5 Continuous Joint Distributions

The *joint distribution* of a pair of random variables X and Y is the probability distribution over the plane defined by

$$P(B) = P((X, Y) \in B)$$

for subsets B of the plane. So $P(B)$ is the probability that the random pair (X, Y) falls in the set B. Joint distributions for discrete random variables were considered in Section 3.1. This chapter shows how these ideas for discrete random variables are extended to two or more continuously distributed random variables with sums replaced by integrals.

Section 5.1 concerns the simplest kind of continuous joint distribution, a *uniform* distribution defined by relative areas. Section 5.2 introduces the concept of a *joint density function.* Joint probabilities are then defined by volumes under a density surface. The important special case of independent normal variables is studied in Section 5.3. Then Section 5.4 deals with a general technique for finding the distribution of a function of two variables.

5.1 Uniform Distributions

The uniform distribution on an interval was discussed in Section 4.1. The idea extends to higher dimensions with relative lengths ıeplaced by relative areas or relative volumes. For example, a random point (X, Y) in the plane has uniform distribution on D, where D is a region of the plane with finite area, if:

(i) (X, Y) is certain to lie in D;

(ii) the chance that (X, Y) falls in a subregion C of D is proportional to the area of C

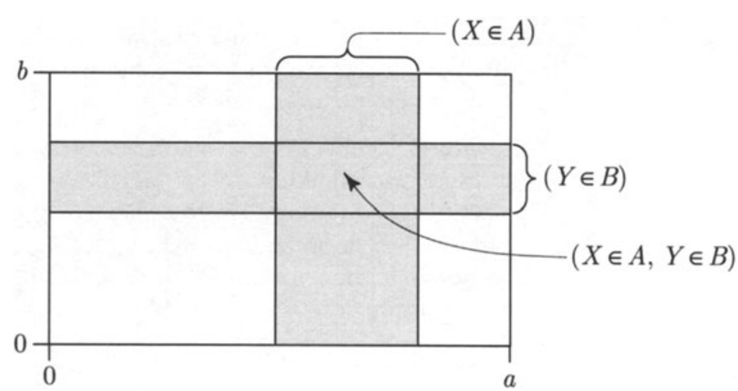
$$P((X, Y) \in C) = \frac{\text{area } (C)}{\text{area } (D)} \quad \text{for} \quad C \subset D$$

Here is an important observation:

Independent Uniform Variables

If X and Y are independent random variables, each uniformly distributed on an interval, then (X, Y) is uniformly distributed on a rectangle.

To see why, suppose X and Y are independent and uniformly distributed on, say, $(0, a)$ and $(0, b)$, respectively. For intervals A and B the event $(X \in A, Y \in B)$ is the event that (X, Y) falls in the rectangle $A \times B$, as shown in the following Venn diagram:



So for any rectangle $A \times B$

$$P((X,Y) \in A \times B) = P(X \in A, Y \in B)$$
$$= P(X \in A)P(Y \in B) \quad \text{by independence of } X \text{ and } Y$$
$$= \frac{\text{length }(A)}{a} \cdot \frac{\text{length }(B)}{b} \quad \text{by assumed uniform distributions of } X \text{ and } Y$$
$$= \frac{\text{area }(A \times B)}{ab}$$

Thus the probability that $(X,Y) \in C$ is the relative area of C in $(0,a) \times (0,b)$ for every rectangle C. The same must then be true for finite unions of rectangles, by the addition rule of probability and for area, hence also for any set C whose area can be defined by approximating with unions of rectangles. *Conclusion*: (X,Y) has uniform distribution on the rectangle $(0,a) \times (0,b)$.

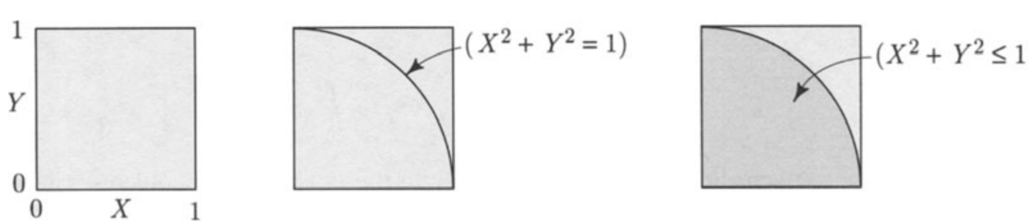
The above observation allows probabilities involving two independent uniform variables X and Y to be found geometrically in terms of areas. The key step is correct identification of areas in the plane corresponding to events in question. Skill at doing this is essential for all further work in this chapter.

Example 1. Probabilities for two independent uniform random variables.
Suppose X and Y are independent uniform $(0,1)$ random variables.

Problem 1. Find $P(X^2 + Y^2 \leq 1)$.

Solution. Proceed by 3 steps as in the diagram below:

- Draw a unit square with coordinates X, Y.
- Notice that $X^2 + Y^2 = 1$ gives the equation of a circle of radius 1.
- Recognize $(X^2 + Y^2 \leq 1)$ as the region inside both the square and circle.
- Use the formula for the area of a circle to get $P(X^2 + Y^2 \leq 1) = \frac{\pi}{4}$.

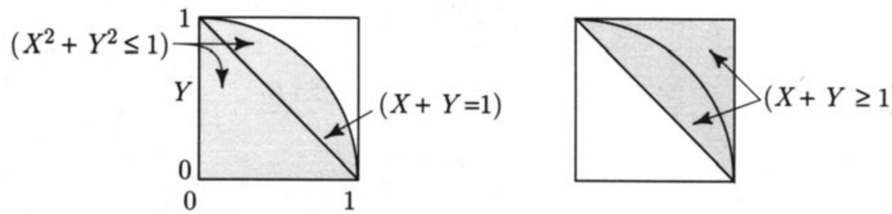


Problem 2. Find the conditional probability $P(X^2 + Y^2 \leq 1 | X + Y \geq 1)$.

Solution. After first identifying $X^2 + Y^2 \leq 1$ as above, next:

- Recognize $(X + Y = 1)$ as the line through the points $(0,1)$ and $(1,0)$.

– Deduce that $(X + Y \geq 1)$ is the shaded region above this line.



– Now compute the required relative area:

$$P(X^2 + Y^2 \leq 1 | X + Y \geq 1) = \frac{P(X^2 + Y^2 \leq 1, X + Y \geq 1)}{P(X + Y \geq 1)}$$
$$= \frac{\pi/4 - 1/2}{1/2} = \frac{\pi}{2} - 1$$

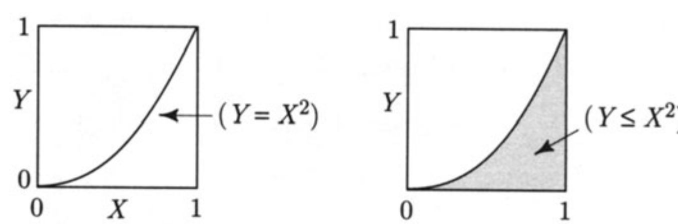
Problem 3. Find $P(Y \leq X^2)$.

Solution.

– Graph $Y = X^2$.

– Recognize $(Y \leq X^2)$ as the region under this graph.

– Compute the area of this region by calculus.



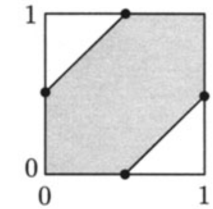
$$P(Y \leq X^2) = \int_0^1 x^2 dx = \frac{1}{3}x^3 \Big|_0^1 = \frac{1}{3}$$

Discussion. Note well how only in the last of these problems was it necessary to resort to calculus to find the area. *Always* sketch the relevant regions first, then look out for familiar shapes, rectangles, triangles, and circles. If all else fails, use calculus.

