

2.2: Continuous Density Functions

In the previous section we have seen how to simulate experiments with a whole continuum of possible outcomes and have gained some experience in thinking about such experiments. Now we turn to the general problem of assigning probabilities to the outcomes and events in such experiments. We shall restrict our attention here to those experiments whose sample space can be taken as a suitably chosen subset of the line, the plane, or some other Euclidean space. We begin with some simple examples.

Spinners

✓ Example 2.2.1

The spinner experiment described in Example 2.1 has the interval $[0, 1)$ as the set of possible outcomes. We would like to construct a probability model in which each outcome is equally likely to occur. We saw that in such a model, it is necessary to assign the probability 0 to each outcome. This does not at all mean that the probability of every event must be zero. On the contrary, if we let the random variable X denote the outcome, then the probability

$$P(0 \leq X \leq 1) \quad (2.2.1)$$

that the head of the spinner comes to rest somewhere in the circle, should be equal to 1. Also, the probability that it comes to rest in the upper half of the circle should be the same as for the lower half, so that

$$P\left(0 \leq X < \frac{1}{2}\right) = P\left(\frac{1}{2} \leq X < 1\right) = \frac{1}{2} \quad (2.2.2)$$

More generally, in our model, we would like the equation

$$P(c \leq X < d) = d - c \quad (2.2.3)$$

to be true for every choice of c and d .

If we let $E = [c, d]$, then we can write the above formula in the form

$$P(E) = \int_E f(x) dx \quad (2.2.4)$$

where $f(x)$ is the constant function with value 1. This should remind the reader of the corresponding formula in the discrete case for the probability of an event:

$$P(E) = \sum_{\omega \in E} m(\omega) \quad (2.2.5)$$

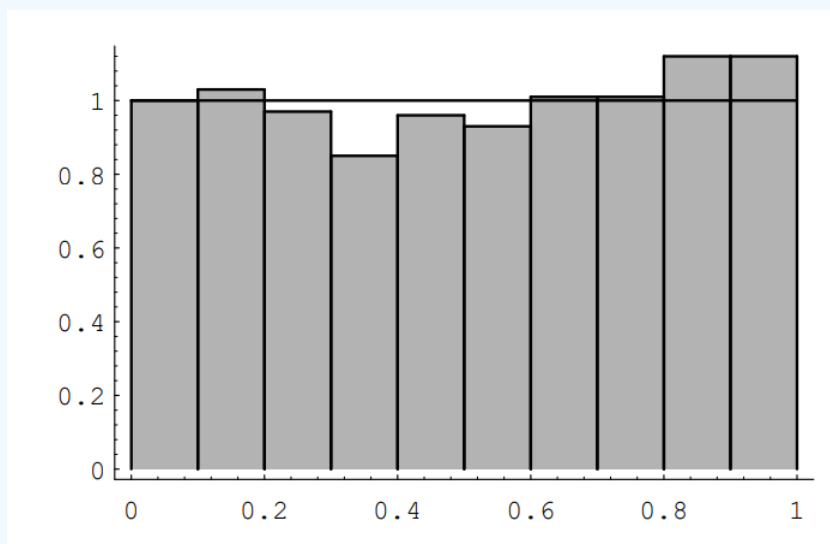


Figure 2.2.1: Spinner experiment.

The difference is that in the continuous case, the quantity being integrated, $f(x)$, is not the probability of the outcome x . (However, if one uses infinitesimals, one can consider $f(x) dx$ as the probability of the outcome x .) In the continuous case, we will use the following convention. If the set of outcomes is a set of real numbers, then the individual outcomes will be referred to by small Roman letters such as x . If the set of outcomes is a subset of \mathbb{R}^2 , then the individual outcomes will be denoted by (x, y) . In either case, it may be more convenient to refer to an individual outcome by using ω , as in Chapter 1. Figure 2.2.1 shows the results of 1000 spins of the spinner. The function $f(x)$ is also shown in the figure. The reader will note that the area under $f(x)$ and above a given interval is approximately equal to the fraction of outcomes that fell in that interval. The function $f(x)$ is called the density function of the random variable X . The fact that the area under $f(x)$ and above an interval corresponds to a probability is the defining property of density functions. A precise definition of density functions will be given shortly.

Darts

✓ Example 2.2.2

A game of darts involves throwing a dart at a circular target of unit radius. Suppose we throw a dart once so that it hits the target, and we observe where it lands. To describe the possible outcomes of this experiment, it is natural to take as our sample space the set Ω of all the points in the target. It is convenient to describe these points by their rectangular coordinates, relative to a coordinate system with origin at the center of the target, so that each pair (x, y) of coordinates with $x^2 + y^2 \leq 1$ describes a possible outcome of the experiment. Then $\Omega = \{(x, y) : x^2 + y^2 \leq 1\}$ is a subset of the Euclidean plane, and the event $E = \{(x, y) : y > 0\}$, for example, corresponds to the statement that the dart lands in the upper half of the target, and so forth. Unless there is reason to believe otherwise (and with experts at the game there may well be!), it is natural to assume that the coordinates are chosen at random. (When doing this with a computer, each coordinate is chosen uniformly from the interval $[-1, 1]$. If the resulting point does not lie inside the unit circle, the point is not counted.) Then the arguments used in the preceding example show that the probability of any elementary event, consisting of a single outcome, must be zero, and suggest that the probability of the event that the dart lands in any subset E of the target should be determined by what fraction of the target area lies in E . Thus,

$$P(E) = \frac{\text{area of } E}{\text{area of target}} = \frac{\text{area of } E}{\pi} \quad (2.2.6)$$

This can be written in the form

$$P(E) = \int_E f(x) dx \quad (2.2.7)$$

where $f(x)$ is the constant function with value $1/\pi$. In particular, if $E = \{(x, y) : x^2 + y^2 \leq a^2\}$ is the event that the dart lands within distance $a < 1$ of the center of the target, then

$$P(E) = \frac{\pi a^2}{\pi} = a^2 \quad (2.2.8)$$

For example, the probability that the dart lies within a distance $1/2$ of the center is $1/4$.

✓ Example 2.2.3

In the dart game considered above, suppose that, instead of observing where the dart lands, we observe how far it lands from the center of the target. In this case, we take as our sample space the set Ω of all circles with centers at the center of the target. It is convenient to describe these circles by their radii, so that each circle is identified by its radius r , $0 \leq r \leq 1$. In this way, we may regard Ω as the subset $[0, 1]$ of the real line.

What probabilities should we assign to the events E of Ω ?

