

## 6.3: Continuous Random Variables

In this section we consider the properties of the expected value and the variance of a continuous random variable. These quantities are defined just as for discrete random variables and share the same properties.

### Expected Value

#### Definition: Term

Let  $X$  be a real-valued random variable with density function  $f(x)$ . The expected value  $\mu = E(X)$  is defined by

$$\mu = E(X) = \int_{-\infty}^{+\infty} x f(x) dx, \quad (6.3.1)$$

provided the integral

$$\int_{-\infty}^{+\infty} |x| f(x) dx \quad (6.3.2)$$

is finite.

The reader should compare this definition with the corresponding one for discrete random variables in Section 6.1. Intuitively, we can interpret  $E(X)$ , as we did in the previous sections, as the value that we should expect to obtain if we perform a large number of independent experiments and average the resulting values of  $X$ .

We can summarize the properties of  $E(X)$  as follows (cf. Theorem 6.2).

#### Theorem 6.3.1

If  $X$  and  $Y$  are real-valued random variables and  $c$  is any constant, then

$$\begin{aligned} E(X + Y) &= E(X) + E(Y) \\ E(cX) &= cE(X). \end{aligned}$$

The proof is very similar to the proof of Theorem 6.2, and we omit it.

More generally, if  $X_1, X_2, \dots, X_n$  are  $n$  real-valued random variables, and  $c_1, c_2, \dots, c_n$  are  $n$  constants, then

$$E(c_1 X_1 + c_2 X_2 + \dots + c_n X_n) = c_1 E(X_1) + c_2 E(X_2) + \dots + c_n E(X_n). \quad (6.3.3)$$

#### Example 6.3.1

Let  $X$  be uniformly distributed on the interval  $[0, 1]$ . Then

$$E(X) = \int_0^1 x dx = 1/2 \quad (6.3.4)$$

It follows that if we choose a large number  $N$  of random numbers from  $[0, 1]$  and take the average, then we can expect that this average should be close to the expected value of  $1/2$ .

### ✓ Example 6.3.2

Let  $Z = (x, y)$  denote a point chosen uniformly and randomly from the unit disk, as in the dart game in Example 2.8 and let  $X = (x^2 + y^2)^{1/2}$  be the distance from  $Z$  to the center of the disk. The density function of  $X$  can easily be shown to equal  $f(x) = 2x$ , so by the definition of expected value,

$$\begin{aligned} E(X) &= \int_0^1 x f(x) dx \\ &= \int_0^1 x(2x) dx \\ &= \frac{2}{3}. \end{aligned}$$

### ✓ Example 6.3.3

In the example of the couple meeting at the Inn (Example 2.16), each person arrives at a time which is uniformly distributed between 5:00 and 6:00 PM. The random variable  $Z$  under consideration is the length of time the first person has to wait until the second one arrives. It was shown that

$$f_Z(z) = 2(1 - z) \quad (6.3.5)$$

for  $0 \leq z \leq 1$ . Hence,

$$E(Z) = \int_0^1 z f_Z(z) dz \quad (6.3.6)$$

$$\begin{aligned} &= \int_0^1 2z(1 - z) dz \\ &= \left[ z^2 - \frac{2}{3} z^3 \right]_0^1 \\ &= \frac{1}{3}. \end{aligned} \quad (6.3.7)$$

## Expectation of a Function of a Random Variable

Suppose that  $X$  is a real-valued random variable and  $\phi(x)$  is a continuous function from  $\mathbf{R}$  to  $\mathbf{R}$ . The following theorem is the continuous analogue of Theorem 6.1.

### Theorem 6.3.2

If  $X$  is a real-valued random variable and if  $\phi : \mathbf{R} \rightarrow \mathbf{R}$  is a continuous real-valued function with domain  $[a, b]$ , then

$$E(\phi(X)) = \int_{-\infty}^{+\infty} \phi(x) f_X(x) dx \quad (6.3.8)$$

provided the integral exists.

For a proof of this theorem, see Ross.<sup>1</sup>

## Expectation of the Product of Two Random Variables

In general, it is not true that  $E(XY) = E(X)E(Y)$ , since the integral of a product is not the product of integrals. But if  $X$  and  $Y$  are independent, then the expectations multiply.

### Theorem 6.3.3

Let  $X$  and  $Y$  be independent real-valued continuous random variables with finite expected values. Then we have

$$E(XY) = E(X)E(Y). \quad (6.3.9)$$

**Proof.** We will prove this only in the case that the ranges of  $X$  and  $Y$  are contained in the intervals  $[a, b]$  and  $[c, d]$ , respectively. Let the density functions of  $X$  and  $Y$  be denoted by  $f_X(x)$  and  $f_Y(y)$ , respectively. Since  $X$  and  $Y$  are independent, the joint density function of  $X$  and  $Y$  is the product of the individual density functions. Hence

$$\begin{aligned} E(XY) &= \int_a^b \int_c^d xy f_X(x) f_Y(y) dy dx \\ &= \int_a^b x f_X(x) dx \int_c^d y f_Y(y) dy \\ &= E(X)E(Y). \end{aligned}$$

The proof in the general case involves using sequences of bounded random variables that approach  $X$  and  $Y$ , and is somewhat technical, so we will omit it.