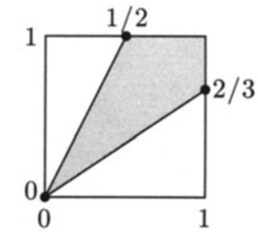
Example 2. **More probabilities for two independent uniform variables.**

Let X and Y be independent random variables, each uniformly distributed on $(0, 1)$. Calculate the following probabilities:

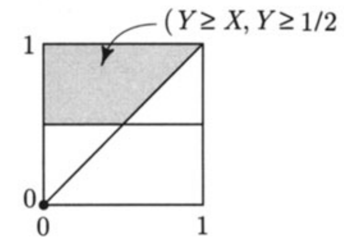
a) $$P(|X - Y| \le 0.5) = \text{indicated area} = 1 - \frac{1}{4} = 0.75$$



b) $$P\left(\left|\frac{X}{Y} - 1\right| \le 0.5\right) = P\left(\frac{2}{3}X \le Y \le 2X\right) = \text{indicated area} = 1 - \frac{1}{2}\left(\frac{1}{2} + \frac{2}{3}\right) = \frac{5}{12}$$



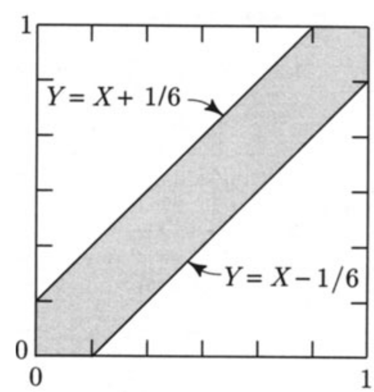
c) $$P\left(Y \ge X | Y \ge \frac{1}{2}\right) = \text{indicated area}/\frac{1}{2} = \left(\frac{1}{2} - \frac{1}{8}\right)/\frac{1}{2} = \frac{3}{4}$$



Example 3. Probability of meeting.

Problem. Two people try to meet at a certain place between 5:00 P.M. and 5:30 P.M. Suppose that each person arrives at a time distributed uniformly at random in this time interval, independent of the other, and waits for the other at most 5 minutes. What is the probability that they meet?

Solution. Let X and Y be the arrival times measured as fractions of the 30 minute interval, starting from 5:00 P.M. Then X and Y are independent uniform $(0, 1)$ random variables. The people meet if and only if $|X - Y| \le 1/6$.



Desired probability = indicated area = $1 - (\frac{5}{6})^2 = \frac{11}{36}$

Uniform Distribution over a Volume

This is the extension of the idea of relative lengths in one dimension and relative areas in two dimensions to relative volumes in three and higher dimensions. If $U_1, \ldots, U_n$ are n independent random variables, with U_i uniformly distributed on an interval (a_i, b_i), then the same argument given earlier for the case $n = 2$ shows that the joint distribution of $(U_1, \ldots, U_n)$ is the uniform distribution defined by relative volumes within the n-dimesional box

$$(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$$

whose n-dimensional volume is the product $(b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$ of the lengths of its sides.

To illustrate, a random point in the *unit cube* $(0, 1) \times (0, 1) \times (0, 1)$, with approximately independent coordinates, is obtained by three successive calls of a pseudo-random number generator, say $(\text{RND}_1, \text{RND}_2, \text{RND}_3)$. For any subvolume B of the unit cube bounded by a reasonably smooth surface (e.g., the portion of a box, pyramid, or sphere that lies inside the unit cube) the long-run frequency of times that $(\text{RND}_1, \text{RND}_2, \text{RND}_3)$ is in B will be approximately the volume of B, that is $P(B)$ for the uniform distribution on the unit cube. For example, the long-run frequency of triples $(\text{RND}_1, \text{RND}_2, \text{RND}_3)$ with

$$(\text{RND}_1 - \tfrac{1}{2})^2 + (\text{RND}_2 - \tfrac{1}{2})^2 + (\text{RND}_3 - \tfrac{1}{2})^2 < 1/4$$

is approximately the volume of the subset of the unit cube

$$\{(x, y, z) : 0{<}x{<}1,\ 0{<}y{<}1,\ 0{<}z{<}1,\ (x - \tfrac{1}{2})^2 + (y - \tfrac{1}{2})^2 + (z - \tfrac{1}{2})^2{<}1/4\}$$

This is the volume of a sphere of radius $\frac{1}{2}$ centered at $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, which is $\frac{4}{3}\pi \left(\frac{1}{2}\right)^3 = \frac{\pi}{6}$.

Exercises 5.1

1. Let (X, Y) have uniform distribution on the set

$$\{(x, y) : 0 < x < 2 \text{ and } 0 < y < 4 \text{ and } x < y\}.$$

Find: a) $P(X < 1)$; b) $P(Y < X^2)$.

2. A metal rod is l inches long. Measurements on the length of this rod are equal to l plus random error. Assume that the errors are uniformly distributed over the range -0.1 inch to $+0.1$ inch, and are independent of each other.

a) Find the chance that a measurement is less than $1/100$ of an inch away from l.

b) Find the chance that two measurements are less than $1/100$ of an inch away from each other.

3. Suppose X and Y are independent and uniformly distributed on the unit interval $(0, 1)$. Find:

$$P(Y \geq \frac{1}{2} \,|\, Y \geq 1 - 2X).$$

4. Let X and Y be independent random variables each uniformly distributed on $(0, 1)$. Find:

a) $P(|X - Y| \leq 0.25)$; b) $P(|X/Y - 1| \leq 0.25)$; c) $P(Y \geq X \,|\, Y \geq 0.25)$.

5. A very large group of students takes a test. Each of them is told his or her percentile rank among all students taking the test.

a) If a student is picked at random from all students taking the test, what is the probability that the student's percentile rank is over 90%?

b) If two students are picked independently at random, what is the probability that their percentile ranks differ by more than 10%?

6. A group of 10 people agree to meet for lunch at a cafe between 12 noon and 12:15 P.M. Assume that each person arrives at the cafe at a time uniformly distributed between noon and 12:15 P.M., and that the arrival times are independent of each other.

a) Jack and Jill are two members of the group. Find the probability that Jack arrives at least two minutes before Jill.

b) Find the probability of the event that the first of the 10 persons to arrive does so by 12:05 P.M., and the last person arrives after 12:10 P.M.

7. Let X and Y be two independent uniform $(0, 1)$ random variables. Let M be the smaller of X and Y. Let $0 < x < 1$.

a) Represent the event $(M \geq x)$ as the region in the plane, and find $P(M \geq x)$ as the area of this region.

b) Use your result in a) to find the c.d.f. and density of M. Sketch the graph of these functions.

8. Let $U_{(1)}, \ldots, U_{(n)}$ be the values of n independent uniform $(0, 1)$ random variables arranged in increasing order. Let $0 \leq x < y \leq 1$.

a) Find and justify a simple formula for $P(U_{(1)} > x \text{ and } U_{(n)} < y)$.

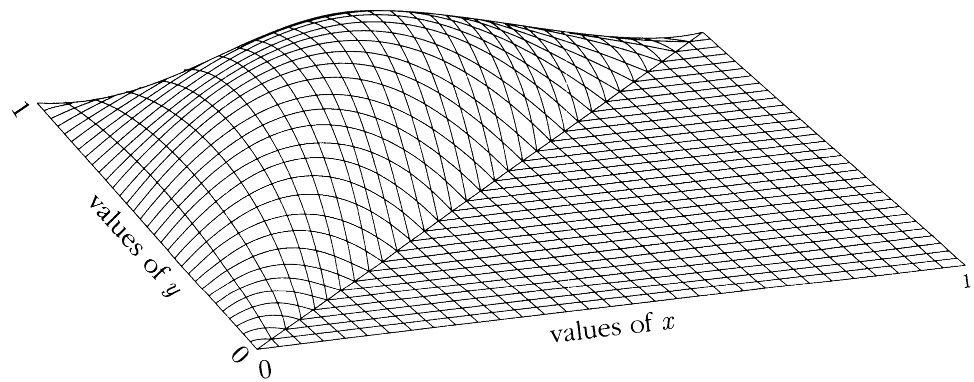
b) Find a formula for $P(U_{(1)} \leq x \text{ and } U_{(n)} < y)$.

9. **A triangle problem.** Suppose a straight stick is broken in three at two points chosen independently at random along its length. What is the chance that the three sticks so formed can be made into the sides of a triangle?

5.2 Densities

The concept of a *joint probability density function* $f(x, y)$ for a pair of random variables X and Y is a natural extension of the idea of a one-dimensional probability density function studied in Chapter 4. The function $f(x, y)$ gives the density of probability per unit area for values of (X, Y) near the point (x, y).