Solution

If

$$E = \{r : 0 \leq r \leq a\}, \quad (2.2.9)$$

then E occurs if the dart lands within a distance a of the center, that is, within the circle of radius a, and we saw in the previous example that under our assumptions the probability of this event is given by

$$P([0, 1]) = a^2 \quad (2.2.10)$$

More generally, if

$$E = \{r : a \leq r \leq b\} \quad (2.2.11)$$

then by our basic assumptions,

$$\begin{aligned} P(E) = P([a, b]) &= P([0, b]) - P([0, a]) \\ &= b^2 - a^2 \\ &= (b - a)(b + a) \\ &= 2(b - a)\frac{(b+a)}{2} \end{aligned} \quad (2.2.12)$$

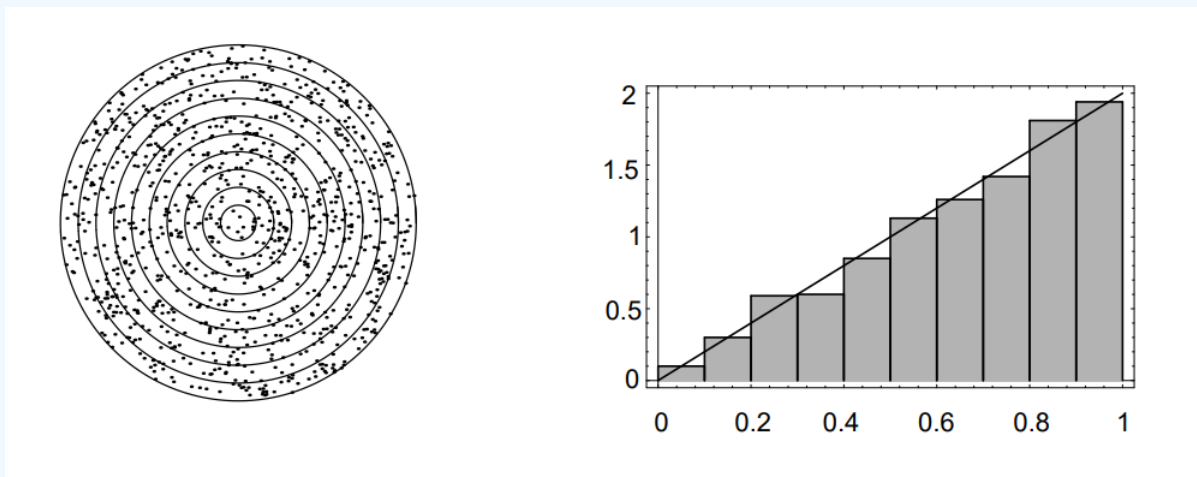


Figure 2.2.2: Distribution of dart distances in 400 throws.

Thus, $P(E) = 2(\text{length of } E)(\text{midpoint of } E)$. Here we see that the probability assigned to the interval E depends not only on its length but also on its midpoint (i.e., not only on how long it is, but also on where it is). Roughly speaking, in this experiment, events of the form $E = [a, b]$ are more likely if they are near the rim of the target and less likely if they are near the center. (A common experience for beginners! The conclusion might well be different if the beginner is replaced by an expert.)

Again we can simulate this by computer. We divide the target area into ten concentric regions of equal thickness. The computer program Darts throws n darts and records what fraction of the total falls in each of these concentric regions.

The program **Areabargraph** then plots a bar graph with the area of the i th bar equal to the fraction of the total falling in the i th region. Running the program for 1000 darts resulted in the bar graph of Figure 2.2.2.

Note that here the heights of the bars are not all equal, but grow approximately linearly with r . In fact, the linear function $y = 2r$ appears to fit our bar graph quite well. This suggests that the probability that the dart falls within a distance a of the center should be given by the area under the graph of the function $y = 2r$ between 0 and a . This area is a^2 , which agrees with the probability we have assigned above to this event.

Sample Space Coordinates

These examples suggest that for continuous experiments of this sort we should assign probabilities for the outcomes to fall in a given interval by means of the area under a suitable function.

More generally, we suppose that suitable coordinates can be introduced into the sample space Ω , so that we can regard Ω as a subset of \mathbb{R}^n . We call such a sample space a continuous sample space. We let X be a random variable which represents the outcome

of the experiment. Such a random variable is called a continuous random variable. We then define a density function for X as follows.

Density Functions of Continuous Random Variables

Definition 2.2.1

Let X be a continuous real-valued random variable. A density function for X is a real-valued function f which satisfies

$$P(a \leq X \leq b) = \int_a^b f(x)dx \quad (2.2.13)$$

for all $a, b \in \mathbb{R}$

We note that it is not the case that all continuous real-valued random variables possess density functions. However, in this book, we will only consider continuous random variables for which density functions exist. In terms of the density $f(x)$, if E is a subset of \mathbb{R} , then

$$P(X \in E) = \int_E f(x)dx \quad (2.2.14)$$

The notation here assumes that E is a subset of \mathbb{R} for which $\int_E f(x)dx$ makes sense.

Example 2.2.3

In the spinner experiment, we choose for our set of outcomes the interval $0 \leq x < 1$, and for our density function

$$f(x) = \begin{cases} 1, & \text{if } 0 \leq x < 1 \\ 0, & \text{otherwise} \end{cases} \quad (2.2.15)$$

Solution

If E is the event that the head of the spinner falls in the upper half of the circle, then $E = \{x : 0 \leq x \leq 1/2\}$, and so

$$P(E) = \int_0^{1/2} 1dx = \frac{1}{2}. \quad (2.2.16)$$

More generally, if E is the event that the head falls in the interval $[a, b]$, then

$$P(E) = \int_a^b 1dx = b - a. \quad (2.2.17)$$

Example 2.2.4

In the first dart game experiment, we choose for our sample space a disc of unit radius in the plane and for our density function the function

$$f(x, y) = \begin{cases} 1/\pi, & \text{if } x^2 + y^2 \leq 1 \\ 0, & \text{otherwise.} \end{cases} \quad (2.2.18)$$

The probability that the dart lands inside the subset E is then given by

$$\begin{aligned} P(E) &= \iint_E \frac{1}{\pi} dx dy \\ &= \frac{1}{\pi} \cdot (\text{area of } E). \end{aligned}$$

In these two examples, the density function is constant and does not depend on the particular outcome. It is often the case that experiments in which the coordinates are chosen at random can be described by constant density functions, and, as in Section 1.2, we call such density functions uniform or equiprobable. Not all experiments are of this type, however.

✓ Example 2.2.5

In the second dart game experiment, we choose for our sample space the unit interval on the real line and for our density the function

$$f(r) = \begin{cases} 2r, & \text{if } 0 < r < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (2.2.19)$$

Then the probability that the dart lands at distance r , $a \leq r \leq b$, from the center of the target is given by

$$\begin{aligned} P([a, b]) &= \int_a^b 2r dr \\ &= b^2 - a^2. \end{aligned}$$

Here again, since the density is small when r is near 0 and large when r is near 1, we see that in this experiment the dart is more likely to land near the rim of the target than near the center. In terms of the bar graph of Example ✓ 2.2.3, the heights of the bars approximate the density function, while the areas of the bars approximate the probabilities of the subintervals (see Figure ✓ 2.2.2).

