot \frac{1}{3} + (2-2)^2 \cdot \frac{1}{3} + (3-2)^2 \cdot \frac{1}{3}$$

$$= (-1)^2 \cdot \frac{1}{3} + 0 + (1)^2 \cdot \frac{1}{3} = \frac{2}{3}.$$

EXAMPLE 3 A cement company plans to bid on a contract for constructing the foundations of new homes in a housing development. The company is considering two bids: a high bid that will produce \$75,000 profit (if the bid is accepted) and a low bid that will produce \$40,000 profit. From past experience the company estimates that the high bid has a 30% chance of acceptance and the low bid a 50% chance. Which bid should the company make?

Solution The standard method of decision is to choose the bid that has the higher expected value. Let X be the amount the company makes if it submits the high bid, and let Y be the amount it makes if it submits the low bid. Then the company must analyze the situation using the probability table shown in Table 1. The expected values are

$$E(X) = (75{,}000)(.30) + 0(.70) = 22{,}500,$$

$$E(Y) = (40{,}000)(.50) + 0(.50) = 20{,}000.$$

If the cement company has many opportunities to bid on similar contracts, a high bid each time will be accepted sufficiently often to produce an average profit of \$22,500 per bid. A consistently low bid will produce an average profit of \$20,000 per bid. Thus the company should submit the high bid. ■

TABLE 1 Bids on a Cement Contract

	High Bid			*Low Bid*	
	Accepted	**Rejected**		**Accepted**	**Rejected**
Value of X	75,000	0	Value of Y	40,000	0
Probability	.30	.70	Probability	.50	.50

When a probability table contains a large number of possible outcomes of an experiment, the associated histogram for the random variable X becomes a valuable aid for visualizing the data in the table. Look at Fig. 9, for example. Since the rectangles that make up the histogram all have the same width, their areas are in the same ratios as their heights. By an appropriate change of scale on the y-axis, we may assume that the *area* (instead of the height) of each rectangle gives the associated probability of X. Such a histogram is sometimes referred to as a *probability density histogram.*

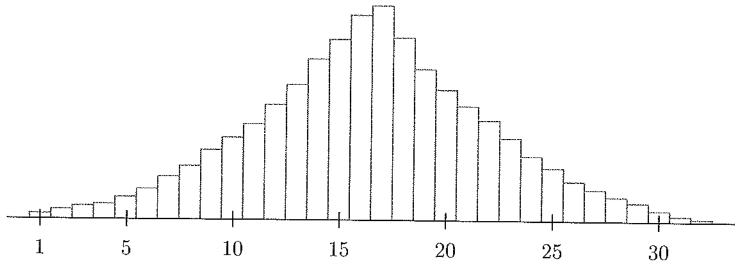


Figure 9. Probabilities displayed as areas.

A histogram that displays probabilities as areas is useful when we wish to visualize the probability that X has a value between two specified numbers. For example, in Fig. 9 suppose that the probabilities associated with $X = 5, X = 6, \ldots, X = 10$ are $p_5, p_6, \ldots, p_{10}$, respectively. Then the probability that X lies between 5 and 10 inclusive is $p_5 + p_6 + \cdots + p_{10}$. In terms of areas, this probability is just the total area of those rectangles over the values $5, 6, \ldots, 10$. (See Fig. 10.) We will consider analogous situations in the next section.

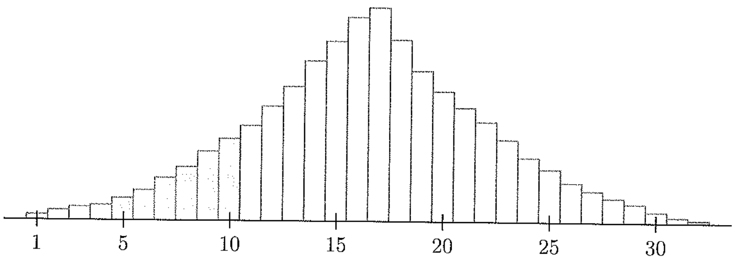


Figure 10. Probability that $5 \leq x \leq 10$.

Practice Problems 12.1

1. Compute the expected value and the variance of the random variable X with Table 2 as its probability table.

TABLE 2

Value of X	-1	0	1	2
Probability	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$

2. The production department at a radio factory sends CB radios to the inspection department in lots of 100. There an inspector examines three radios at random from each lot. If at least one of the three radios is defective and needs adjustment, the entire lot is sent back to the production department. Records of the inspection department show that the number X of defective radios in a sample of three radios has Table 3 as its probability table.

TABLE 3 Quality Control Data

Defectives	0	1	2	3
Probability	.7265	.2477	.0251	.0007

(a) What percentage of the lots does the inspection department reject?