FIGURE 1. A joint density surface. Here a particular joint density function given by the formula $f(x, y) = 5!\,x(y - x)(1 - y) \quad (0 < x < y < 1)$, is viewed as the height of a surface over the unit square $0 \le x \le 1, 0 \le y \le 1$. As explained later in Example 3, two random variables X and Y with this joint density are the second and fourth smallest of five independent uniform $(0, 1)$ variables. But for now the source and special form of this density are not important. Just view it as a typical joint density surface.

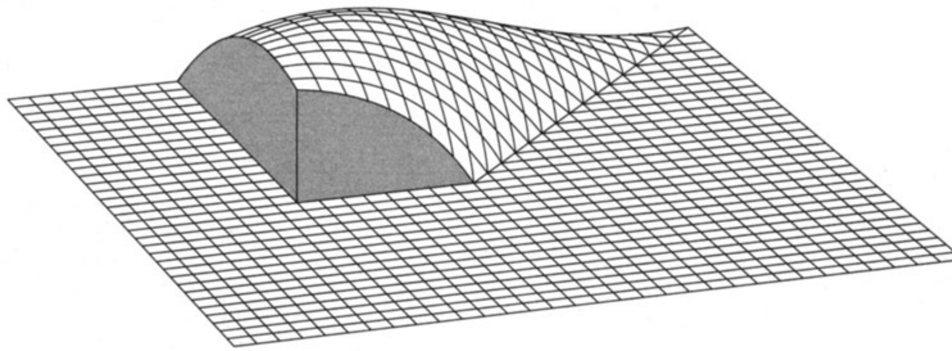


Examples in the previous section show how any event determined by two random variables X and Y, like the event $(X > 0.25$ and $Y > 0.5)$, corresponds to a region of the plane. Now instead of a uniform distribution defined by relative areas, the probability of region B is defined by the volume under the density surface over B. This volume is an integral

$$P((X, Y) \in B) = \iint_B f(x, y)\,dx\,dy$$

This is the analog of the familiar area under the curve interpretation for probabilities obtained from densities on a line. Examples to follow show how such integrals can be computed by repeated integration, change of variables, or symmetry arguments. Uniform distribution over a region is now just the special case when $f(x, y)$ is constant over the region and zero elsewhere. As a general rule, formulae involving joint densities are analogous to corresponding formulae for discrete joint distributions described in Section 3.1. See pages 348 and 349 for a summary.

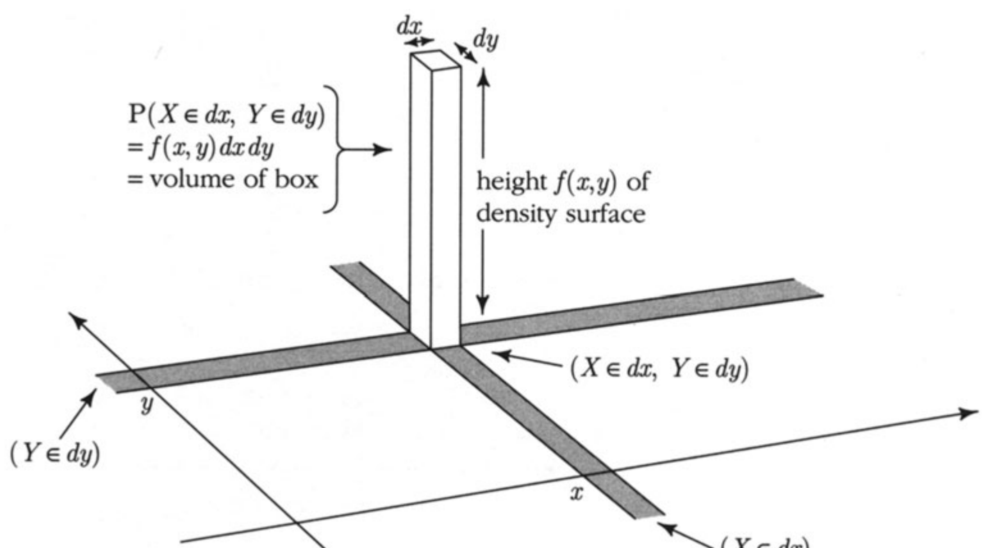
FIGURE 2. **Volume representing a probability.** The probability $P(X > 0.25$ and $Y > 0.5)$, for random variables X and Y with the joint density of Figure 1. The set B in this case is $\{(x, y) : x > 0.25$ and $y > 0.5\}$. You can see the volume is about half the total volume under the surface. The exact value, found later in Example 3, is $27/64$.



Informally, if (X, Y) has joint density $f(x, y)$, then there is the *infinitesimal probability formula*

$$P(X \in dx, Y \in dy) = f(x, y)dx\,dy$$

This means that the probability that the pair (X, Y) falls in an infinitesimal rectangle of width dx and height dy near the point (x, y) is the probability density at (x, y) multiplied by the area $dx\,dy$ of the rectangle.



Discrete Joint Distribution

Probability of a point:

$$P(X = x, Y = y) = P(x, y)$$

The joint probability $P(x, y)$ is the probability of the single point (x, y).

Probability of a set B: The sum of probabilities of points in B

$$P((X, Y) \in B) = \sum_{(x,y)\in B} P(x, y)$$

Constraints: Non-negative with total sum 1

$$P(x, y) \geq 0 \qquad \text{and} \qquad \sum_{\text{all } x} \sum_{\text{all } y} P(x, y) = 1$$

Marginals:

$$P(X = x) = \sum_{\text{all } y} P(x, y)$$

$$P(Y = y) = \sum_{\text{all } x} P(x, y)$$

Independence: $P(x, y) = P(X = x)P(Y = y)$ (for all x and y)

Expectation of a function g of (X,Y), e.g., XY,

$$E(g(X, Y)) = \sum_{\text{all } x} \sum_{\text{all } y} g(x, y)P(x, y)$$

provided the sum converges absolutely.

Joint Distribution Defined by a Density

Infinitesimal probability:

$$P(X \in dx, Y \in dy) = f(x,y)\,dx\,dy$$

The joint density $f(x,y)$ is the probability per unit area for values near (x,y).

Probability of a set B: The volume under the density surface over B

$$P((X,Y) \in B) = \iint_B f(x,y)\,dx\,dy$$

Constraints: Non-negative with total integral 1

$$f(x,y) \geq 0 \qquad \text{and} \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

Marginals:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

Independence: $f(x,y) = f_X(x)f_Y(y)$ (for all x and y)

Expectation of a function g of (X,Y), e.g., XY

$$E(g(X,Y)) = \iint g(x,y)f(x,y)dx\,dy$$

provided the integral converges absolutely.

The infinitesimal probability formula

$$P(X \in dx, Y \in dy) = f(x,y)dx\,dy$$

is really shorthand for a limiting statement about the ratio of probability per unit area for small areas, which, strictly speaking, holds only at points (x,y) such that the joint density is continuous at (x,y). But the infinitesimal formula conveys the right intuitive idea, and can be manipulated to obtain useful formulae which turn out to be valid even without assuming that the joint density is continuous.

Marginal densities. If (X,Y) has a joint density $f(x,y)$ in the plane, then each of the random variables X and Y has a density on the line. These are called the *marginal densities*. As shown in the preceding display, the marginal densities can be calculated from the joint density by integral analogs of the discrete formulae for marginal probabilities as row and column sums in a joint distribution table. Probabilities of discrete points are replaced by densities, and sums by integrals.

Independence. In general, random variables X and Y are called independent if

(1) $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ for all choices of sets A and B.

Joint Density for Independent Variables

Random variables X and Y with joint density $f(x,y)$ are independent if and only if the joint density is the product of the two marginal densities:

(2) $$f(x,y) = f_X(x)f_Y(y) \qquad \text{(for all } x \text{ and } y\text{)}$$

Intuitively (2) follows from (1) by taking A to be a small interval $(x, x+dx)$ near x, B a small interval $(y, y+dy)$ near y, to obtain

(3) $$P(X \in dx, Y \in dy) = P(X \in dx)P(Y \in dy)$$

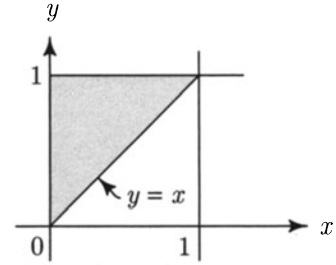
$$\text{so} \quad f(x,y)\,dx\,dy = f_X(x)\,dx\,f_Y(y)\,dy$$

Cancelling the differentials dx and dy leaves the product formula for densities. Conversely, (1) is obtained from (2) by integration.