In the same way, one can show that if  $X_1, X_2, \dots, X_n$  are  $n$  mutually independent real-valued random variables, then

$$E(X_1 X_2 \cdots X_n) = E(X_1) E(X_2) \cdots E(X_n). \quad (6.3.10)$$

### ✓ Example 6.3.4

Let  $Z = (X, Y)$  be a point chosen at random in the unit square. Let  $A = X^2$  and  $B = Y^2$ . Then Theorem 4.3 implies that  $A$  and  $B$  are independent. Using Theorem ✓ 6.3.2, the expectations of  $A$  and  $B$  are easy to calculate:

$$\begin{aligned} E(A) &= E(B) = \int_0^1 x^2 dx \\ &= \frac{1}{3}. \end{aligned}$$

Using Theorem ✓ 6.3.3, the expectation of  $AB$  is just the product of  $E(A)$  and  $E(B)$ , or  $1/9$ . The usefulness of this theorem is demonstrated by noting that it is quite a bit more difficult to calculate  $E(AB)$  from the definition of expectation. One finds that the density function of  $AB$  is

$$f_{AB}(t) = \frac{-\log(t)}{4\sqrt{t}}, \quad (6.3.11)$$

so

$$\begin{aligned} E(AB) &= \int_0^1 t f_{AB}(t) dt \\ &= \frac{1}{9}. \end{aligned}$$

### ✓ Example 6.3.4

Again let  $Z = (X, Y)$  be a point chosen at random in the unit square, and let  $W = X + Y$ . Then  $Y$  and  $W$  are not independent, and we have

$$\begin{aligned}
 E(Y) &= \frac{1}{2}, \\
 E(W) &= 1, \\
 E(YW) &= E(XY + Y^2) = E(X)E(Y) + \frac{1}{3} = \frac{7}{12} \neq E(Y)E(W).
 \end{aligned}$$

We turn now to the variance.

## Variance

### Definition: Variance

Let  $X$  be a real-valued random variable with density function  $f(x)$ . The variance  $\sigma^2 = V(X)$  is defined by

$$\sigma^2 = V(X) = E((X - \mu)^2). \quad (6.3.12)$$

The next result follows easily from Theorem 6.1.1. There is another way to calculate the variance of a continuous random variable, which is usually slightly easier. It is given in Theorem 6.3.6.

### Theorem 6.3.4

If  $X$  is a real-valued random variable with  $E(X) = \mu$ , then

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx. \quad (6.3.13)$$

The properties listed in the next three theorems are all proved in exactly the same way that the corresponding theorems for discrete random variables were proved in Section 6.2.

### Theorem 6.3.5

If  $X$  is a real-valued random variable defined on  $\Omega$  and  $c$  is any constant, then (cf. Theorem 6.2.7)

$$\begin{aligned}
 V(cX) &= c^2 V(X), \\
 V(X + c) &= V(X).
 \end{aligned}$$

### Theorem 6.3.6

If  $X$  is a real-valued random variable with  $E(X) = \mu$ , then (cf. Theorem 6.2.6)

$$V(X) = E(X^2) - \mu^2. \quad (6.3.14)$$

### Theorem 6.3.7

If  $X$  and  $Y$  are independent real-valued random variables on  $\Omega$ , then (cf. Theorem 6.2.8)

$$V(X + Y) = V(X) + V(Y). \quad (6.3.15)$$

### ✓ Example 6.3.5

If  $X$  is uniformly distributed on  $[0, 1]$ , then, using Theorem ✓ 6.3.6, we have

$$V(X) = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = \frac{1}{12}. \quad (6.3.16)$$

### ✓ Example 6.3.6

Let  $X$  be an exponentially distributed random variable with parameter  $\lambda$ . Then the density function of  $X$  is

$$f_X(x) = \lambda e^{-\lambda x}. \quad (6.3.17)$$

From the definition of expectation and integration by parts, we have

$$\begin{aligned} E(X) &= \int_0^\infty x f_X(x) dx \\ &= \lambda \int_0^\infty x e^{-\lambda x} dx \\ &= -x e^{-\lambda x} \Big|_0^\infty + \int_0^\infty e^{-\lambda x} dx \\ &= 0 + \frac{e^{-\lambda x}}{-\lambda} \Big|_0^\infty = \frac{1}{\lambda}. \end{aligned}$$