We see in this example that, unlike the case of discrete sample spaces, the value $f(x)$ of the density function for the outcome x is not the probability of x occurring (we have seen that this probability is always 0) and in general $f(x)$ is not a probability at all. In this example, if we take $\lambda = 2$ then $f(3/4) = 3/2$, which being bigger than 1, cannot be a probability.

Nevertheless, the density function f does contain all the probability information about the experiment, since the probabilities of all events can be derived from it. In particular, the probability that the outcome of the experiment falls in an interval $[a, b]$ is given by

$$P([a, b]) = \int_a^b f(x) dx, \quad (2.2.20)$$

that is, by the area under the graph of the density function in the interval $[a, b]$. Thus, there is a close connection here between probabilities and areas. We have been guided by this close connection in making up our bar graphs; each bar is chosen so that its area, and not its height, represents the relative frequency of occurrence, and hence estimates the probability of the outcome falling in the associated interval.

In the language of the calculus, we can say that the probability of occurrence of an event of the form $[x, x + dx]$, where dx is small, is approximately given by

$$P([x, x + dx]) \approx f(x) dx, \quad (2.2.21)$$

that is, by the area of the rectangle under the graph of f . Note that as $dx \rightarrow 0$, this probability $\rightarrow 0$, so that the probability $P(\{x\})$ of a single point is again 0, as in Example 2.2.1.

A glance at the graph of a density function tells us immediately which events of an experiment are more likely. Roughly speaking, we can say that where the density is large the events are more likely, and where it is small the events are less likely. In Example 2.4 the density function is largest at 1. Thus, given the two intervals $[0, a]$ and $[1, 1 + a]$, where a is a small positive real number, we see that X is more likely to take on a value in the second interval than in the first.

Cumulative Distribution Functions of Continuous Random Variables

We have seen that density functions are useful when considering continuous random variables. There is another kind of function, closely related to these density functions, which is also of great importance. These functions are called cumulative distribution functions.

Definition: 2.2.2

Let X be a continuous real-valued random variable. Then the cumulative distribution function of X is defined by the equation

$$F_X(x) = P(X \leq x) \quad (2.2.22)$$

If X is a continuous real-valued random variable which possesses a density function, then it also has a cumulative distribution function, and the following theorem shows that the two functions are related in a very nice way.

Theorem 2.2.1

Let X be a continuous real-valued random variable with density function $f(x)$. Then the function defined by

$$F(x) = \int_{-\infty}^x f(t) dt \quad (2.2.23)$$

is the cumulative distribution function of X . Furthermore, we have

$$\frac{d}{dx} F(x) = f(x) \quad (2.2.24)$$

Proof. By definition,

$$F(x) = P(X \leq x) \quad (2.2.25)$$

Let $E = (-\infty, x]$. Then

$$P(X \leq x) = P(X \in E), \quad (2.2.26)$$

which equals

$$\int_{-\infty}^x f(t) dt \quad (2.2.27)$$

Applying the Fundamental Theorem of Calculus to the first equation in the statement of the theorem yields the second statement.

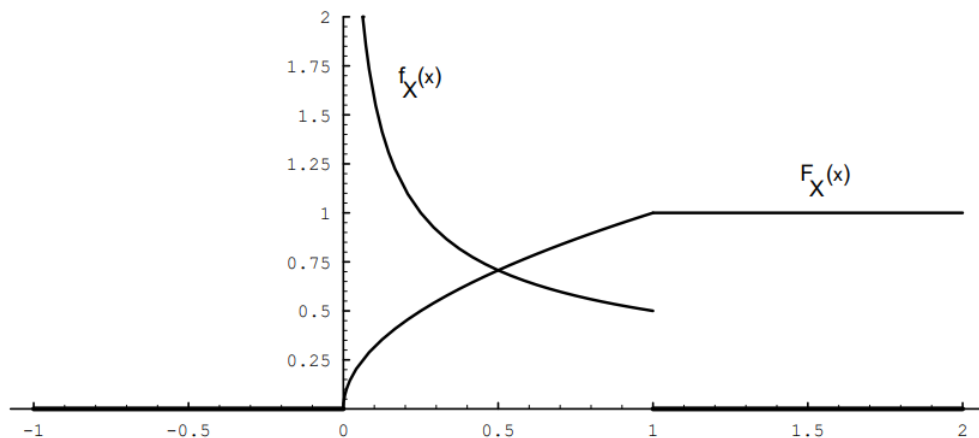


Figure 2.2.3: Distribution and density for $X = U^2$

In many experiments, the density function of the relevant random variable is easy to write down. However, it is quite often the case that the cumulative distribution function is easier to obtain than the density function. (Of course, once we have the cumulative distribution function, the density function can easily be obtained by differentiation, as the above theorem shows.) We now give some examples which exhibit this phenomenon.

✓ Example 2.2.6

A real number is chosen at random from $[0, 1]$ with uniform probability, and then this number is squared. Let X represent the result. What is the cumulative distribution function of X ? What is the density of X ?

Solution

We begin by letting U represent the chosen real number. Then $X = U^2$. If $0 \leq x \leq 1$, then we have

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P(U^2 \leq x) \\ &= P(U \leq \sqrt{x}) \\ &= \sqrt{x}. \end{aligned}$$

It is clear that X always takes on a value between 0 and 1, so the cumulative distribution function of X is given by

$$F_X(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ \sqrt{x}, & \text{if } 0 \leq x \leq 1 \\ 1, & \text{if } x \geq 1 \end{cases} \quad (2.2.28)$$

From this we easily calculate that the density function of X is

$$f_X(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ 1/(2\sqrt{x}), & \text{if } 0 \leq x \leq 1, \\ 0, & \text{if } x > 1 \end{cases} \quad (2.2.29)$$

Note that $F_X(x)$ is continuous, but $f_X(x)$ is not. (See Figure 2.13.)