(b) Find the mean number of defective radios in the samples of three radios.

(c) Based on the evidence in part (b), estimate the average number of defective radios in each lot of 100 radios.

EXERCISES 12.1

1. Table 4 is the probability table for a random variable X. Find $E(X)$, $Var(X)$, and the standard deviation of X.

TABLE 4

Outcome	0	1
Probability	$\frac{1}{5}$	$\frac{4}{5}$

2. Find $E(X)$, $Var(X)$, and the standard deviation of X, where X is the random variable whose probability table is given in Table 5.

TABLE 5

Outcome	1	2	3
Probability	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$

3. Compute the variances of the three random variables whose probability tables are given in Table 6. Relate the sizes of the variances to the spread of the values of the random variable.

TABLE 6

	Outcome	Probability
(a)	4	.5
	6	.5
(b)	3	.5
	7	.5
(c)	1	.5
	9	.5

4. Compute the variances of the two random variables whose probability tables are given in Table 7. Relate the sizes of the variances to the spread of the values of the random variables.

TABLE 7

	Outcome	Probability
(a)	2	.1
	4	.4
	6	.4
	8	.1
(b)	2	.3
	4	.2
	6	.2
	8	.3

5. The number of accidents per week at a busy intersection was recorded for a year. There were 11 weeks with no accidents, 26 weeks with one accident, 13 weeks with two accidents, and 2 weeks with three accidents. A week is to be selected at random and the number of accidents noted. Let X be the outcome. Then X is a random variable taking on the values 0, 1, 2, and 3.

(a) Write out a probability table for X.

(b) Compute $E(X)$. (c) Interpret $E(X)$.

6. The number of phone calls coming into a telephone switchboard during each minute was recorded during an entire hour. During 30 of the 1-minute intervals there were no calls, during 20 intervals there was one call, and during 10 intervals there were two calls. A 1-minute interval is to be selected at random and the number of calls noted. Let X be the outcome. Then X is a random variable taking on the values 0, 1, and 2.

(a) Write out a probability table for X.

(b) Compute $\mathrm{E}(X)$. (c) Interpret $\mathrm{E}(X)$.

7. Consider a circle with radius 1.

(a) What percentage of the points lies within $\frac{1}{2}$ unit of the center?

(b) Let c be a constant with $0 < c < 1$. What percentage of the points lies within c units of the center?

8. Consider a circle with circumference 1. An arrow (or spinner) is attached at the center so that, when flicked, it spins freely. Upon stopping, it points to a particular point on the circumference of the circle. Determine the likelihood that the point is

(a) On the top half of the circumference.

(b) On the top quarter of the circumference.

(c) On the top one-hundredth of the circumference.

(d) Exactly at the top of the circumference.

9. A citrus grower anticipates a profit of \$100,000 this year if the nightly temperatures remain mild. Unfortunately, the weather forecast indicates a 25% chance that the temperatures will drop below freezing during the next week. Such freezing weather will destroy 40% of the crop and reduce the profit to \$60,000. However, the grower can protect the citrus fruit against the possible freezing (using smudge pots, electric fans, and so on) at a cost of \$5000. Should the grower spend the \$5000 and thereby reduce the profit to \$95,000? [*Hint*: Compute $\mathrm{E}(X)$, where X is the profit the grower will get if he does nothing to protect the fruit.]

10. Suppose that the weather forecast in Exercise 9 indicates a 10% chance that cold weather will reduce the citrus grower's profit from \$100,000 to \$85,000 and a 10% chance that cold weather will reduce the profit to \$75,000. Should the grower spend \$5000 to protect the citrus fruit against the possible bad weather?

Solutions to Practice Problems 12.1

1. $\mathrm{E}(X) = (-1)\cdot\frac{1}{8} + 0\cdot\frac{1}{8} + 1\cdot\frac{3}{8} + 2\cdot\frac{3}{8} = 1,$

$$\begin{aligned}\mathrm{Var}(X) &= (-1-1)^2\cdot\tfrac{1}{8} + (0-1)^2\cdot\tfrac{1}{8} \\ &\quad + (1-1)^2\cdot\tfrac{3}{8} + (2-1)^2\cdot\tfrac{3}{8} \\ &= 4\cdot\tfrac{1}{8} + 1\cdot\tfrac{1}{8} + 0 + 1\cdot\tfrac{3}{8} = 1.\end{aligned}$$

2. (a) In three cases a lot will be rejected: $X = 1$, 2, or 3. Adding the corresponding probabilities, we find that the probability of rejecting a lot is $.2477 + .0251 + .0007 = .2735$, or 27.35%. (An alternative method of solution uses the fact that the sum of the probabilities for *all* possible cases must be 1. From the table we see that the probability of accepting a lot is .7265, so the probability of rejecting a lot is $1 - .7265 = .2735$.)