Similarly, using Theorems 6.3.1 and 6.3.6, we have

$$\begin{aligned} V(X) &= \int_0^\infty x^2 f_X(x) dx - \frac{1}{\lambda^2} \\ &= \lambda \int_0^\infty x^2 e^{-\lambda x} dx - \frac{1}{\lambda^2} \\ &= -x^2 e^{-\lambda x} \Big|_0^\infty + 2 \int_0^\infty x e^{-\lambda x} dx - \frac{1}{\lambda^2} \\ &= -x^2 e^{-\lambda x} \Big|_0^\infty - \frac{2x e^{-\lambda x}}{\lambda} \Big|_0^\infty - \frac{2}{\lambda^2} e^{-\lambda x} \Big|_0^\infty - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}. \end{aligned}$$

In this case, both  $E(X)$  and  $V(X)$  are finite if  $\lambda > 0$ .

### ✓ Example 6.3.7

Let  $Z$  be a standard normal random variable with density function

$$f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad (6.3.18)$$

Since this density function is symmetric with respect to the  $y$ -axis, then it is easy to show that

$$\int_{-\infty}^\infty x f_Z(x) dx \quad (6.3.19)$$

has value 0. The reader should recall however, that the expectation is defined to be the above integral only if the integral

$$\int_{-\infty}^{\infty} |x| f_Z(x) dx \quad (6.3.20)$$

is finite. This integral equals

$$2 \int_0^{\infty} x f_Z(x) dx \quad (6.3.21)$$

which one can easily show is finite. Thus, the expected value of  $Z$  is 0 .

To calculate the variance of  $Z$ , we begin by applying Theorem 6.3.6:

$$V(Z) = \int_{-\infty}^{+\infty} x^2 f_Z(x) dx - \mu^2 \quad (6.3.22)$$

If we write  $x^2$  as  $x \cdot x$ , and integrate by parts, we obtain

$$\frac{1}{\sqrt{2\pi}} \left( -x e^{-x^2/2} \right) \Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2} dx. \quad (6.3.23)$$

The first summand above can be shown to equal 0 , since as  $x \rightarrow \pm\infty$ ,  $e^{-x^2/2}$  gets small more quickly than  $x$  gets large. The second summand is just the standard normal density integrated over its domain, so the value of this summand is 1 . Therefore, the variance of the standard normal density equals 1 .

Now let  $X$  be a (not necessarily standard) normal random variable with parameters  $\mu$  and  $\sigma$ . Then the density function of  $X$  is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}. \quad (6.3.24)$$

We can write  $X = \sigma Z + \mu$ , where  $Z$  is a standard normal random variable. Since  $E(Z) = 0$  and  $V(Z) = 1$  by the calculation above, Theorems 6.10 and 6.14 imply that

$$\begin{aligned} E(X) &= E(\sigma Z + \mu) = \mu, \\ V(X) &= V(\sigma Z + \mu) = \sigma^2. \end{aligned}$$

#### ✓ Example 6.3.8

Let  $X$  be a continuous random variable with the Cauchy density function

$$f_X(x) = \frac{a}{\pi} \frac{1}{a^2 + x^2}. \quad (6.3.25)$$

Then the expectation of  $X$  does not exist, because the integral

$$\frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{|x| dx}{a^2 + x^2} \quad (6.3.26)$$

diverges. Thus the variance of  $X$  also fails to exist. Densities whose variance is not defined, like the Cauchy density, behave quite differently in a number of important respects from those whose variance is finite. We shall see one instance of this difference in Section 8.2.

## Independent Trials

### Corollary 6.3.1

If  $X_1, X_2, \dots, X_n$  is an independent trials process of real-valued random variables, with  $E(X_i) = \mu$  and  $V(X_i) = \sigma^2$ , and if

$$\begin{aligned} S_n &= X_1 + X_2 + \dots + X_n, \\ A_n &= \frac{S_n}{n}, \end{aligned} \quad (6.3.27)$$

then

$$\begin{aligned} E(S_n) &= n\mu, \\ E(A_n) &= \mu, \\ V(S_n) &= n\sigma^2, \\ V(A_n) &= \frac{\sigma^2}{n}. \end{aligned}$$

It follows that if we set

$$S_n^* = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}, \quad (6.3.28)$$

then

$$\begin{aligned} E(S_n^*) &= 0, \\ V(S_n^*) &= 1. \end{aligned} \quad (6.3.29)$$

We say that  $S_n^*$  is a standardized version of  $S_n$  (see Exercise 12 in Section 6.2).