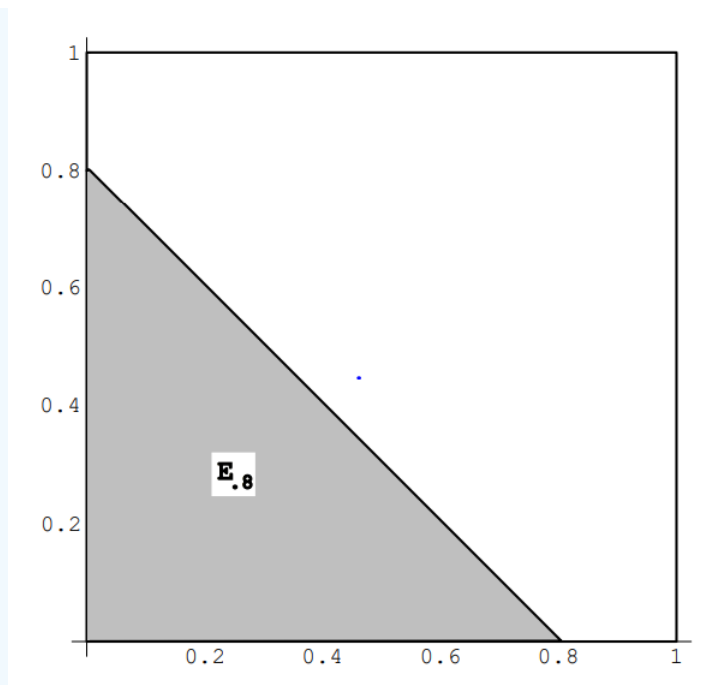


Figure 2.2.4: Calculation of distribution function for Example 2.14

When referring to a continuous random variable X (say with a uniform density function), it is customary to say that " X is uniformly distributed on the interval $[a, b]$." It is also customary to refer to the cumulative distribution function of X as the distribution function of X . Thus, the word "distribution" is being used in several different ways in the subject of probability. (Recall that it also has a meaning when discussing discrete random variables.) When referring to the cumulative distribution function of a continuous random variable X , we will always use the word "cumulative" as a modifier, unless the use of another modifier, such as "normal" or "exponential," makes it clear. Since the phrase "uniformly distributed on the interval $[a, b]$ " is not acceptable English, we will have to say "uniformly distributed" instead

✓ Example 2.2.7

In Example 2.4, we considered a random variable, defined to be the sum of two random real numbers chosen uniformly from $[0, 1]$. Let the random variables X and Y denote the two chosen real numbers. Define $Z = X + Y$. We will now derive expressions for the cumulative distribution function and the density function of Z .

Solution

Here we take for our sample space Ω the unit square in \mathbf{R}^2 with uniform density. A point $\omega \in \Omega$ then consists of a pair (x, y) of numbers chosen at random. Then $0 \leq Z \leq 2$. Let E_z denote the event that $Z \leq z$. In Figure 2.14, we show the set $E_{.8}$. The event E_z , for any z between 0 and 1, looks very similar to the shaded set in the figure. For $1 < z \leq 2$, the set E_z looks like the unit square with a triangle removed from the upper right-hand corner. We can now calculate the probability distribution F_Z of Z ; it is given by

$$\begin{aligned} F_Z(z) &= P(Z \leq z) \\ &= \text{Area of } E_z \end{aligned}$$

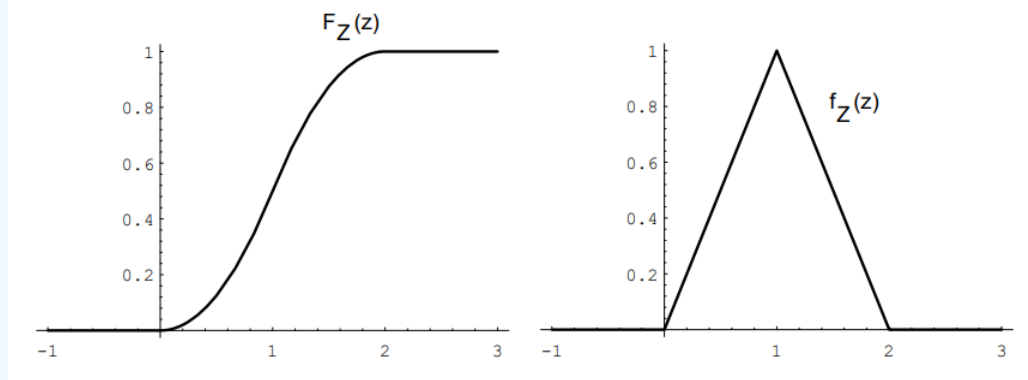


Figure 2.2.5: Distribution and density functions for Example \(\PageIndex{5}\).

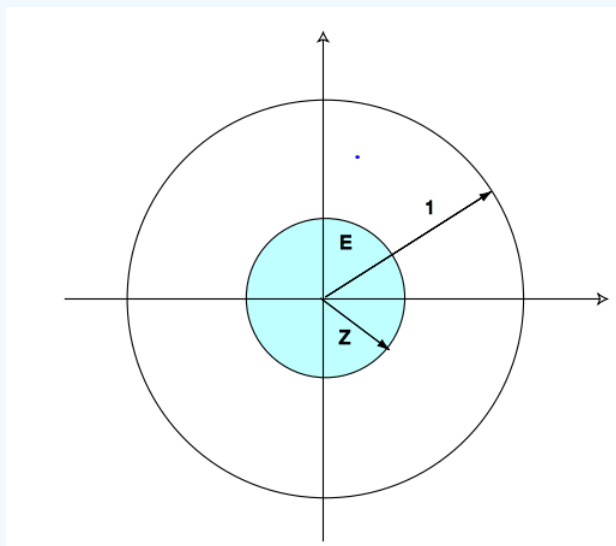


Figure 2.2.6: Calculation of F_Z for Example 2.2.6

$$= \begin{cases} 0, & \text{if } z < 0, \\ (1/2)z^2, & \text{if } 0 \leq z \leq 1, \\ 1 - (1/2)(2 - z)^2, & \text{if } 1 \leq z \leq 2, \\ 1, & \text{if } 2 < z \end{cases} \quad (2.2.30)$$

The density function is obtained by differentiating this function:

$$f_Z(z) = \begin{cases} 0, & \text{if } z < 0 \\ z, & \text{if } 0 \leq z \leq 1 \\ 2 - z, & \text{if } 1 \leq z \leq 2 \\ 0, & \text{if } 2 < z \end{cases} \quad (2.2.31)$$

The reader is referred to Figure 2.2.5 for the graphs of these functions.

✓ Example 2.2.8

In the dart game described in Example ✓ 2.2.2, what is the distribution of the distance of the dart from the center of the target? What is its density?

Solution

Here, as before, our sample space Ω is the unit disk in \mathbf{R}^2 , with coordinates (X, Y) . Let $Z = \sqrt{X^2 + Y^2}$ represent the distance from the center of the target.