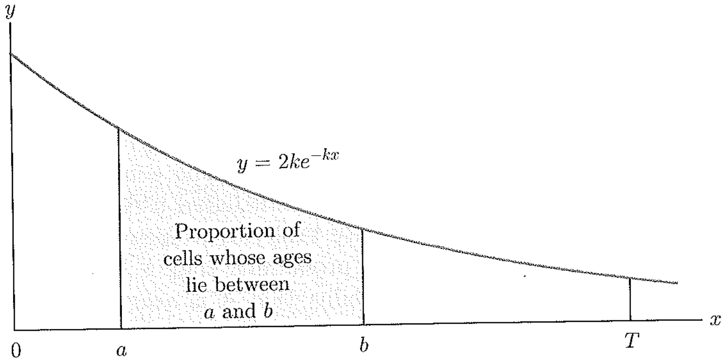
(b) $$\begin{aligned}\mathrm{E}(X) &= 0(.7265) + 1(.2477) \\ &\quad + 2(.0251) + 3(.0007) \\ &= .3000.\end{aligned}$$

(c) In part (b) we found that an average of .3 radio in every sample of three radios is defective. Thus about 10% of the radios in the sample are defective. Since the samples are chosen at random, we may assume that about 10% of *all* the radios are defective. Thus we estimate that on the average 10 out of each lot of 100 radios will be defective.

12.2 Continuous Random Variables

Consider a cell population that is growing vigorously. When a cell is T days old it divides and forms two new daughter cells. If the population is sufficiently large, it will contain cells of many different ages between 0 and T. It turns out that the proportion of cells of various ages remains constant. That is, if a and b are any two numbers between 0 and T, with $a < b$, the proportion of cells whose ages lie between a and b is essentially constant from one moment to the next, even though individual cells are aging and new cells are being formed all the time. In fact, biologists have found that under the ideal circumstances described, the proportion of cells whose ages are between a and b is given by the area under the graph of the function $f(x) = 2ke^{-kx}$ from $x = a$ to $x = b$, where $k = (\ln 2)/T$.* (See Fig. 1.)

*See J. R. Cook and T. W. James, "Age Distribution of Cells in Logarithmically Growing Cell Populations," in *Synchrony in Cell Division and Growth*, Erik Zeuthen, ed. (New York: John Wiley & Sons, 1964), pp. 485–495.



1. Age distribution in a population.

Now consider an experiment in which we select a cell at random from the population and observe its age, X. Then the probability that X lies between a and b is given by the area under the graph of $f(x) = 2ke^{-kx}$ from a to b. (See Fig. 1.) Let us denote this probability by $\Pr(a \le X \le b)$. Using the fact that the area under the graph of $f(x)$ is given by a definite integral, we have

$$\Pr(a \le X \le b) = \int_a^b f(x)\,dx = \int_a^b 2ke^{-kx}\,dx. \tag{1}$$

Since X can assume any one of the (infinitely many) numbers in the continuous interval from 0 to T, we say that X is a *continuous random variable.* The function $f(x)$ that determines the probability in (1) for each a and b is called the (*probability*) *density function* of X (or of the experiment whose outcome is X).

More generally, consider an experiment whose outcome may be any value between A and B. The outcome of the experiment, denoted X, is called a *continuous random variable.* For the cell population described previously, $A = 0$ and $B = T$. Another typical experiment might consist of choosing a number X at random between $A = 5$ and $B = 6$. Or we could observe the duration X of a random telephone call passing through a given telephone switchboard. If we have no way of knowing how long a call might last, X might be any nonnegative number. In this case it is convenient to say that X lies between 0 and ∞ and to take $A = 0$ and $B = \infty$. On the other hand, if the possible values of X for some experiment include rather large negative numbers, we sometimes takes $A = -\infty$.

Given an experiment whose outcome is a continuous random variable X, the probability $\Pr(a \le X \le b)$ is a measure of the likelihood that an outcome of the experiment will lie between a and b. If the experiment is repeated many times, the proportion of times X has a value between a and b should be close to $\Pr(a \le X \le b)$. In experiments of practical interest involving a continuous random variable X, it is usually possible to find a function $f(x)$ such that

$$\Pr(a \le X \le b) = \int_a^b f(x)\,dx \tag{2}$$

for all a and b in the range of possible values of X. Such a function $f(x)$ is called a *probability density function* and satisfies the following properties:

(I) $f(x) \ge 0$ for $A \le x \le B$.