Example 1. Uniform on a triangle.

Suppose (X, Y) is uniformly distributed over the region $\{(x, y) : 0 < x < y < 1\}$.

Problem 1. Find the joint density of (X, Y).



Solution. By the assumption, $f(x, y) = c$ for $0 < x < y < 1$ and 0 elsewhere. Because the triangle has area $\frac{1}{2}$, $c = 2$.

Problem 2. Find the marginal densities $f_X(x)$ and $f_Y(y)$.

Solution.

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x, y)\, dy \\ &= \int_{y=x}^{y=1} 2dy \qquad \text{since } f(x, y) = 2 \text{ for } 0 < x < y < 1, \quad 0 \text{ elsewhere} \\ &= 2(1 - x) \qquad \text{for } 0 < x < 1 \quad \text{and 0 elsewhere.} \end{aligned}$$

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x, y)\, dx \\ &= \int_{x=0}^{x=y} 2dx \qquad \text{since } f(x, y) = 2 \text{ for } 0 < x < y < 1 \quad 0 \text{ elsewhere} \\ &= 2y \qquad \text{for } 0 < y < 1 \quad \text{and 0 elsewhere.} \end{aligned}$$

Problem 3. Are X and Y independent?

Solution. No, since $f(x, y) \neq f_X(x) f_Y(y)$.

Problem 4. Find $E(X)$ and $E(Y)$.

Solution.

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^1 2x(1 - x)\, dx = \frac{1}{3}$$

$$E(Y) = \int_{-\infty}^{\infty} y f_Y(y)\, dy = \int_0^1 2y^2\, dy = \frac{2}{3}$$

Problem 5. Find $E(XY)$.

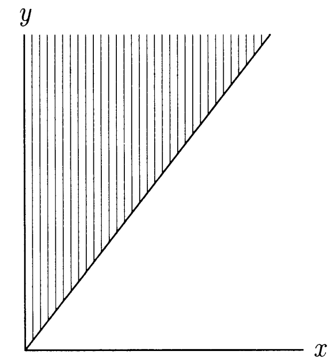
Solution.

$$E(XY) = \iint_{R^2} xy f(x, y)\, dx\, dy = 2 \int_{y=0}^{1} dy \int_{x=0}^{y} xy\, dx = 2 \int_{y=0}^{1} \frac{y^3}{2}\, dy = \frac{1}{4}$$

Remark. You can show that the joint distribution of X and Y considered here is that of $X = \min(U, V)$, $Y = \max(U, V)$, where U and V are independent uniform $(0, 1)$ variables. Example 3 gives a more difficult derivation of this kind.

Example 2. Independent exponential variables.

Problem. Let X and Y be independent and exponentially distributed random variables with parameters λ and μ, respectively. Calculate $P(X < Y)$.

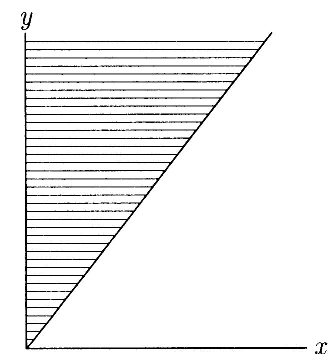


Solution. The joint density is

$$f(x, y) = (\lambda e^{-\lambda x})(\mu e^{-\mu y}) = \lambda\mu e^{-\lambda x - \mu y}$$

by independence. And $P(X < Y)$ is found by integration of this joint density over the set $\{(x, y) : x < y\}$:

$$\begin{aligned} P(X < Y) &= \iint_{x<y} \lambda\mu\, e^{-\lambda x - \mu y} dx\, dy \\ &= \int_{x=0}^{\infty} dx \int_{y=x}^{\infty} \lambda\mu\, e^{-\lambda x - \mu y} dy \\ &= \int_{x=0}^{\infty} \lambda e^{-\lambda x - \mu x} dx \\ &= \frac{\lambda}{\lambda + \mu} \end{aligned}$$



Remark. Done in the other order, the integral is

$$\int_{y=0}^{\infty} dy \int_{x=0}^{y} \lambda\mu\, e^{-\lambda x - \mu y} dx$$

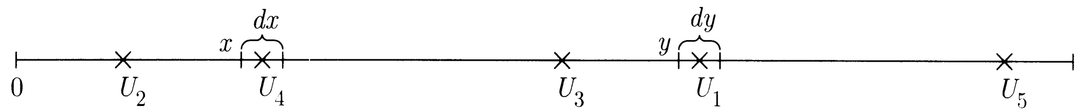
which simplifies to the same answer. As a general rule, provided the integrand is positive, as always when finding probabilities, double integrals done in either order produce the same result.

Example 3. Joint distribution of order statistics.

Suppose $U_{(1)} < U_{(2)} < \cdots < U_{(5)}$ are the order statistics of 5 independent uniform $(0, 1)$ variables $U_1, \ldots, U_5$, so $U_{(i)}$ is the ith smallest of $U_1, \ldots, U_5$, as, for example, in the following diagram:

Problem 1. Find the joint density of $U_{(2)}$ and $U_{(4)}$.

Solution. This is very like the calculation of the density of $U_{(i)}$ done in Section 4.6. The following diagram shows one way of getting $U_{(2)}$ in dx and $U_{(4)}$ in dy for $0 < x < y < 1$:



$$\begin{aligned}
&P(U_{(2)} \in dx, U_{(4)} \in dy) \\
&= P(\text{ one } U_i \text{ in } (0,x), \text{ one in } dx, \text{ one in } (x,y), \text{ one in } dy, \text{ one in } (y,1)\ ) \\
&= 5!\, P(U_2 \in (0,x), U_4 \in dx, U_3 \in (x,y), U_1 \in dy, U_5 \in (y,1)) \\
&= 5!\, x\, dx (y-x)\, dy (1-y)
\end{aligned}$$

Here the 5! is the number of different ways of deciding which variables fall in which intervals. The conclusion is that the joint density of $U_{(2)}$ and $U_{(4)}$ is

$$P(U_{(2)} \in dx, U_{(4)} \in dy)/dx\, dy = \begin{cases} 5!\, x(y-x)(1-y) & \text{for } 0 < x < y < 1 \\ 0 & \text{elsewhere} \end{cases}$$

This is the density surface shown in Figure 1 on page 346.

Problem 2. Find $P(U_{(2)} > 1/4 \text{ and } U_{(4)} > 1/2)$.

Solution. The volume representing this probability is shown in Figure 2 on page 347. This is the volume under the density surface over the area shaded in the diagram at right. This area is the intersection of:

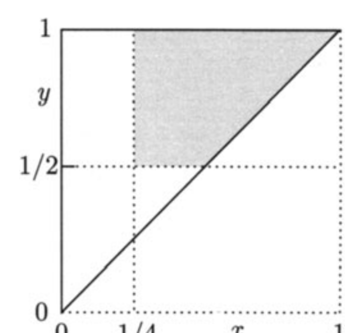
(i) the region representing the event; and

(ii) the region where the density is strictly positive.

This determines the ranges of integration. The required probability is thus

$$\begin{aligned}
&5! \int_{y=1/2}^{1} \int_{x=1/4}^{y} x(y-x)(1-y)\, dx\, dy \\
&= 5! \int_{y=1/2}^{1} (1-y)\, dy \left[\frac{1}{2}x^2 y - \frac{1}{3}x^3\right]\Bigg|_{1/4}^{y} \\
&= 5! \int_{y=1/2}^{1} (1-y)\, dy \left[\frac{y^3}{6} - \frac{y}{2^5} - \frac{1}{3 \times 2^6}\right] = \frac{27}{64}
\end{aligned}$$

by straightforward integration of the polynomial.