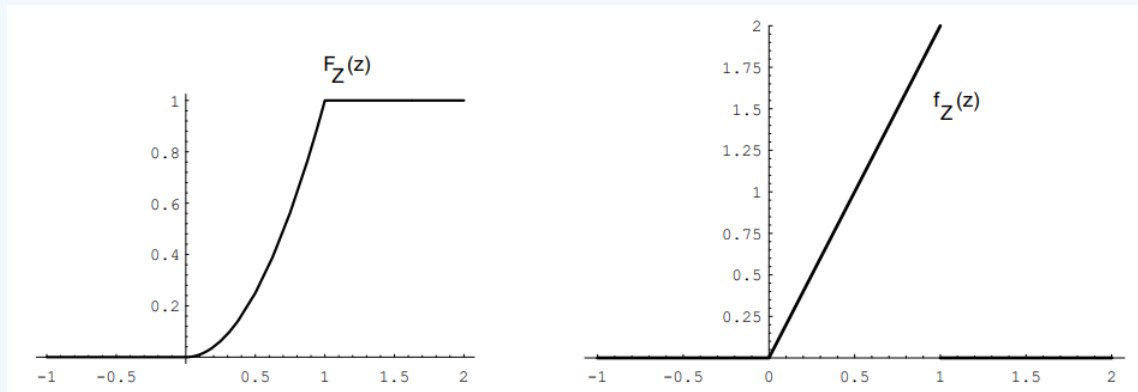


Figure 2.2.7: Distribution and density for $Z = \sqrt{X^2 + Y^2}$

Let E be the event $\{Z \leq z\}$. Then the distribution function F_Z of Z (see Figure 2.16) is given by

$$F_Z(z) = P(Z \leq z) = \frac{\text{Area of } E}{\text{Area of target}}$$

Thus, we easily compute that

$$F_Z(z) = \begin{cases} 0, & \text{if } z \leq 0 \\ z^2, & \text{if } 0 \leq z \leq 1 \\ 1, & \text{if } z > 1 \end{cases} \quad (2.2.32)$$

The density $f_Z(z)$ is given again by the derivative of $F_Z(z)$:

$$f_Z(z) = \begin{cases} 0, & \text{if } z \leq 0 \\ 2z, & \text{if } 0 \leq z \leq 1 \\ 0, & \text{if } z > 1 \end{cases} \quad (2.2.33)$$

The reader is referred to Figure 2.2.2.7 for the graphs of these functions.

We can verify this result by simulation, as follows: We choose values for X and Y at random from $[0, 1]$ with uniform distribution, calculate $Z = \sqrt{X^2 + Y^2}$, check whether $0 \leq Z \leq 1$, and present the results in a bar graph (see Figure 2.2.8).

✓ Example 2.2.9

Suppose Mr. and Mrs. Lockhorn agree to meet at the Hanover Inn between 5:00 and 6:00 P.M. on Tuesday. Suppose each arrives at a time between 5:00 and 6:00 chosen at random with uniform probability. What is the distribution function for the length of time that the first to arrive has to wait for the other? What is the density function?

Solution

Here again we can take the unit square to represent the sample space, and (X, Y) as the arrival times (after 5:00 P.M.) for the Lockhorns. Let $Z = |X - Y|$. Then we have $F_X(x) = x$ and $F_Y(y) = y$. Moreover (see Figure 2.19),

$$\begin{aligned} F_Z(z) &= P(Z \leq z) \\ &= P(|X - Y| \leq z) \\ &= \text{Area of } E. \end{aligned}$$

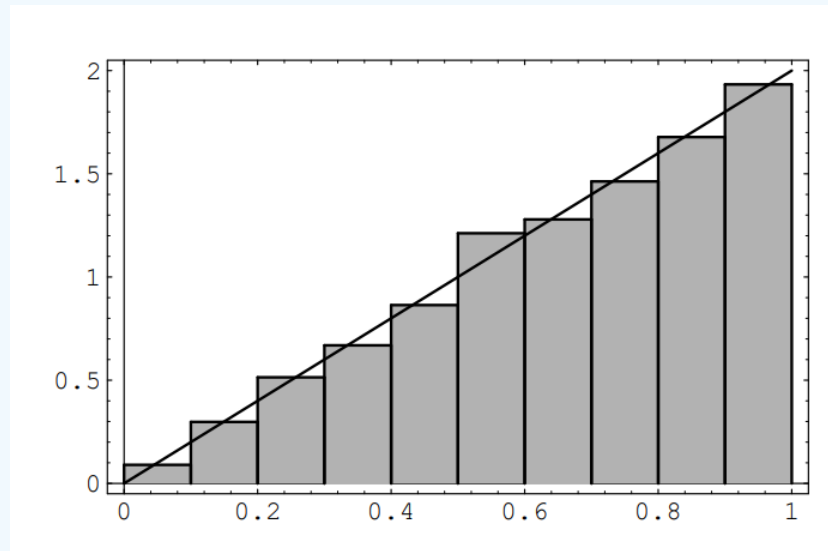


Figure 2.2.8: Simulation results for Example 2.2.7

Thus, we have

$$F_Z(z) = \begin{cases} 0, & \text{if } z \leq 0 \\ 1 - (1 - z)^2, & \text{if } 0 \leq z \leq 1, \\ 1, & \text{if } z > 1 \end{cases} \quad (2.2.34)$$

The density $f_Z(z)$ is again obtained by differentiation:

$$f_Z(z) = \begin{cases} 0, & \text{if } z \leq 0 \\ 2(1 - z), & \text{if } 0 \leq z \leq 1 \\ 0, & \text{if } z > 1 \end{cases} \quad (2.2.35)$$

✓ Example 2.2.10

There are many occasions where we observe a sequence of occurrences which occur at "random" times. For example, we might be observing emissions of a radioactive isotope, or cars passing a milepost on a highway, or light bulbs burning out. In such cases, we might define a random variable X to denote the time between successive occurrences. Clearly, X is a continuous random variable whose range consists of the non-negative real numbers. It is often the case that we can model X by using the exponential density. This density is given by the formula

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{if } t \geq 0 \\ 0, & \text{if } t < 0 \end{cases} \quad (2.2.36)$$

The number λ is a non-negative real number, and represents the reciprocal of the average value of X . (This will be shown in Chapter 6.) Thus, if the average time between occurrences is 30 minutes, then $\lambda = 1/30$. A graph of this density function with $\lambda = 1/30$ is shown in Figure 2.20. One can see from the figure that even though the average value is 30, occasionally much larger values are taken on by X .

Suppose that we have bought a computer that contains a Warp 9 hard drive. The salesperson says that the average time between breakdowns of this type of hard drive is 30 months.