(II) $\int_A^B f(x) = 1.$

Indeed, Property I means that, for x between A and B, the graph of $f(x)$ must lie on or above the x-axis. Property II simply says that there is probability 1 that X has a value between A and B. (Of course, if $B = \infty$ and/or $A = -\infty$, the integral in Property II is an improper integral.) Properties I and II characterize probability density functions, in the sense that any function $f(x)$ satisfying I and II is the probability density function for some continuous random variable X. Moreover, $\Pr(a \le X \le b)$ can then be calculated using equation (2).

Unlike a probability table for a discrete random variable, a density function $f(x)$ does *not* give the probability that X has a certain value. Instead, $f(x)$ can be used to find the probability that X is *near* a specific value in the following sense. If x_0 is a number between A and B and if Δx is the width of a small interval centered at x_0, the probability that X is between $x_0 - \frac{1}{2}\Delta x$ and $x_0 + \frac{1}{2}\Delta x$ is approximately $f(x_0)\Delta x$, that is, the area of the rectangle shown in Fig. 2.

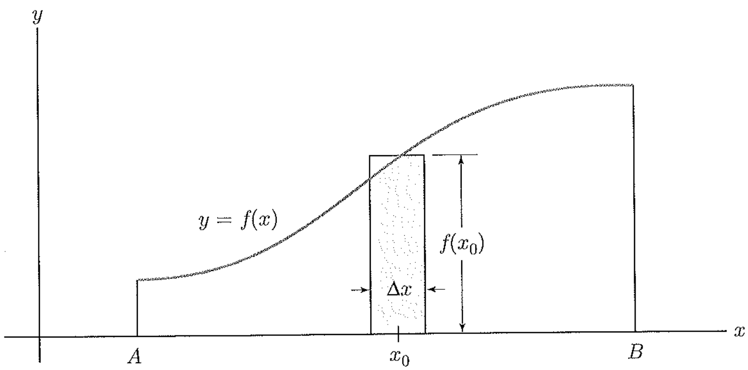


Figure 2. Area of rectangle gives approximate probability that X is near x_0.

EXAMPLE 1 Consider the cell population described earlier. Let $f(x) = 2ke^{-kx}$, where $k = (\ln 2)/T$. Show that $f(x)$ is indeed a probability density function on $0 \le x \le T$.

Solution Clearly, $f(x) \ge 0$, since $\ln 2$ is positive and the exponential function is never negative. Thus Property I is satisfied. For Property II we check that

$$\int_0^T f(x)\,dx = \int_0^T 2ke^{-kx}\,dx = -2e^{-kx}\Big|_0^T = -2e^{-kT} + 2e^0$$

$$= -2e^{-[(\ln 2)/T]T} + 2 = -2e^{-\ln 2} + 2$$

$$= -2e^{\ln(1/2)} + 2 = -2\left(\frac{1}{2}\right) + 2 = 1.$$

EXAMPLE 2 Let $f(x) = kx^2$.

(a) Find the value of k that makes $f(x)$ a probability density function on $0 \le x \le 4$.

(b) Let X be a continuous random variable whose density function is $f(x)$. Compute $\Pr(1 \le X \le 2)$.

Solution **(a)** We must have $k \ge 0$ so that Property I is satisfied. For Property II, we calculate

$$\int_0^4 f(x)\,dx = \int_0^4 kx^2\,dx = \frac{1}{3}kx^3\Big|_0^4 = \frac{1}{3}k(4)^3 - 0 = \frac{64}{3}k.$$

To satisfy Property II we must have $\frac{64}{3}k = 1$, or $k = \frac{3}{64}$. Thus $f(x) = \frac{3}{64}x^2$.

(b) $\Pr(1 \le X \le 2) = \int_1^2 f(x)\,dx = \int_1^2 \frac{3}{64}x^2\,dx = \frac{1}{64}x^3\Big|_1^2 = \frac{8}{64} - \frac{1}{64} = \frac{7}{64}.$

The area corresponding to this probability is shown in Fig. 3. ■

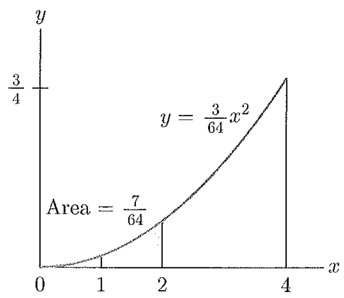


Figure 3

The density function in the next example is a special case of what statisticians sometimes call a ***beta* probability density**.

EXAMPLE 3 The parent corporation for a franchised chain of fast-food restaurants claims that the proportion of their new restaurants that make a profit during their first year of operation has the probability density function

$$f(x) = 12x(1-x)^2, \qquad 0 \le x \le 1.$$

(a) What is the probability that less than 40% of the restaurants opened this year will make a profit during their first year of operation?

(b) What is the probability that more than 50% of the restaurants will make a profit during their first year of operation?