## Queues

### Example 6.3.1

Let us consider again the queueing problem, that is, the problem of the customers waiting in a queue for service (see Example 5.7). We suppose again that customers join the queue in such a way that the time between arrivals is an exponentially distributed random variable  $X$  with density function

$$f_X(t) = \lambda e^{-\lambda t}. \quad (6.3.30)$$

Then the expected value of the time between arrivals is simply  $1/\lambda$  (see Example 6.26), as was stated in Example 5.7. The reciprocal  $\lambda$  of this expected value is often referred to as the arrival rate. The service time of an individual who is first in line is defined to be the amount of time that the person stays at the head of the line before leaving. We suppose that the customers are served in such a way that the service time is another exponentially distributed random variable  $Y$  with density function

$$f_Y(t) = \mu e^{-\mu t}. \quad (6.3.31)$$

Then the expected value of the service time is

$$E(X) = \int_0^{\infty} t f_X(t) dt = \frac{1}{\mu}. \quad (6.3.32)$$

The reciprocal  $\mu$  if this expected value is often referred to as the service rate

We expect on grounds of our everyday experience with queues that if the service rate is greater than the arrival rate, then the average queue size will tend to stabilize, but if the service rate is less than the arrival rate, then the queue will tend to increase in length without limit (see Figure 5.7). The simulations in Example 5.7 tend to bear out our everyday experience. We can make this conclusion more precise if we introduce the traffic intensity as the product

$$\rho = (\text{arrival rate})(\text{average service time}) = \frac{\lambda}{\mu} = \frac{1/\mu}{1/\lambda}. \quad (6.3.33)$$

The traffic intensity is also the ratio of the average service time to the average time between arrivals. If the traffic intensity is less than 1 the queue will perform reasonably, but if it is greater than 1 the queue will grow indefinitely large. In the critical case of  $\rho = 1$ , it can be shown that the queue will become large but there will always be times at which the queue is empty.<sup>22</sup>

In the case that the traffic intensity is less than 1 we can consider the length of the queue as a random variable  $Z$  whose expected value is finite,

$$E(Z) = N. \quad (6.3.34)$$

The time spent in the queue by a single customer can be considered as a random variable  $W$  whose expected value is finite,

$$E(W) = T. \quad (6.3.35)$$

Then we can argue that, when a customer joins the queue, he expects to find  $N$  people ahead of him, and when he leaves the queue, he expects to find  $\lambda T$  people behind him. Since, in equilibrium, these should be the same, we would expect to find that

$$N = \lambda T. \quad (6.3.36)$$

This last relationship is called Little's law for queues.<sup>23</sup> We will not prove it here. A proof may be found in Ross.<sup>24</sup> Note that in this case we are counting the waiting time of all customers, even those that do not have to wait at all. In our simulation in Section 4.2, we did not consider these customers.

If we knew the expected queue length then we could use Little's law to obtain the expected waiting time, since

$$T = \frac{N}{\lambda}. \quad (6.3.37)$$

The queue length is a random variable with a discrete distribution. We can estimate this distribution by simulation, keeping track of the queue lengths at the times at which a customer arrives. We show the result of this simulation (using the program Queue) in Figure 6.3.1.



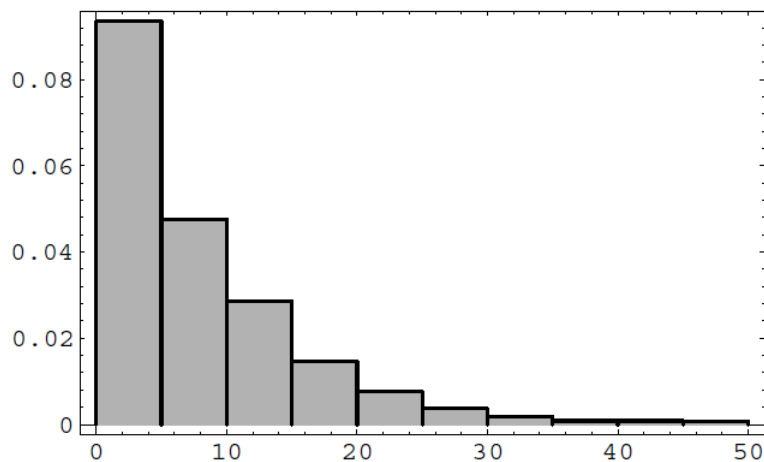


Figure 6.3.1: Distribution of queue lengths.

We note that the distribution appears to be a geometric distribution. In the study of queueing theory it is shown that the distribution for the queue length in equilibrium is indeed a geometric distribution with

$$s_j = (1 - \rho)\rho^j \quad \text{for } j = 0, 1, 2, \dots, \quad (6.3.38)$$

if  $\rho < 1$ . The expected value of a random variable with this distribution is

$$N = \frac{\rho}{(1 - \rho)} \quad (6.3.39)$$

(see Example 6.4). Thus by Little's result the expected waiting time is

$$T = \frac{\rho}{\lambda(1 - \rho)} = \frac{1}{\mu - \lambda}, \quad (6.3.40)$$

where  $\mu$  is the service rate,  $\lambda$  the arrival rate, and  $\rho$  the traffic intensity.