Exercises 5.2

1. Suppose that (X, Y) is uniformly distributed over the region $\{(x, y) : 0 < |y| < x < 1\}$. Find:

a) the joint density of (X, Y); b) the marginal densities $f_X(x)$ and $f_Y(y)$.

c) Are X and Y independent? d) Find $E(X)$ and $E(Y)$.

2. Repeat Exercise 1 for (X, Y) with uniform distribution over $\{(x, y) : 0 < |x| + |y| < 1\}$.

3. A random point (X, Y) in the unit square has joint density $f(x, y) = c(x^2 + 4xy)$ for $0 < x < 1$ and $0 < y < 1$, for some constant c.

a) Evaluate c. b) Find $P(X \leq a)$, $0 < a < 1$. c) Find $P(Y \leq b)$, $0 < b < 1$.

4. For random variables X and Y with joint density function

$$f(x, y) = 6e^{-2x-3y} \qquad (x, y > 0)$$

and $f(x, y) = 0$ otherwise, find:

a) $P(X \leq x, Y \leq y)$; b) $f_X(x)$; c) $f_Y(y)$.

d) Are X and Y independent? Give a reason for your answer.

5. Let X be exponentially distributed with rate λ, independent of Y, which is exponentially distributed with rate μ. Find $P(X \geq 3Y)$.

6. Let X and Y have joint density

$$f(x, y) = \begin{cases} 90(y - x)^8 & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

a) Find $P(Y > 2X)$. b) Find the marginal density of X.

c) Fill in the blanks (explain briefly):
The joint density f above is the joint density of the ________ and ________ of ten independent uniform (0, 1) random variables.

7. Two points are picked independently and uniformly at random from the region inside a circle. Let R_1 and R_2 be the distances of these points from the center of the circle. Find $P(R_2 \leq R_1/2)$.

8. Random variables X and Y have joint density

$$f_{X,Y}(x, y) = \begin{cases} c(y^2 - x^2)e^{-y} & -y \leq x \leq y, \quad y > 0 \\ 0 & \text{otherwise} \end{cases}$$

Here c is a constant.

a) Show that Y has a gamma density, and hence deduce that $c = 1/8$.

b) Find the density of $4Y^3$.

c) Explain why $E(|X|)$ is at most 4.

9. **Minimum and maximum of two independent exponentials.** Let $X = \min(S, T)$ and $Y = \max(S, T)$ for independent exponential(λ) variables S and T. Let $Z = Y - X$.

a) Find the joint density of X and Y. Are X and Y independent?

b) Find the joint density of X and Z. Are X and Z independent?

c) Identify the marginal distributions of X and Z.

10. Minimum and maximum of n independent exponentials. Let $X_1, X_2, \ldots, X_n$ be independent, each with exponential (λ) distribution. Let $V = \min(X_1, X_2, \ldots, X_n)$ and $W = \max(X_1, X_2, \ldots, X_n)$. Find the joint density of V and W.

11. Suppose X and Y are independent random variables such that X has uniform $(0, 1)$ distribution, Y has exponential distribution with mean 1. Calculate:

a) $E(X+Y)$; b) $E(XY)$; c) $E[(X-Y)^2]$; d) $E(X^2 e^{2Y})$.

12. Let T_1 and T_5 be the times of the first and fifth arrivals in a Poisson process with rate λ, as in Section 4.2. Find the joint density of T_1 and T_5.

13. Uniform spacings. Let $X = \min(U, V)$ and $Y = \max(U, V)$ for independent uniform(0, 1) variables U and V. Find the distributions of

a) X; b) $1 - Y$; c) $Y - X$.

14. Let U_1, U_2, U_3, U_4, U_5 be independent, each with uniform distribution on $(0, 1)$. Let R be the distance between the minimum and the maximum of the U_i's. Find

a) $E(R)$;

b) the joint density of the minimum and maximum of the U_i's;

c) $P(R > 0.5)$

15. C.d.f.'s in two dimensions. The *cumulative joint distribution function* of random variables X and Y is the function of x and y defined by $F(x, y) = P(X \le x, Y \le y)$.

a) Find a formula in terms of $F(x, y)$ for $P(a < X \le b,\ c < Y \le d)$.

b) For X and Y with joint density $f(x, y)$, express $F(x, y)$ in terms of f.

c) For X and Y with joint density $f(x, y)$, express $f(x, y)$ in terms of F.

These are analogs of formulae of Section 4.5 for cumulative distribution functions in one dimension. They are not used much, as there are few joint distributions for which there is an explicit formula for $F(x, y)$. But here are two examples.

d) Find $F(x, y)$ in terms of the marginal c.d.f's for independent X and Y.

e) Find $F(x, y)$ for X the minimum and Y the maximum of n independent uniform $(0, 1)$ variables, and $0 < x < y < 1$. Deduce the joint density of X and Y.

16. Suppose X_1, X_2, X_3 are independent exponential random variables with parameters $\lambda_1, \lambda_2, \lambda_3$ respectively. Evaluate $P(X_1 < X_2 < X_3)$.

17. Let (X, Y) be picked uniformly from the unit disc $R^2 \le 1$, where $R^2 = X^2 + Y^2$. Find:

a) the joint density of R and X;

b) repeat a) for a point (X, Y, Z) picked at random from inside the unit sphere $R^2 \le 1$, where now $R^2 = X^2 + Y^2 + Z^2$.

18. Suppose X_1, X_2 are independent random variables with the same density function.

a) Evaluate $P(X_1 < X_2)$.

b) Continuing, suppose X_1, X_2, X_3 are independent random variables with the same density function. Evaluate $P(X_{i_1} < X_{i_2} < X_{i_3})$ where (i_1, i_2, i_3) is a given permutation of $(1, 2, 3)$.

19. Let Lat be the latitude, Lon the longitude of the point of impact of the next meteorite that strikes the Earth's surface. Measure Lat in degrees from $-90°$ (South Pole) to $+90°$ (North Pole), and measure Lon similarly from $-180°$ to $+180°$. Assuming the point of impact is uniformly distributed over the Earth's surface, find

a) the density of Lon; b) the density of Lat;

c) the joint density of Lat and Lon. d) Are Lat and Lon independent?

20. Let X and Y be independent and uniform $(0, 1)$ and let $R = \sqrt{X^2 + Y^2}$. Show that:

a) $$f_R(r) = \begin{cases} \frac{\pi}{2} r & 0 \le r \le 1 \\ 2r\left[\frac{\pi}{4} - \arccos(1/r)\right] & 1 \le r \le \sqrt{2} \end{cases}$$

b) $$F_R(r) = \begin{cases} \frac{1}{4}\pi r^2 & 0 \le r \le 1 \\ \sqrt{r^2 - 1} + \left[\frac{\pi}{4} - \arccos(1/r)\right] r^2 & 1 \le r \le \sqrt{2} \end{cases}$$

c) Show without explicitly calculating $E(R)$ that

$$\sqrt{\frac{1}{2}} < E(R) < \sqrt{\frac{2}{3}}$$

d) (*Hard.*) Show that $E(R) \approx 0.765$.

21. Suppose two points are picked at random from the unit square. Let D be the distance between them. The main point of this problem is to find $E(D)$. This is hard to do exactly by calculus. But some information about $E(D)$ can be obtained as follows.

a) It is intuitively clear that $E(D)$ must be greater than $E(D_{\text{center}})$, where D_{center} is the distance from one point picked at random to the center of the square, and less than $E(D_{\text{corner}})$, the expected distance of one point from a particular corner of the square. Assuming this to be the case, find the values of these bounds on $E(D)$ using the results of Exercise 20.

b) Compute $E(D^2)$ exactly.

c) Deduce from b) a better upper bound for $E(D)$.

d) Computer simulation of 10,000 pairs of points gave mean distance 0.5197, and mean square distance 0.3310. Use these results to find an approximate 95% confidence interval for the unknown value of $E(D)$.