Solution Let X be the proportion of new restaurants opened this year that make a profit during their first year of operation. Then the possible values of X range between 0 and 1.

(a) The probability that X is less than .4 equals the probability that X is between 0 and .4. We note that $f(x) = 12x(1 - 2x + x^2) = 12x - 24x^2 + 12x^3$ and therefore,

$$\Pr(0 \le X \le .4) = \int_0^{.4} f(x)\,dx = \int_0^{.4} (12x - 24x^2 + 12x^3)\,dx$$

$$= (6x^2 - 8x^3 + 3x^4)\Big|_0^{.4} = .5248.$$

(b) The probability that X is greater than .5 equals the probability that X is between .5 and 1. Thus

$$\Pr(.5 \le X \le 1) = \int_{.5}^{1} (12x - 24x^2 + 12x^3)\,dx$$

$$= (6x^2 - 8x^3 + 3x^4)\Big|_{.5}^{1} = .3125.$$ ■

Each probability density function is closely related to another important function called a cumulative distribution function. To describe this relationship, let us consider an experiment whose outcome is a continuous random variable X, with values between A and B, and let $f(x)$ be the associated density function. For each number x between A and B, let $F(x)$ be the probability that X is less than or equal to the number x. Sometimes we write $F(x) = \Pr(X \le x)$; however, since X is never less than A, we may also write

$$F(x) = \Pr(A \le X \le x). \tag{3}$$

Graphically, $F(x)$ is the area under the graph of the probability density function $f(x)$ from A to x. (See Fig. 4.) The function $F(x)$ is called the *cumulative distribution function* of the random variable X (or of the experiment whose outcome is X). Note that $F(x)$ also has the properties

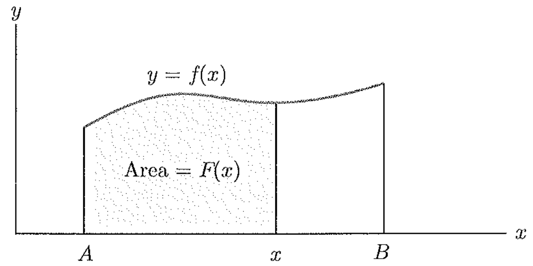


Figure 4. The cumulative distribution function $F(x)$.

$$F(A) = \Pr(A \le X \le A) = 0, \tag{4}$$
$$F(B) = \Pr(A \le X \le B) = 1. \tag{5}$$

Since $F(x)$ is an area function that gives the area under the graph of $f(x)$ from A to x, we know from Section 6.3 that $F(x)$ is an antiderivative of $f(x)$. That is,

$$F'(x) = f(x), \qquad A \le x \le B. \tag{6}$$

It follows that we may use $F(x)$ to compute probabilities, since

$$\Pr(a \le X \le b) = \int_a^b f(x)\,dx = F(b) - F(a), \tag{7}$$

for any a and b between A and B.

The relation (6) between $F(x)$ and $f(x)$ makes it possible to find one of these functions when the other is known, as we see in the following two examples.

EXAMPLE 4 Let X be the age of a cell selected at random from the cell population described earlier. The probability density function for X is $f(x) = 2ke^{-kx}$, where $k = (\ln 2)/T$. (See Fig. 5.) Find the cumulative distribution function $F(x)$ for X.

Solution Since $F(x)$ is an antiderivative of $f(x) = 2ke^{-kx}$, we have $F(x) = -2e^{-kx} + C$ for some constant C. Now $F(x)$ is defined for $0 \le x \le T$. Thus (4) implies that $F(0) = 0$. Setting $F(0) = -2e^0 + C = 0$, we find that $C = 2$, so

$$F(x) = -2e^{-kx} + 2.$$

(See Fig. 6.) ■

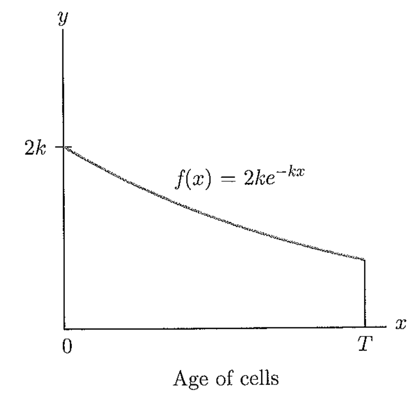


Figure 5. Probability density function.

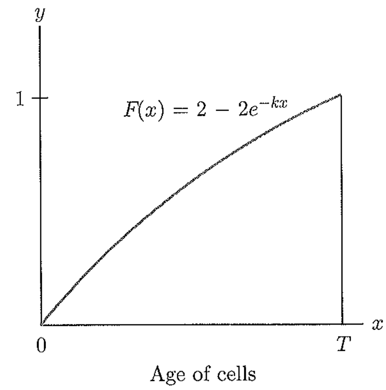


Figure 6. Cumulative distribution function.