In our simulation, the arrival rate is 1 and the service rate is 1.1. Thus, the traffic intensity is  $1/1.1 = 10/11$ , the expected queue size is

$$\frac{10/11}{(1 - 10/11)} = 10 \quad (6.3.41)$$

and the expected waiting time is

$$\frac{1}{1.1 - 1} = 10 \quad (6.3.42)$$

In our simulation the average queue size was 8.19 and the average waiting time was 7.37. In Figure 6.3.2, we show the histogram for the waiting times. This histogram suggests that the density for the waiting times is exponential with parameter  $\mu - \lambda$ , and this is the case.

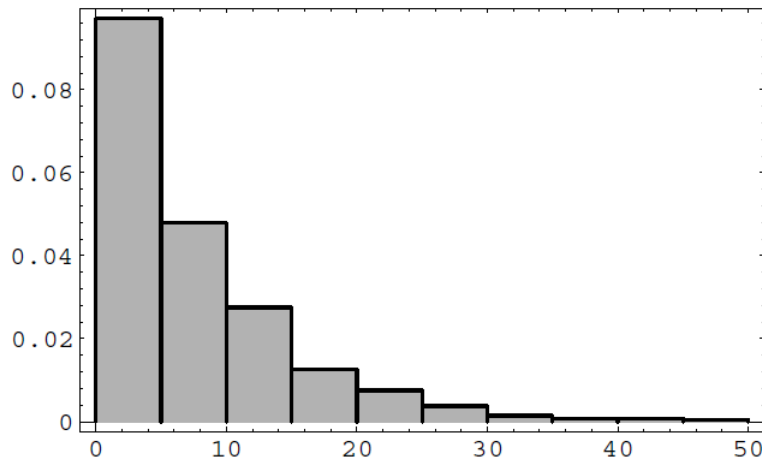


Figure 6.3.1: Distribution of queue waiting times.

<sup>1</sup> Ross, A First Course in Probability, (New York: Macmillan, 1984), pgs. 241-245

<sup>2</sup> L. Kleinrock, Queueing Systems, vol. 2 (New York: John Wiley and Sons, 1975).

<sup>3</sup> *ibid.*, p. 17.

<sup>4</sup> S. M. Ross, Applied Probability Models with Optimization Applications, (San Francisco:Holden-Day, 1970)

## Exercises

### 6.3.1

Let  $X$  be a random variable with range  $[-1, 1]$  and let  $f_X(x)$  be the density function of  $X$ . Find  $\mu(X)$  and  $\sigma^2(X)$  if, for  $|x| < 1$ ,

- (a)  $f_X(x) = 1/2$ .
- (b)  $f_X(x) = |x|$ .
- (c)  $f_X(x) = 1 - |x|$ .
- (d)  $f_X(x) = (3/2)x^2$ .

### 6.3.2

Let  $X$  be a random variable with range  $[-1, 1]$  and  $f_X$  its density function. Find  $\mu(X)$  and  $\sigma^2(X)$  if, for  $|x| > 1$ ,  $f_X(x) = 0$ , and for  $|x| < 1$ ,

- (a)  $f_X(x) = (3/4)(1 - x^2)$ .
- (b)  $f_X(x) = (\pi/4)\cos(\pi x/2)$ .
- (c)  $f_X(x) = (x+1)/2$ .
- (d)  $f_X(x) = (3/8)(x+1)^2$ .

### 6.3.3

The lifetime, measure in hours, of the ACME super light bulb is a random variable  $T$  with density function  $f_T(t) = \lambda^2 t e^{-\lambda t}$ , where  $\lambda = .05$ . What is the expected lifetime of this light bulb? What is its variance?

### 6.3.4

Let  $X$  be a random variable with range  $[-1, 1]$  and density function  $f_X(x) = ax + b$  if  $|x| < 1$ .

- (a) Show that if  $\int_{-1}^{+1} f_X(x) dx = 1$ , then  $b = 1/2$ .
- (b) Show that if  $f_X(x) \geq 0$ , then  $-1/2 \leq a \leq 1/2$ .
- (c) Show that  $\mu = (2/3)a$ , and hence that  $-1/3 \leq \mu \leq 1/3$ .
- (d) Show that  $\sigma^2(X) = (2/3)b - (4/9)a^2 = 1/3 - (4/9)a^2$ .

### 6.3.5

Let  $X$  be a random variable with range  $[-1, 1]$  and density function  $f_X(x) = ax^2 + bx + c$  if  $|x| < 1$  and 0 otherwise.

- Show that  $2a/3 + 2c = 1$  (see Exercise 4).
- Show that  $2b/3 = \mu(X)$ .
- Show that  $2a/5 + 2c/3 = \sigma^2(X)$ .
- Find  $a, b$ , and  $c$  if  $\mu(X) = 0$ ,  $\sigma^2(X) = 1/15$ , and sketch the graph of  $f_X$ .
- Find  $a, b$ , and  $c$  if  $\mu(X) = 0$ ,  $\sigma^2(X) = 1/2$ , and sketch the graph of  $f_X$ .