5.3 Independent Normal Variables

The most important properties of the normal distribution involve two or more independent normal variables. Suppose first that X and Y are independent, each with standard normal density function

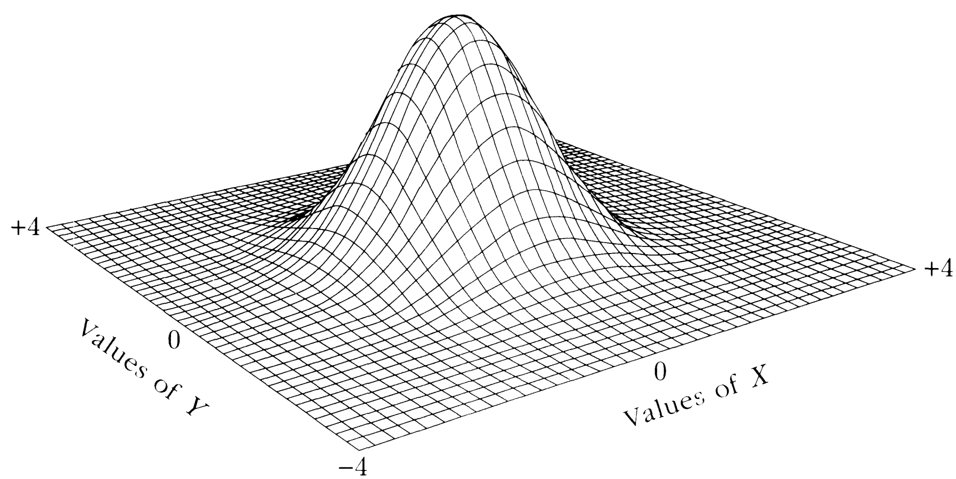
(a) $$\phi(z) = ce^{-\frac{1}{2}z^2} \quad \text{where the formula} \quad c = \frac{1}{\sqrt{2\pi}}$$

taken for granted up to now, will be verified in this section. The joint density of X and Y is given by

(b) $$f(x, y) = \phi(x)\phi(y) = c^2 e^{-\frac{1}{2}(x^2+y^2)}$$

The key property of this joint density is that it is a function of $r^2 = x^2 + y^2$, where r is the radial distance from the origin of the point (x, y). This makes the graph of this joint density a round bell-shaped surface over the (x, y) plane, with cross sections proportional to the standard normal curve.

FIGURE 1. Perspective plot of the joint density of X and Y.



The rotational symmetry of this bivariate distribution obtained from two independent normal variables is a very special property. It can be shown that this property distinguishes the normal distribution from all other probability distributions on the line. And this rotational symmetry is the key to understanding several important properties of the normal distribution, now considered in turn.

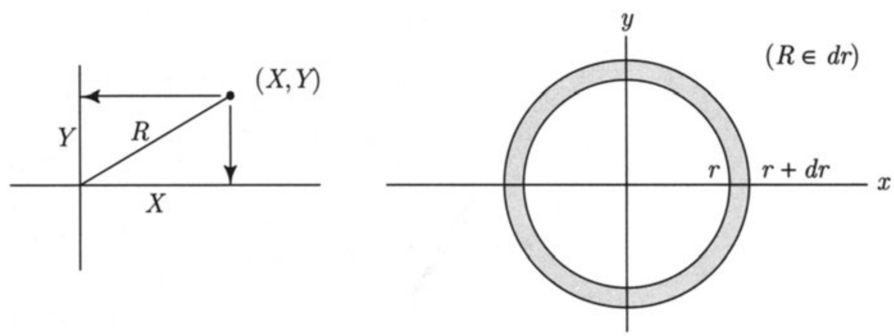
Evaluation of the Constant of Integration

The value of the constant c in the normal density (a) is found as a byproduct of calculating the distribution of the random variable

$$R = \sqrt{X^2 + Y^2}$$

which is the distance from the origin of a random point (X, Y) with joint density $\phi(x)\phi(y)$.

FIGURE 2. Geometry of X, Y, and R.



The event $(R \in dr)$ corresponds to (X, Y) falling in an annulus of infinitesimal width dr, radius r, circumference $2\pi r$, and area $2\pi r\, dr$, as in Figure 2. And $P(R \in dr)$ is the volume over this infinitesimal annulus beneath the joint density. But on the annulus the joint density has nearly constant value

$$\phi(x)\phi(y) = c^2 e^{-\frac{1}{2}(x^2+y^2)} = c^2 e^{-\frac{1}{2}r^2}$$

so the volume in question is just this nearly constant value times the area of the annulus. Thus

$$P(R \in dr) = 2\pi\, r\, dr\, c^2 e^{-\frac{1}{2}r^2} \qquad (r > 0)$$

This shows that R has probability density function

$$f_R(r) = 2\pi r c^2 e^{-\frac{1}{2}r^2}$$

The integral of this density from 0 to ∞ must be 1:

$$1 = \int_0^\infty 2\pi r c^2 e^{-\frac{1}{2}r^2}\, dr = -2\pi c^2 e^{-\frac{1}{2}r^2}\Big|_0^\infty = 2\pi c^2$$

This makes

$$2\pi c^2 = 1 \qquad \text{and} \qquad c = 1/\sqrt{2\pi}$$

So the constant of integration in the normal density involves π, due to the fact that the joint density of two independent standard normal variables is constant on circles centered at the origin.

The distribution of R appearing here, with density function

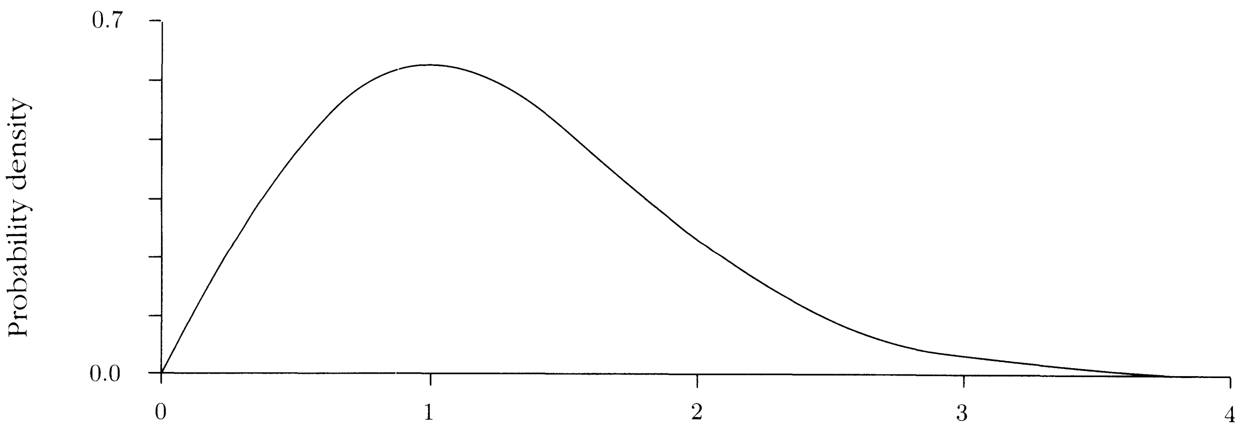
(c1) $$f_R(r) = r\,e^{-\frac{1}{2}r^2} \qquad (r > 0)$$

and c.d.f.

(c2) $$F_R(r) = \int_0^r se^{-\frac{1}{2}s^2}ds = 1 - e^{-\frac{1}{2}r^2} \qquad (r > 0)$$

is called the *Rayleigh distribution.*

FIGURE 3. Density of the Rayleigh distribution of R.



Calculating the Variance of the Standard Normal Distribution

Since $E(X) = 0$ by symmetry, the variance of a standard normal random variable X is

$$\sigma^2 = E(X^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{1}{2}x^2} dx$$

This integral can be reduced by an integration by parts to the integral of the standard normal density (exercise). But two independent standard normal variables X and Y

can also be used to show that $\sigma^2 = 1$. This, too, involves the radial random variable R. Because $R^2 = X^2 + Y^2$,

$$E(R^2) = E(X^2) + E(Y^2) = 2E(X^2)$$

using the fact that X and Y have the same distribution. So

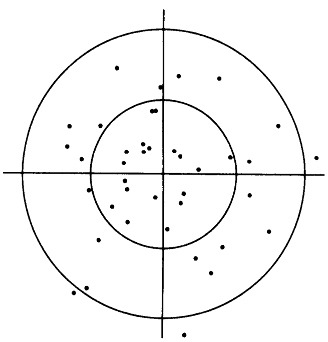
$$\sigma^2 = E(X^2) = \frac{1}{2}E(R^2)$$

But $S = R^2$ has density given by the change of variable formula

$$\begin{aligned} f_S(s) &= f_R(r) \Big/ \frac{ds}{dr} \qquad (s = r^2 > 0) \\ &= re^{-\frac{1}{2}r^2}/2r \qquad (s = r^2 > 0) \qquad \text{by (c1)} \\ &= \frac{1}{2}e^{-\frac{1}{2}s} \qquad (s > 0) \end{aligned}$$

Since this is the exponential density with parameter $\lambda = 1/2$,

$$E(R^2) = E(S) = 1/\lambda = 2 \qquad \text{so} \quad \sigma^2 = 1$$



Example 1. **Shots at a target.**

An expert marksman firing at a target produces a random scatter of shots which is roughly symmetrically distributed about the center of the bull's eye, with approximately 50% of the shots in the bull's eye, as in the diagram.