EXAMPLE 5 Let X be the random variable associated with the experiment that consists of selecting a point at random from a circle of radius 1 and observing its distance from the center. Find the probability density function $f(x)$ and cumulative distribution function $F(x)$ of X.

Solution The distance of a point from the center of the unit circle is a number between 0 and 1. Suppose that $0 \leq x \leq 1$. Let us first compute the cumulative distribution function $F(x) = \Pr(0 \leq X \leq x)$. That is, let us find the probability that a point selected at random lies within x units of the center of the circle, in other words, lies inside the circle of radius x. See the shaded region in Fig. 7(b). Since the area of this shaded region is πx^2 and the area of the entire unit circle is $\pi \cdot 1^2 = \pi$, the proportion of points inside the shaded region is $\pi x^2/\pi = x^2$. Thus the probability is x^2 that a point selected at random will be in this shaded region. Hence

$$F(x) = x^2.$$

Differentiating, we find that the probability density function for X is

$$f(x) = F'(x) = 2x.$$

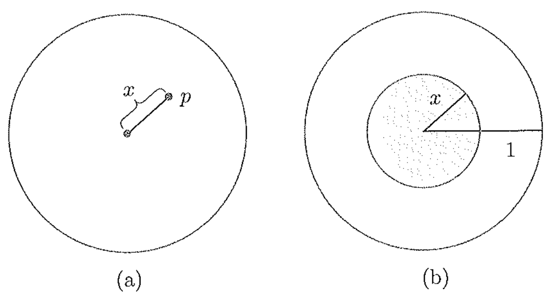


Figure 7

Our final example involves a continuous random variable X whose possible values lie between $A = 1$ and $B = \infty$; that is, X is any number greater than or equal to 1.

EXAMPLE 6 Let $f(x) = 3x^{-4}$, $x \geq 1$.

(a) Show that $f(x)$ is the probability density function of some random variable X.

(b) Find the cumulative distribution function $F(x)$ of X.

(c) Compute $\Pr(X \le 4)$, $\Pr(4 \le X \le 5)$, and $\Pr(4 \le X)$.

Solution (a) It is clear that $f(x) \ge 0$ for $x \ge 1$. Thus Property I holds. To check for Property II, we must compute

$$\int_1^\infty 3x^{-4}\,dx.$$

But

$$\int_1^b 3x^{-4}\,dx = -x^{-3}\Big|_1^b = -b^{-3} + 1 \to 1$$

as $b \to \infty$. Thus

$$\int_1^\infty 3x^{-4}\,dx = 1,$$

and Property II holds.

(b) Since $F(x)$ is an antiderivative of $f(x) = 3x^{-4}$, we have

$$F(x) = \int 3x^{-4}\,dx = -x^{-3} + C.$$

Since X has values greater than or equal to 1, we must have $F(1) = 0$. Setting $F(1) = -1 + C = 0$, we find that $C = 1$, so

$$F(x) = 1 - x^{-3}.$$

(c) $\Pr(X \le 4) = F(4) = 1 - 4^{-3} = 1 - \frac{1}{64} = \frac{63}{64}$.

Since we know $F(x)$, we may use it to compute $\Pr(4 \le X \le 5)$, as follows:

$$\Pr(4 \le X \le 5) = F(5) - F(4) = (1 - 5^{-3}) - (1 - 4^{-3})$$

$$= \frac{1}{4^3} - \frac{1}{5^3} \approx .0076.$$

We may compute $\Pr(4 \le X)$ directly by evaluating the improper integral

$$\int_4^\infty 3x^{-4}\,dx.$$

However, there is a simpler method. We know that

$$\int_1^4 3x^{-4}\,dx + \int_4^\infty 3x^{-4}\,dx = \int_1^\infty 3x^{-4}\,dx = 1. \tag{8}$$

In terms of probabilities, (8) may be written as

$$\Pr(X \le 4) + \Pr(4 \le X) = 1.$$

Hence

$$\Pr(4 \le X) = 1 - \Pr(X \le 4) = 1 - \frac{63}{64} = \frac{1}{64}.$$

Practice Problems 12.2

1. In a certain farming region and in a certain year, the number of bushels of wheat produced per acre is a random variable X with a density function

$$f(x) = \frac{x - 30}{50}, \qquad 30 \le x \le 40.$$

(a) What is the probability that an acre selected at random produced less than 35 bushels of wheat?

(b) If the farming region had 20,000 acres of wheat, how many acres produced less than 35 bushels of wheat per acre?