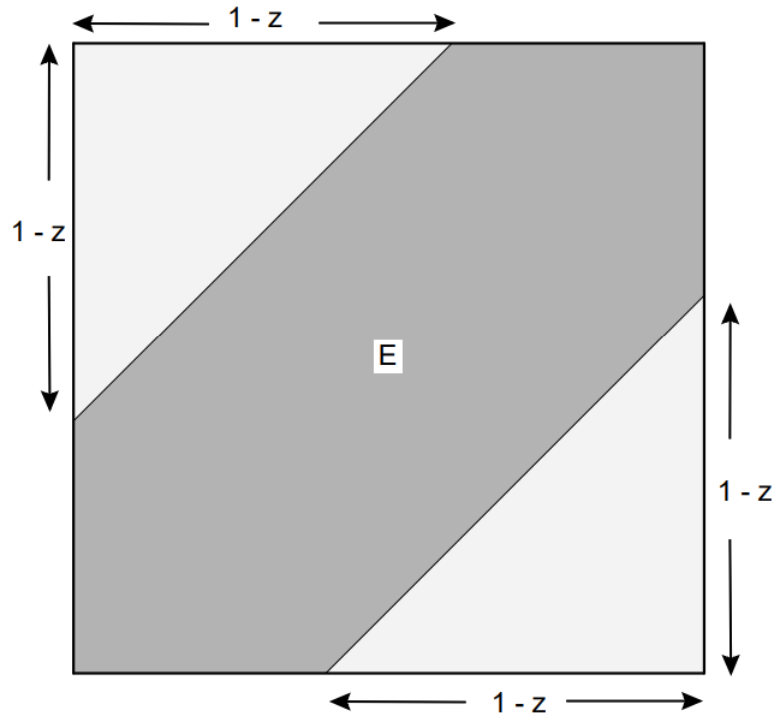


Figure 2.2.9: Calculation of FZ .

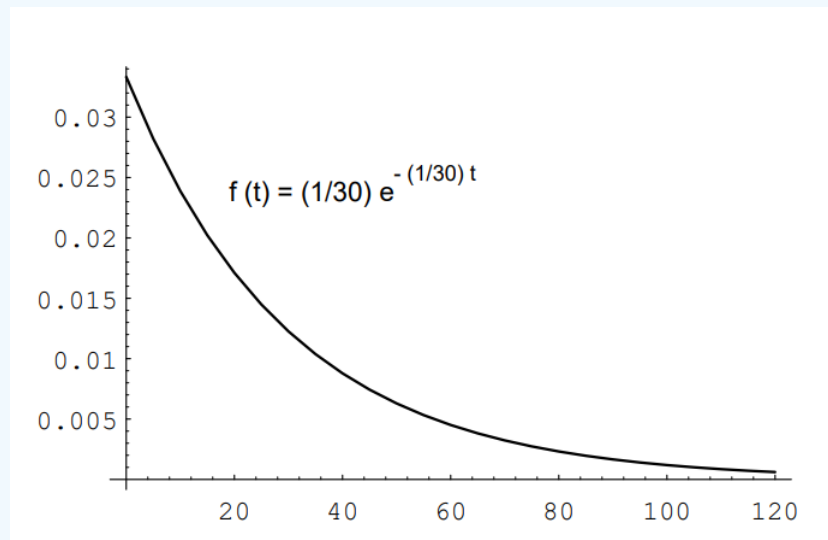


Figure 2.2.10: Exponential density with $\lambda = 1/30$

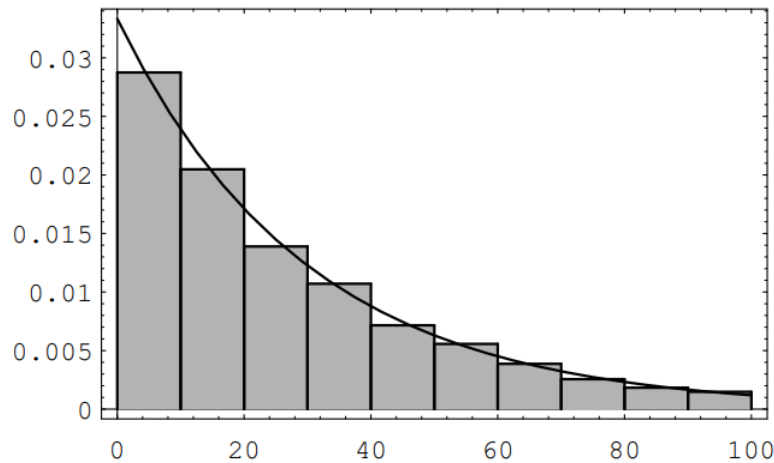


Figure 2.2.11: Residual lifespan of a hard drive

It is often assumed that the length of time between breakdowns is distributed according to the exponential density. We will assume that this model applies here, with $\lambda = 1/30$.

Now suppose that we have been operating our computer for 15 months. We assume that the original hard drive is still running. We ask how long we should expect the hard drive to continue to run. One could reasonably expect that the hard drive will run, on the average, another 15 months. (One might also guess that it will run more than 15 months, since the fact that it has already run for 15 months implies that we don't have a lemon.) The time which we have to wait is a new random variable, which we will call Y . Obviously, $Y = X - 15$. We can write a computer program to produce a sequence of simulated Y -values. To do this, we first produce a sequence of X 's, and discard those values which are less than or equal to 15 (these values correspond to the cases where the hard drive has quit running before 15 months). To simulate a value of X , we compute the value of the expression

$$\left(-\frac{1}{\lambda}\right) \log(rnd), \quad (2.2.37)$$

where rnd represents a random real number between 0 and 1. (That this expression has the exponential density will be shown in Chapter 4.3.) Figure 2.2.11 shows an area bar graph of 10,000 simulated Y -values.

The average value of Y in this simulation is 29.74, which is closer to the original average life span of 30 months than to the value of 15 months which was guessed above. Also, the distribution of Y is seen to be close to the distribution of X . It is in fact the case that X and Y have the same distribution. This property is called the memoryless property, because the amount of time that we have to wait for an occurrence does not depend on how long we have already waited. The only continuous density function with this property is the exponential density.

Assignment of Probabilities

A fundamental question in practice is: How shall we choose the probability density function in describing any given experiment? The answer depends to a great extent on the amount and kind of information available to us about the experiment. In some cases, we can see that the outcomes are equally likely. In some cases, we can see that the experiment resembles another already described by a known density. In some cases, we can run the experiment a large number of times and make a reasonable guess at the density on the basis of the observed distribution of outcomes, as we did in Chapter 1. In general, the problem of choosing the right density function for a given experiment is a central problem for the experimenter and is not always easy to solve (see Example 2.6). We shall not examine this question in detail here but instead shall assume that the right density is already known for each of the experiments under study.

The introduction of suitable coordinates to describe a continuous sample space, and a suitable density to describe its probabilities, is not always so obvious, as our final example shows.