### 6.3.6

Let  $T$  be a random variable with range  $[0, \infty]$  and  $f_T$  its density function. Find  $\mu(T)$  and  $\sigma^2(T)$  if, for  $t < 0$ ,  $f_T(t) = 0$ , and for  $t > 0$ ,

- $f_T(t) = 3e^{-3t}$ .
- $f_T(t) = 9te^{-3t}$ .
- $f_T(t) = 3/(1+t)^4$ .

### 6.3.7

Let  $X$  be a random variable with density function  $f_X$ . Show, using elementary calculus, that the function

$$\phi(a) = E((X - a)^2) \quad (6.3.43)$$

takes its minimum value when  $a = \mu(X)$ , and in that case  $\phi(a) = \sigma^2(X)$ .

### 6.3.8

Let  $X$  be a random variable with mean  $\mu$  and variance  $\sigma^2$ . Let  $Y = aX^2 + bX + c$ . Find the expected value of  $Y$ .

### 6.3.9

Let  $X, Y$ , and  $Z$  be independent random variables, each with mean  $\mu$  and variance  $\sigma^2$ .

- Find the expected value and variance of  $S = X + Y + Z$ .
- Find the expected value and variance of  $A = (1/3)(X + Y + Z)$ .
- Find the expected value of  $S^2$  and  $A^2$ .

### 6.3.10

Let  $X$  and  $Y$  be independent random variables with uniform density functions on  $[0, 1]$ . Find

- $E(|X - Y|)$ .
- $E(\max(X, Y))$ .
- $E(\min(X, Y))$ .
- $E(X^2 + Y^2)$ .
- $E((X + Y)^2)$ .

### 6.3.11

The Pilsdorff Beer Company runs a fleet of trucks along the 100 mile road from Hangtown to Dry Gulch. The trucks are old, and are apt to break down at any point along the road with equal probability. Where should the company locate a garage so as to minimize the expected distance from a typical breakdown to the garage? In other words, if  $X$  is a random variable giving the location of the breakdown, measured, say, from Hangtown, and  $b$  gives the location of the garage, what choice of  $b$  minimizes  $E(|X - b|)$ ? Now suppose  $X$  is not distributed uniformly over  $[0, 100]$ , but instead has density function  $f_X(x) = 2x/10,000$ . Then what choice of  $b$  minimizes  $E(|X - b|)$ ?

### 6.3.12

Find  $E(X^Y)$ , where  $X$  and  $Y$  are independent random variables which are uniform on  $[0, 1]$ . Then verify your answer by simulation.

### 6.3.13

Let  $X$  be a random variable that takes on nonnegative values and has distribution function  $F(x)$ . Show that

$$E(X) = \int_0^{\infty} (1 - F(x))dx. \quad (6.3.44)$$

Hint: Integrate by parts.

Illustrate this result by calculating  $E(X)$  by this method if  $X$  has an exponential distribution  $F(x) = 1 - e^{-\lambda x}$  for  $x \geq 0$ , and  $F(x) = 0$  otherwise.

### 6.3.14

Let  $X$  be a continuous random variable with density function  $f_X(x)$ . Show that if

$$\int_{-\infty}^{+\infty} x^2 f_X(x)dx < \infty \quad (6.3.45)$$

then

$$\int_{-\infty}^{+\infty} |x| f_X(x)dx < \infty. \quad (6.3.46)$$

Hint: Except on the interval  $[-1, 1]$ , the first integrand is greater than the second integrand.

### 6.3.15

Let  $X$  be a random variable distributed uniformly over  $[0, 20]$ . Define a new random variable  $Y$  by  $Y = \lfloor X \rfloor$  (the greatest integer in  $X$ ). Find the expected value of  $Y$ . Do the same for  $Z = \lfloor X + .5 \rfloor$ . Compute  $E(|X - Y|)$  and  $E(|X - Z|)$ . (Note that  $Y$  is the value of  $X$  rounded off to the nearest smallest integer, while  $Z$  is the value of  $X$  rounded off to the nearest integer. Which method of rounding off is better? Why?)