2. The density function for a continuous random variable X on the interval $1 \le x \le 2$ is $f(x) = 8/(3x^3)$. Find the corresponding cumulative distribution function for X.

EXERCISES 12.2

Verify that each of the following functions is a probability density function.

1. $f(x) = \frac{1}{18}x$, $0 \le x \le 6$

2. $f(x) = 2(x-1)$, $1 \le x \le 2$

3. $f(x) = \frac{1}{4}$, $1 \le x \le 5$

4. $f(x) = \frac{8}{9}x$, $0 \le x \le \frac{3}{2}$

5. $f(x) = 5x^4$, $0 \le x \le 1$

6. $f(x) = \frac{3}{2}x - \frac{3}{4}x^2$, $0 \le x \le 2$

In Exercises 7–12, find the value of k that makes the given function a probability density function on the specified interval.

7. $f(x) = kx$, $1 \le x \le 3$

8. $f(x) = kx^2$, $0 \le x \le 2$

9. $f(x) = k$, $5 \le x \le 20$

10. $f(x) = k/\sqrt{x}$, $1 \le x \le 4$

11. $f(x) = kx^2(1-x)$, $0 \le x \le 1$

12. $f(x) = k(3x - x^2)$, $0 \le x \le 3$

13. The density function of a continuous random variable X is $f(x) = \frac{1}{8}x$, $0 \le x \le 4$. Sketch the graph of $f(x)$ and shade in the areas corresponding to (a) $\Pr(X \le 1)$; (b) $\Pr(2 \le X \le 2.5)$; (c) $\Pr(3.5 \le X)$.

14. The density function of a continuous random variable X is $f(x) = 3x^2$, $0 \le x \le 1$. Sketch the graph of $f(x)$ and shade in the areas corresponding to (a) $\Pr(X \le .3)$; (b) $\Pr(.5 \le X \le .7)$; (c) $\Pr(.8 \le X)$.

15. Find $\Pr(1 \le X \le 2)$ when X is a random variable whose density function is given in Exercise 1.

16. Find $\Pr(1.5 \le X \le 1.7)$ when X is a random variable whose density function is given in Exercise 2.

17. Find $\Pr(X \le 3)$ when X is a random variable whose density function is given in Exercise 3.

18. Find $\Pr(1 \le X)$ when X is a random variable whose density function is given in Exercise 4.

19. Suppose that the lifetime X (in hours) of a certain type of flashlight battery is a random variable on the interval $30 \le x \le 50$ with density function $f(x) = \frac{1}{20}$, $30 \le x \le 50$. Find the probability that a battery selected at random will last at least 35 hours.

20. At a certain supermarket the amount of wait time at the express lane is a random variable with density function $f(x) = 11/[10(x+1)^2]$, $0 \le x \le 10$. (See Fig. 8.) Find the probability of having to wait less than 4 minutes at the express lane.

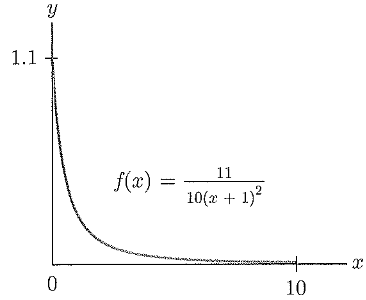


Figure 8. A density function.

21. The cumulative distribution function for a random variable X on the interval $1 \le x \le 5$ is $F(x) = \frac{1}{2}\sqrt{x-1}$. (See Fig. 9.) Find the corresponding density function.

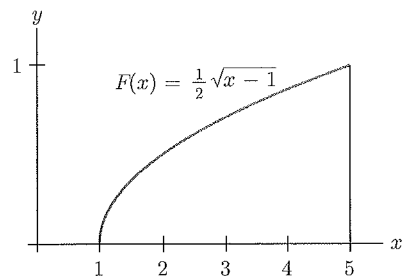


Figure 9. A cumulative distribution function.

22. The cumulative distribution function for a random variable X on the interval $1 \le x \le 2$ is $F(x) = \frac{4}{3} - 4/(3x^2)$. Find the corresponding density function.

23. Compute the cumulative distribution function corresponding to the density function $f(x) = \frac{1}{5}$, $2 \le x \le 7$.

24. Compute the cumulative distribution function corresponding to the density function $f(x) = \frac{1}{2}(3-x)$, $1 \le x \le 3$.

25. The time (in minutes) required to complete a certain subassembly is a random variable X with density function $f(x) = \frac{1}{21}x^2$, $1 \le x \le 4$.

(a) Use $f(x)$ to compute $\Pr(2 \le X \le 3)$.

(b) Find the corresponding cumulative distribution function $F(x)$.

(c) Use $F(x)$ to compute $\Pr(2 \le X \le 3)$.