✓ Example 2.2.11

Consider an experiment in which a fair coin is tossed repeatedly, without stopping. We have seen in Example 1.6 that, for a coin tossed n times, the natural sample space is a binary tree with n stages. On this evidence we expect that for a coin tossed repeatedly, the natural sample space is a binary tree with an infinite number of stages, as indicated in Figure 2.2.12

It is surprising to learn that, although the n -stage tree is obviously a finite sample space, the unlimited tree can be described as a continuous sample space. To see how this comes about, let us agree that a typical outcome of the unlimited coin tossing experiment can be described by a sequence of the form $\omega = \{HHTHTTH \dots\}$. If we write 1 for H and 0 for T, then $\omega = \{1 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \dots\}$. In this way, each outcome is described by a sequence of 0's and 1's.

Now suppose we think of this sequence of 0's and 1's as the binary expansion of some real number $x = .1101001\dots$ lying between 0 and 1. (A binary expansion is like a decimal expansion but based on 2 instead of 10.) Then each outcome is described by a value of x , and in this way x becomes a coordinate for the sample space, taking on all real values between 0 and 1. (We note that it is possible for two different sequences to correspond to the same real number; for example, the sequences $\{THHHHH\dots\}$ and $\{HTTTTTT\dots\}$ both correspond to the real number $1/2$. We will not concern ourselves with this apparent problem here.)

What probabilities should be assigned to the events of this sample space? Consider, for example, the event E consisting of all outcomes for which the first toss comes up heads and the second tails. Every such outcome has the form $.10*****$, where $*$ can be either 0 or 1.

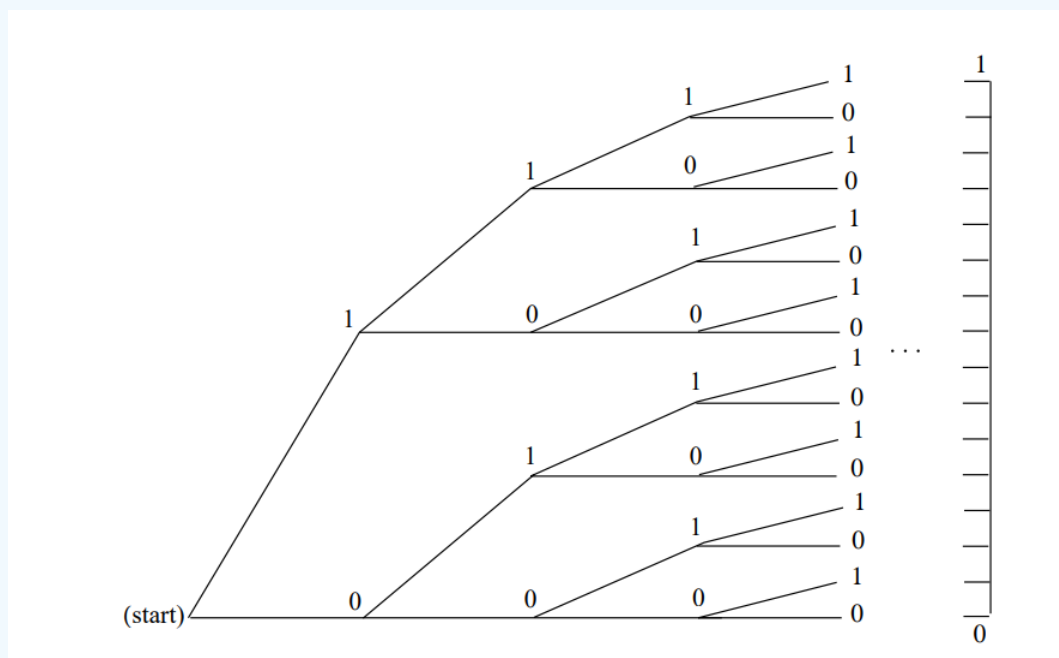


Figure 2.2.12: Tree for infinite number of tosses of a coin.

Now if x is our real-valued coordinate, then the value of x for every such outcome must lie between $1/2 = .10000\dots$ and $3/4 = .11000\dots$; and moreover, every value of x between $1/2$ and $3/4$ has a binary expansion of the form $.10*****$. This means that $\omega \in E$ if and only if $1/2 \leq x < 3/4$, and in this way we see that we can describe E by the interval $[1/2, 3/4)$. More generally, every event consisting of outcomes for which the results of the first n tosses are prescribed is described by a binary interval of the form $[k/2^n, (k+1)/2^n)$.

We have already seen in Section 1.2 that in the experiment involving n tosses, the probability of any one outcome must be exactly $1/2^n$. It follows that in the unlimited toss experiment, the probability of any event consisting of outcomes for which the results of the first n tosses are prescribed must also be $1/2^n$. But $1/2^n$ is exactly the length of the interval of x -values describing E ! Thus we see that, just as with the spinner experiment, the probability of an event E is determined by what fraction of the unit interval lies in E .

Consider again the statement: The probability is $1/2$ that a fair coin will turn up heads when tossed. We have suggested that one interpretation of this statement is that if we toss the coin indefinitely the proportion of heads will approach $1/2$. That is, in our correspondence with binary sequences we expect to get a binary sequence with the proportion of 1's tending to $1/2$. The event E of binary sequences for which this is true is a proper subset of the set of all possible binary sequences. It does not contain, for example, the sequence 011011011 ... (i.e., (011) repeated again and again). The event E is actually a very complicated subset of the binary sequences, but its probability can be determined as a limit of probabilities for events with a finite number of outcomes whose probabilities are given by finite tree measures. When the probability of E is computed in this way, its value is found to be 1. This remarkable result is known as the Strong Law of Large Numbers (or Law of Averages) and is one justification for our frequency concept of probability. We shall prove a weak form of this theorem in Chapter 8.

Exercises

Exercise 2.2.1:

Suppose you choose at random a real number X from the interval $[2, 10]$.

- Find the density function $f(x)$ and the probability of an event E for this experiment, where E is a subinterval $[a, b]$ of $[2, 10]$.
- From (a), find the probability that $X > 5$, that $5 < X < 7$, and that $X^2 - 12X + 35 > 0$.

Exercise 2.2.2:

Suppose you choose a real number X from the interval $[2, 10]$ with a density function of the form

$$f(x) = Cx, \quad (2.2.38)$$

where C is a constant.

- Find C .
- Find $P(E)$, where $E = [a, b]$ is a subinterval of $[2, 10]$.
- Find $P(X > 5)$, $P(X < 7)$, and $P(X^2 - 12X + 35 > 0)$.

Exercise 2.2.3:

Same as Exercise 2.2.2, but suppose

$$f(x) = \frac{C}{x}. \quad (2.2.39)$$

Exercise 2.2.4:

Suppose you throw a dart at a circular target of radius 10 inches. Assuming that you hit the target and that the coordinates of the outcomes are chosen at random, find the probability that the dart falls

- within 2 inches of the center.
- within 2 inches of the rim.
- within the first quadrant of the target.
- within the first quadrant and within 2 inches of the rim.