Problem 1. What is the approximate fraction of shots inside a circle with the same center as the bull's eye, but twice the radius?

Solution. Suppose that the marksman's shots are distributed approximately like (X, Y), where X and Y are independent normal random variables with mean 0 and variance σ^2. This would give such a symmetric distribution. By measuring distances in standard units, that is, relative to σ, we may as well assume $\sigma = 1$. Then the formulae obtained above for the distribution of $R = \sqrt{X^2 + Y^2}$ apply directly. Let r denote the radius of the bull's eye, measured in standard units. Using the normal approximation, the probability of each shot hitting the bull's eye would be

$$F_R(r) = 1 - e^{-\frac{1}{2}r^2}$$

from formula (c2) on page 359. Estimating this probability as 50% from the empirical data gives

$$e^{-\frac{1}{2}r^2} = 1/2 \qquad \text{so} \qquad r = \sqrt{2\log(2)} = 1.177\ldots \text{standard units}$$

Similarly, the fraction of shots inside a circle of twice the radius of the bull's eye should be approximately

$$F_R(2r) = 1 - e^{-\frac{1}{2}(2r)^2} = 1 - (e^{-\frac{1}{2}r^2})^4 = 1 - (1/2)^4 = \frac{15}{16} = 0.9375$$

Problem 2. What is the approximate average distance of the marksman's shots from the center of the bull's eye?

Solution. Using the law of large numbers, this average should be approximately

$$\begin{aligned} E(R) &= \int_0^\infty r f_R(r)dr = \int_0^\infty r^2 e^{-\frac{1}{2}r^2} dr \qquad \text{by (c1) on page 359} \\ &= \frac{1}{2}\int_{-\infty}^\infty x^2 e^{-\frac{1}{2}x^2} dx \qquad \text{by symmetry} \\ &= \frac{\sqrt{2\pi}}{2}\int_{-\infty}^\infty x^2 \phi(x) dx \qquad \text{by definition of } \phi(x) \\ &= \sqrt{\frac{\pi}{2}} \qquad \text{because standard normal variance is 1} \\ &\approx 1.253 \text{ standard units} \\ &\approx 1.253/1.177 = 1.065 \text{ times the bull's eye radius } r \end{aligned}$$

Linear Combinations and Rotations

Linear combinations of independent normal variables are always normally distributed. This important fact is another consequence of the rotational symmetry of the joint distribution of independent standard normal random variables X and Y. To see why, let X_θ be the first coordinate of (X, Y) relative to new coordinate axes set up at angle θ relative to the original X and Y axes, as in Figure 4.

As the diagram shows,

$$X_\theta = X\cos\theta + Y\sin\theta$$

But due to the rotational symmetry of the joint distribution, it is clear without calculation that the probability distribution of X_θ must be the same as that of X, namely, standard normal, no matter what the angle θ of rotation. For example, the event $x \le X_\theta \le x + \Delta x$ corresponding to (X, Y), falling in the area shaded in the left diagram of Figure 5, must have the same probability as the event $x \le X \le x + \Delta x$ corresponding to (X, Y), falling in the area shaded in the right diagram, because the shape of the bivariate normal density is the same over the two shaded regions. So,

$$P(x \le X_\theta \le x + \Delta x) = P(x \le X \le x + \Delta x)$$

FIGURE 4. Projection X_θ onto axis at angle θ to X-axis: $X_\theta = X\ \cos\theta + Y\ \sin\theta$.

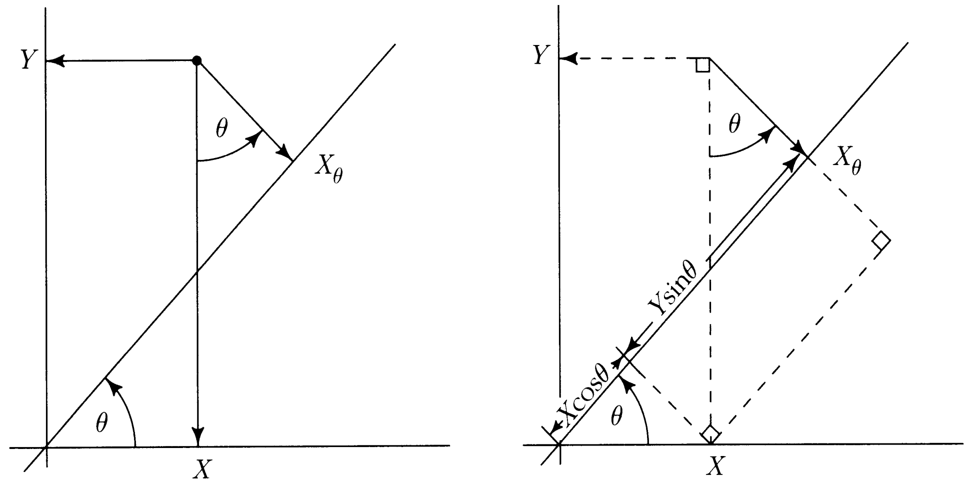
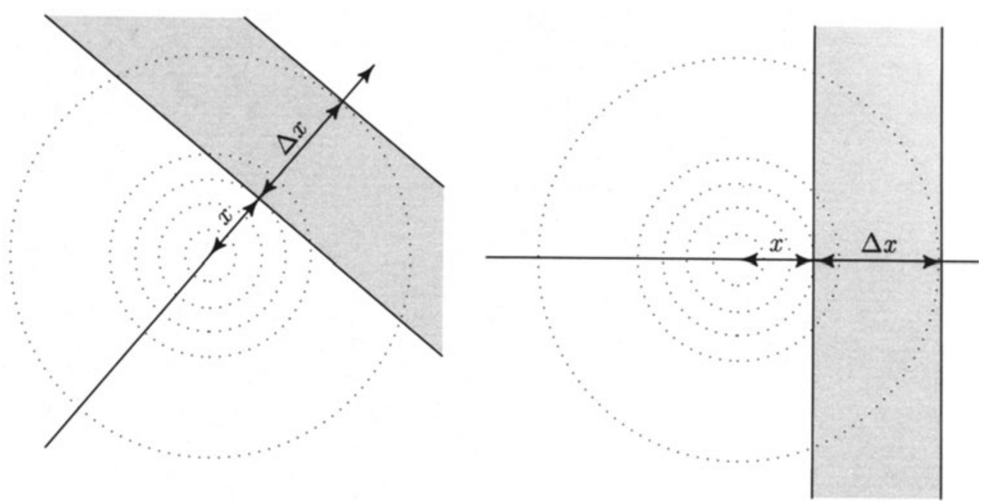


FIGURE 5. Events $(x \le X_\theta \le x + \Delta x)$ and $(x \le X \le x + \Delta x)$. Rotational symmetry of the joint density implies these two events have the same probability.



for every x and Δx. This shows that X_θ has normal $(0,1)$ distribution, for every θ. Since $\cos\theta$ and $\sin\theta$ may be arbitrary numbers α and β, subject only to the constraint that $\alpha^2+\beta^2=1$, the rotational symmetry of the joint distribution of two independent normal variables X and Y implies that:

(d) *If X and Y are two independent normal $(0,1)$ random variables, then $\alpha X+\beta Y$ has normal $(0,1)$ distribution for all α and β with $\alpha^2+\beta^2=1$.*

In particular, taking $\alpha=\beta=1/\sqrt{2}$, corresponding to rotation by 45°:

(e) *If X and Y are independent normal $(0,1)$ random variables, then $(X+Y)/\sqrt{2}$ has normal $(0,1)$ distribution.*

If Z has normal $(0,1)$ distribution, then σZ has normal $(0,\sigma^2)$ distribution. Taking $\sigma=\sqrt{2}$, (e) implies:

(f) *If X and Y are independent normal $(0,1)$ random variables, then $X+Y$ has normal $(0,2)$ distribution.*

This argument extends to give the following general conclusion, which includes (d), (e), and (f), as special cases.