26. The density function for a continuous random variable X on the interval $1 \le x \le 4$ is $f(x) = \frac{4}{9}x - \frac{1}{9}x^2$.

(a) Use $f(x)$ to compute $\Pr(3 \le X \le 4)$.

(b) Find the corresponding cumulative distribution function $F(x)$.

(c) Use $F(x)$ to compute $\Pr(3 \le X \le 4)$.

An experiment consists of selecting a point at random from the square in Fig. 10(a). Let X be the maximum of the coordinates of the point.

27. Show that the cumulative distribution function of X is $F(x) = x^2/4$, $0 \le x \le 2$.

28. Find the corresponding density function of X.

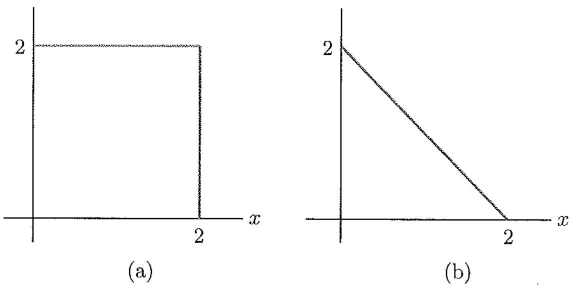


Figure 10

An experiment consists of selecting a point at random from the triangle in Fig. 10(b). Let X be the sum of the coordinates of the point.

29. Show that the cumulative distribution function of X is $F(x) = x^2/4$, $0 \le x \le 2$.

30. Find the corresponding density function of X.

In a certain cell population, cells divide every 10 days, and the age of a cell selected at random is a random variable X with the density function $f(x) = 2ke^{-kx}$, $0 \le x \le 10$, $k = (\ln 2)/10$.

31. Find the probability that a cell is at most 5 days old.

32. Upon examination of a slide, 10% of the cells are found to be undergoing mitosis (a change in the cell leading to division). Compute the length of time required for mitosis; that is, find the number M such that

$$\int_{10-M}^{10} 2ke^{-kx}\,dx = .10.$$

33. A random variable X has a density function $f(x) = \frac{1}{3}$, $0 \le x \le 3$. Find b such that $\Pr(0 \le X \le b) = .6$.

34. A random variable X has a density function $f(x) = \frac{2}{3}x$ on $1 \le x \le 2$. Find a such that $\Pr(a \le X) = \frac{1}{3}$.

35. A random variable X has a cumulative distribution function $F(x) = \frac{1}{4}x^2$ on $0 \le x \le 2$. Find b such that $\Pr(X \le b) = .09$.

36. A random variable X has a cumulative distribution function $F(x) = (x-1)^2$ on $1 \le x \le 2$. Find b such that $\Pr(X \le b) = \frac{1}{4}$.

37. Let X be a continuous random variable with values between $A = 1$ and $B = \infty$, and with the density function $f(x) = 4x^{-5}$.

(a) Verify that $f(x)$ is a probability density function for $x \ge 1$.

(b) Find the corresponding cumulative distribution function $F(x)$.

(c) Use $F(x)$ to compute $\Pr(1 \le X \le 2)$ and $\Pr(2 \le X)$.

38. Let X be a continuous random variable with the density function $f(x) = 2(x+1)^{-3}$, $x \ge 0$.

(a) Verify that $f(x)$ is a probability density function for $x \ge 0$.

(b) Find the cumulative distribution function for X.

(c) Compute $\Pr(1 \le X \le 2)$ and $\Pr(3 \le X)$.

Solutions to Practice Problems 12.2

1. (a) $$\Pr(X \le 35) = \int_{30}^{35} \frac{x-30}{50}\,dx = \frac{(x-30)^2}{100}\Bigg|_{30}^{35} = \frac{5^2}{100} - 0 = .25.$$

(b) Using part (a), we see that 25% of the 20,000 acres, or 5000 acres, produced less that 35 bushels of wheat per acre.

2. The cumulative distribution function $F(x)$ is an antiderivative of $f(x) = 8/(3x^3) = \frac{8}{3}x^{-3}$. Thus $F(x) = -\frac{4}{3}x^{-2} + C$ for some constant C. Since X varies over the interval $1 \le x \le 2$, we must have $F(1) = 0$; that is, $-\frac{4}{3}(1)^{-2} + C = 0$. Thus $C = \frac{4}{3}$, and

$$F(x) = \frac{4}{3} - \frac{4}{3}x^{-2}.$$

12.3 Expected Value and Variance

When studying the cell population described in Section 12.2, we might reasonably ask for the average age of the cells. In general, given an experiment described by a random variable X and a probability density function $f(x)$, it is often important to know the average outcome of the experiment and the degree to which the