### 6.3.16

Assume that the lifetime of a diesel engine part is a random variable  $X$  with density  $f_X$ . When the part wears out, it is replaced by another with the same density. Let  $N(t)$  be the number of parts that are used in time  $t$ . We want to study the random variable  $N(t)/t$ . Since parts are replaced on the average every  $E(X)$  time units, we expect about  $t/E(X)$  parts to be used in time  $t$ . That is, we expect that

$$\lim_{t \rightarrow \infty} E\left(\frac{N(t)}{t}\right) = \frac{1}{E(X)}. \quad (6.3.47)$$

This result is correct but quite difficult to prove. Write a program that will allow you to specify the density  $f_X$ , and the time  $t$ , and simulate this experiment to find  $N(t)/t$ . Have your program repeat the experiment 500 times and plot a bar graph for the random outcomes of  $N(t)/t$ . From this data, estimate  $E(N(t)/t)$  and compare this with  $1/E(X)$ . In particular, do this for  $t = 100$  with the following two densities:

- (a)  $f_X = e^{-t}$ .
- (b)  $f_X = te^{-t}$ .

### 6.3.17

Let  $X$  and  $Y$  be random variables. The covariance  $\text{Cov}(X, Y)$  is defined by (see Exercise 6.2.23)

$$\text{cov}(X, Y) = E((X - \mu(X))(Y - \mu(Y))). \quad (6.3.48)$$

- (a) Show that  $\text{cov}(X, Y) = E(XY) - E(X)E(Y)$ .
- (b) Using (a), show that  $\text{cov}(X, Y) = 0$ , if  $X$  and  $Y$  are independent. (Caution: the converse is not always true.)
- (c) Show that  $V(X + Y) = V(X) + V(Y) + 2\text{cov}(X, Y)$ .

### 6.3.18

Let  $X$  and  $Y$  be random variables with positive variance. The correlation of  $X$  and  $Y$  is defined as

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{V(X)V(Y)}}. \quad (6.3.49)$$

(a) Using Exercise 17(c), show that

$$0 \leq V\left(\frac{X}{\sigma(X)} + \frac{Y}{\sigma(Y)}\right) = 2(1 + \rho(X, Y)). \quad (6.3.50)$$

(b) Now show that

$$0 \leq V\left(\frac{X}{\sigma(X)} - \frac{Y}{\sigma(Y)}\right) = 2(1 - \rho(X, Y)). \quad (6.3.51)$$

(c) Using (a) and (b), show that

$$-1 \leq \rho(X, Y) \leq 1. \quad (6.3.52)$$

### 6.3.19

Let  $X$  and  $Y$  be independent random variables with uniform densities in  $[0, 1]$ . Let  $Z = X + Y$  and  $W = X - Y$ . Find

(a)  $\rho(X, Y)$  (see Exercise 18).

(b)  $\rho(X, Z)$ .

(c)  $\rho(Y, W)$ .

(d)  $\rho(Z, W)$ .

### 6.3.20

When studying certain physiological data, such as heights of fathers and sons, it is often natural to assume that these data (e.g., the heights of the fathers and the heights of the sons) are described by random variables with normal densities. These random variables, however, are not independent but rather are correlated. For example, a two-dimensional standard normal density for correlated random variables has the form

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot e^{-(x^2 - 2\rho xy + y^2)/2(1-\rho^2)}. \quad (6.3.53)$$

(a) Show that  $X$  and  $Y$  each have standard normal densities.

(b) Show that the correlation of  $X$  and  $Y$  (see Exercise 18) is  $\rho$ .

### 6.3.21

For correlated random variables  $X$  and  $Y$  it is natural to ask for the expected value for  $X$  given  $Y$ . For example, Galton calculated the expected value of the height of a son given the height of the father. He used this to show that tall men can be expected to have sons who are less tall on the average. Similarly, students who do very well on one exam can be expected to do less well on the next exam, and so forth. This is called regression on the mean. To define this conditional expected value, we first define a conditional density of  $X$  given  $Y = y$  by

$$f_{X|Y}(x | y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}, \quad (6.3.54)$$

where  $f_{X,Y}(x, y)$  is the joint density of  $X$  and  $Y$ , and  $f_Y$  is the density for  $Y$ . Then the conditional expected value of  $X$  given  $Y$  is

$$E(X | Y = y) = \int_a^b x f_{X|Y}(x | y) dx. \quad (6.3.55)$$

For the normal density in Exercise 20, show that the conditional density of  $f_{X|Y}(x|y)$  is normal with mean  $\rho y$  and variance  $1 - \rho^2$ . From this we see that if  $X$  and  $Y$  are positively correlated ( $0 < \rho < 1$ ), and if  $y > E(Y)$ , then the expected value for  $X$  given  $Y = y$  will be less than  $y$  (i.e., we have regression on the mean).

### 6.3.22

A point  $Y$  is chosen at random from  $[0, 1]$ . A second point  $X$  is then chosen from the interval  $[0, Y]$ . Find the density for  $X$ . Hint: Calculate  $f_{X|Y}$  as in Exercise 21 and then use

$$f_X(x) = \int_x^1 f_{X|Y}(x|y) f_Y(y) dy. \quad (6.3.56)$$

Can you also derive your result geometrically?