Exercise 2.2.5:

Suppose you are watching a radioactive source that emits particles at a rate described by the exponential density

$$f(t) = \lambda e^{-\lambda t}, \quad (2.2.40)$$

where $\lambda = 1$, so that the probability $P(0, T)$ that a particle will appear in the next T seconds is $P([0, T]) = \int_0^T \lambda e^{-\lambda t} dt$. Find the probability that a particle (not necessarily the first) will appear

- within the next second.
- within the next 3 seconds.
- between 3 and 4 seconds from now.
- after 4 seconds from now.

Exercise 2.2.6:

Assume that a new light bulb will burn out after t hours, where t is chosen from $[0, \infty)$ with an exponential density

$$f(t) = \lambda e^{-\lambda t}. \quad (2.2.41)$$

In this context, λ is often called the failure rate of the bulb.

- (a) Assume that $\lambda = 0.01$, and find the probability that the bulb will not burn out before T hours. This probability is often called the reliability of the bulb.
- (b) For what T is the reliability of the bulb $= 1/2$?

Exercise 2.2.7:

Choose a number B at random from the interval $[0,1]$ with uniform density. Find the probability that

- (a) $1/3 < B < 2/3$.
- (b) $|B - 1/2| \leq 1/4$.
- (c) $B < 1/4$ or $1 - B < 1/4$.
- (d) $3B^2 < B$.

Exercise 2.2.8:

Choose independently two numbers B and C at random from the interval $[0, 1]$ with uniform density. Note that the point (B, C) is then chosen at random in the unit square. Find the probability that

- (a) $B + C < 1/2$.
- (b) $BC < 1/2$.
- (c) $|B - C| < 1/2$.
- (d) $\max\{B, C\} < 1/2$.
- (e) $\min\{B, C\} < 1/2$.
- (f) $B < 1/2$ and $1 - C < 1/2$.
- (g) conditions (c) and (f) both hold.
- (h) $B^2 + C^2 \leq 1/2$.
- (i) $(B - 1/2)^2 + (C - 1/2)^2 < 1/4$.

Exercise 2.2.9:

Suppose that we have a sequence of occurrences. We assume that the time X between occurrences is exponentially distributed with $\lambda = 1/10$, so on the average, there is one occurrence every 10 minutes (see Example 2.17). You come upon this system at time 100, and wait until the next occurrence. Make a conjecture concerning how long, on the average, you will have to wait. Write a program to see if your conjecture is right.

Exercise 2.2.10:

As in Exercise 9, assume that we have a sequence of occurrences, but now assume that the time X between occurrences is uniformly distributed between 5 and 15. As before, you come upon this system at time 100, and wait until the next occurrence. Make a conjecture concerning how long, on the average, you will have to wait. Write a program to see if your conjecture is right.

Exercise 2.2.11:

For examples such as those in Exercises 9 and 10, it might seem that at least you should not have to wait on average more than 10 minutes if the average time between occurrences is 10 minutes. Alas, even this is not true. To see why, consider the following assumption about the times between occurrences. Assume that the time between occurrences is 3 minutes with probability .9 and 73 minutes with probability .1. Show by simulation that the average time between occurrences is 10 minutes, but that if you come upon this system at time 100, your average waiting time is more than 10 minutes.

Exercise 2.2.12:

Take a stick of unit length and break it into three pieces, choosing the break points at random. (The break points are assumed to be chosen simultaneously.) What is the probability that the three pieces can be used to form a triangle? Hint: The sum of the lengths of any two pieces must exceed the length of the third, so each piece must have length $< 1/2$. Now use Exercise 8(g).

Exercise 2.2.13:

Take a stick of unit length and break it into two pieces, choosing the break point at random. Now break the longer of the two pieces at a random point. What is the probability that the three pieces can be used to form a triangle?

Exercise 2.2.14:

Choose independently two numbers B and C at random from the interval $[-1, 1]$ with uniform distribution, and consider the quadratic equation

$$x^2 + Bx + C = 0. \quad (2.2.42)$$

Find the probability that the roots of this equation

(a) are both real.

(b) are both positive.

Hints: (a) requires $0 \leq B^2 - 4C$, (b) requires $0 \leq B^2 - 4C$, $B \leq 0$, $0 \leq C$.

Exercise 2.2.15:

At the Tunbridge World's Fair, a coin toss game works as follows. Quarters are tossed onto a checkerboard. The management keeps all the quarters, but for each quarter landing entirely within one square of the checkerboard the management pays a dollar. Assume that the edge of each square is twice the diameter of a quarter, and that the outcomes are described by coordinates chosen at random. Is this a fair game?

Exercise 2.2.16:

Three points are chosen at random on a circle of unit circumference. What is the probability that the triangle defined by these points as vertices has three acute angles? Hint: One of the angles is obtuse if and only if all three points lie in the same semicircle. Take the circumference as the interval $[0, 1]$. Take one point at 0 and the others at B and C .

Exercise 2.2.17:

Write a program to choose a random number X in the interval $[2, 10]$ 1000 times and record what fraction of the outcomes satisfy $X > 5$, what fraction satisfy $5 < X < 7$, and what fraction satisfy $x^2 - 12x + 35 > 0$. How do these results compare with Exercise 1?

Exercise 2.2.18:

Write a program to choose a point (X, Y) at random in a square of side 20 inches, doing this 10,000 times, and recording what fraction of the outcomes fall within 19 inches of the center; of these, what fraction fall between 8 and 10 inches of the center; and, of these, what fraction fall within the first quadrant of the square. How do these results compare with those of Exercise 4?

Exercise 2.2.19:

Write a program to simulate the problem described in Exercise 7 (see Exercise 17). How do the simulation results compare with the results of Exercise 7?

Exercise 2.2.20:

Write a program to simulate the problem described in Exercise 12.

Exercise 2.2.21:

Write a program to simulate the problem described in Exercise 16.

Exercise 2.2.22:

Write a program to carry out the following experiment. A coin is tossed 100 times and the number of heads that turn up is recorded. This experiment is then repeated 1000 times. Have your program plot a bar graph for the proportion of the 1000 experiments in which the number of heads is n , for each n in the interval $[35, 65]$. Does the bar graph look as though it can be fit with a normal curve?

Exercise 2.2.23:

Write a program that picks a random number between 0 and 1 and computes the negative of its logarithm. Repeat this process a large number of times and plot a bar graph to give the number of times that the outcome falls in each interval of length 0.1 in $[0, 10]$. On this bar graph plot a graph of the density $f(x) = e^{-x}$. How well does this density fit your graph?

This page titled [2.2: Continuous Density Functions](#) is shared under a [GNU Free Documentation License 1.3](#) license and was authored, remixed, and/or curated by [Charles M. Grinstead & J. Laurie Snell](#) ([American Mathematical Society](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.