### 6.3.23

Let  $X$  and  $V$  be two standard normal random variables. Let  $\rho$  be a real number between -1 and 1.

(a) Let  $Y = \rho X + \sqrt{1 - \rho^2} V$ . Show that  $E(Y) = 0$  and  $\text{Var}(Y) = 1$ . We shall see later (see Example 7.5 and Example 10.17), that the sum of two independent normal random variables is again normal. Thus, assuming this fact, we have shown that  $Y$  is standard normal.

(b) Using Exercises 17 and 18, show that the correlation of  $X$  and  $Y$  is  $\rho$ .

(c) In Exercise 20, the joint density function  $f_{X,Y}(x, y)$  for the random variable  $(X, Y)$  is given. Now suppose that we want to know the set of points  $(x, y)$  in the  $xy$ -plane such that  $f_{X,Y}(x, y) = C$  for some constant  $C$ . This set of points is called a set of constant density. Roughly speaking, a set of constant density is a set of points where the outcomes  $(X, Y)$  are equally likely to fall. Show that for a given  $C$ , the set of points of constant density is a curve whose equation is

$$x^2 - 2\rho xy + y^2 = D, \quad (6.3.57)$$

where  $D$  is a constant which depends upon  $C$ . (This curve is an ellipse.)

(d) One can plot the ellipse in part (c) by using the parametric equations

$$\begin{aligned} x &= \frac{r \cos \theta}{\sqrt{2(1-\rho)}} + \frac{r \sin \theta}{\sqrt{2(1+\rho)}}, \\ y &= \frac{r \cos \theta}{\sqrt{2(1-\rho)}} - \frac{r \sin \theta}{\sqrt{2(1+\rho)}}. \end{aligned} \quad (6.3.58)$$

Write a program to plot 1000 pairs  $(X, Y)$  for  $\rho = -1/2, 0, 1/2$ . For each plot, have your program plot the above parametric curves for  $r = 1, 2, 3$ .

### 6.3.24

Following Galton, let us assume that the fathers and sons have heights that are dependent normal random variables. Assume that the average height is 68 inches, standard deviation is 2.7 inches, and the correlation coefficient is .5 (see Exercises 20 and 21). That is, assume that the heights of the fathers and sons have the form  $2.7X + 68$  and  $2.7Y + 68$ , respectively, where  $X$  and  $Y$  are correlated standardized normal random variables, with correlation coefficient .5.

(a) What is the expected height for the son of a father whose height is 72 inches?

(b) Plot a scatter diagram of the heights of 1000 father and son pairs. Hint: You can choose standardized pairs as in Exercise 23 and then plot  $(2.7X + 68, 2.7Y + 68)$ .

### 6.3.25

When we have pairs of data  $(x_i, y_i)$  that are outcomes of the pairs of dependent random variables  $X, Y$  we can estimate the correlation coefficient  $\rho$  by

$$\bar{r} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_X s_Y}, \quad (6.3.59)$$

where  $\bar{x}$  and  $\bar{y}$  are the sample means for  $X$  and  $Y$ , respectively, and  $s_X$  and  $s_Y$  are the sample standard deviations for  $X$  and  $Y$  (see Exercise 6.2.17). Write a program to compute the sample means, variances, and correlation for such dependent data. Use your program to compute these quantities for Galton's data on heights of parents and children given in Appendix B.

Plot the equal density ellipses as defined in Exercise 23 for  $r = 4, 6$ , and  $8$ , and on the same graph print the values that appear in the table at the appropriate points. For example, print 12 at the point  $(70.5, 68.2)$  indicating that there were 12 cases where the parent's height was 70.5 and the child's was 68.12. See if Galton's data is consistent with the equal density ellipses.

### 6.3.26

(from Hamming <sup>25</sup>) Suppose you are standing on the bank of a straight river.

(a) Choose, at random, a direction which will keep you on dry land, and walk 1 km in that direction. Let  $P$  denote your position. What is the expected distance from  $P$  to the river?

(b) Now suppose you proceed as in part (a), but when you get to  $P$ , you pick a random direction (from among all directions) and walk 1 km. What is the probability that you will reach the river before the second walk is completed?

### 6.3.27

(from Hamming <sup>26</sup>) A game is played as follows: A random number  $X$  is chosen uniformly from  $[0, 1]$ . Then a sequence  $Y_1, Y_2, \dots$  of random numbers is chosen independently and uniformly from  $[0, 1]$ . The game ends the first time that  $Y_i > X$ . You are then paid  $(i - 1)$  dollars. What is a fair entrance fee for this game?

### 6.3.28

A long needle of length  $L$  much bigger than 1 is dropped on a grid with horizontal and vertical lines one unit apart. Show that the average number  $a$  of lines crossed is approximately

$$a = \frac{4L}{\pi}. \quad (6.3.60)$$

---

This page titled [6.3: Continuous Random Variables](#) 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.