Sums of Independent Normal Variables

If X and Y are independent with normal (λ,σ^2) and normal (μ,τ^2) distributions, then $X+Y$ has normal $(\lambda+\mu,\sigma^2+\tau^2)$ distribution.

Proof. Recall that X has normal (λ,σ^2) distribution if and only if $(X-\lambda)/\sigma$ has normal $(0,1)$ distribution. Transform all the variables to standard units by letting

$$U=(X-\lambda)/\sigma \quad \text{and} \quad V=(Y-\mu)/\tau \quad \text{and} \quad W=\frac{X+Y-(\lambda+\mu)}{\sqrt{\sigma^2+\tau^2}}$$

Then U and V are independent normal $(0,1)$ random variables. By algebra,

$$W=\alpha U+\beta V \quad \text{where} \quad \alpha^2=\frac{\sigma^2}{\sigma^2+\tau^2} \quad \text{and} \quad \beta^2=\frac{\tau^2}{\sigma^2+\tau^2} \quad \text{so} \quad \alpha^2+\beta^2=1$$

Apply (d) above with (U,V) instead of (X,Y) to deduce that W has normal $(0,1)$ distribution. So $X+Y=(\lambda+\mu)+\sqrt{\sigma^2+\tau^2}W$ has normal $(\lambda+\mu,\sigma^2+\tau^2)$ distribution.
□

Several Independent Normal Variables

The result that the sum of two independent normal variables is normal extends to sums and linear combinations of several independent normal random variables, by repeated applications of the result for two variables. For example, if $X_1, \ldots, X_n$ are independent and normal $(0, 1)$, then $X_1 + \cdots + X_n$ has normal $(0, n)$ distribution, with standard deviation $\sqrt{n}$.

Example 2. **Linear combinations of normals.**

For $\sigma = 1, 2, 3$ suppose X_σ has normal $(0, \sigma^2)$ distribution, and these three random variables are independent.

Problem 1. Find $P(X_1 + X_2 + X_3 < 4)$.

Solution. Let $S = X_1 + X_2 + X_3$. Then S has normal $(0, 1^2 + 2^2 + 3^2)$ distribution, and if $Z = S/\sqrt{14}$ is S standardized, the problem is just to find

$$P(S < 4) = P(Z < 4/\sqrt{14}) = \Phi(4/\sqrt{14}) \approx 0.857$$

Problem 2. Find $P(4X_1 - 10 < X_2 + X_3)$.

Solution. Rearranging the statement of the inequality shows this is the same as

$$P(4X_1 - X_2 - X_3 < 10) = P(L < 10) \qquad \text{where } L = 4X_1 - X_2 - X_3$$

Since the linear combination L has normal distribution with mean 0 and variance $4^2 \times 1^2 + (-1)^2 \times 2^2 + (-1)^2 \times 3^2 = 29$, the probability is

$$P(L < 10) = \Phi(10/\sqrt{29}) \approx 0.968$$

The Chi-Square Distribution

By the same calculation as in two dimensions, the joint density of n independent normal variables at every point on the sphere of radius r in n-dimensional space is $(1/\sqrt{2\pi})^n \exp(-\frac{1}{2}r^2)$. This joint density is symmetric with respect to arbitrary rotations of the coordinates in n-dimensional space, or *spherically symmetric*. So a cloud of points (or a galaxy of stars), in ordinary 3-dimensional space, with approximately independent normally distributed coordinates with common variance, appears spherical when viewed at a distance, from any perspective. For independent standard normal Z_i let

$$R_n = \sqrt{Z_1^2 + \cdots + Z_n^2}$$

denote the distance of $(Z_1, \ldots, Z_n)$ from the origin in n-dimensional space. The n-dimensional volume of a thin spherical shell of thickness dr at radius r is $c_n r^{n-1}\, dr$

where c_n is the $(n-1)$-dimensional volume of the "surface" of a sphere of radius 1 in n dimensions. (For $n = 3$, $c_3 = 4\pi$, by the formula $4\pi r^2$ for the surface area of a sphere of radius r in 3 dimensions.) The same argument used in two dimensions shows that

$$P(R_n \in dr) = c_n\, r^{n-1} \left(1/\sqrt{2\pi}\right)^n e^{-\frac{1}{2}r^2} dr \qquad (r > 0) \tag{1}$$

A change of variable allows the constant c_n to be evaluated by recognizing that the density of $R_n^2 = Z_1^2 + \cdots + Z_n^2$ is the gamma $(n/2, 1/2)$ density introduced in Section 4.2:

$$f_{R_n^2}(t) = (2^{n/2}\Gamma(n/2))^{-1} t^{(n/2)-1} e^{-t/2} \qquad (t > 0) \tag{2}$$

Exercise 15 and Chapter 5 Review Exercise 26 give formulae for c_n and $\Gamma(n/2)$.

Statisticians call this gamma $(n/2, 1/2)$ distribution of R_n^2 the ***chi-square*** distribution with n ***degrees of freedom***. The chi-square distribution provides a useful test of ***goodness of fit***, that is, how well data from an empirical distribution of n observations conform to the model of random sampling from a particular theoretical distribution. If there are only two categories, say success and failure, the model of independent trials with probability p of success is tested using the normal approximation to the binomial distribution. But for data in several categories the problem is how to combine the tests for different categories in a reasonable way. This problem was solved as follows by the statistician Karl Pearson (1857–1936). For a finite number of categories m, let N_i denote the number of results in category i. Under the hypothesis that the N_i are counting results of independent trials with probability p_i for category i on each trial, it turns out that no matter what the probabilities p_i, for large enough n the so-called ***chi-square statistic***

$$\sum_{i=1}^{m} \frac{(N_i - n\,p_i)^2}{n\,p_i}$$

that is the sum over categories of (observed $-$ expected)^2/expected, has distribution that is approximately chi-square with $m-1$ degrees of freedom. In statistical jargon, a value of the statistic higher than the 95th percentile point on the chi-square distribution with $m-1$ degrees of freedom would "reject the hypothesis at the 5% level". Unusually small values of the chi-square statistic are sometimes taken as evidence to suggest that an observer fudged the data to suit the hypothesis. The exact joint distribution of the N_i under the hypothesis of randomness is multinomial with parameters n and $p_1, \ldots, p_m$. The above result can be derived from a multivariate form of the normal approximation to the binomial. The joint distribution of $N_1, \ldots, N_m$ is essentially $m-1$ dimensional due to the constraint $N_1 + \cdots + N_m = n$. This is why the relevant chi-square distribution has $m-1$ degrees of freedom.

For tables of the chi-square distribution, and similar chi-square tests of other hypotheses such as independence, consult a statistics book. The mean, standard deviation and skewness of the chi-square distribution of R_n^2 with n degrees of freedom are easily calculated (Exercise 15):

$$E(R_n^2) = n, \qquad SD(R_n^2) = \sqrt{2n} \qquad \text{and} \qquad \text{Skewness}(R_n^2) = 4/\sqrt{2n}$$

For large n the chi-square distribution is approximately normal, by the central limit theorem. Because the skewness is quite large even for moderate values of n, the normal approximation with skewness correction gives the better approximation

$$P(R_n^2 \le x) \approx \Phi(z) - \frac{\sqrt{2}}{3\sqrt{n}}(z^2 - 1)\phi(z) \quad \text{where} \quad z = (x - n)/\sqrt{2n} \text{ and } x > 0$$

TABLE 1. Distribution of radial distance in three dimensions. The probability that a point with independent standard normal coordinates in three dimensions lies inside a sphere of radius r, that is, $P(R_3 \le r) = P(R_3^3 \le r^2)$, was obtained by numerical integration of the density. These probabilities are shown along with their approximations obtained using the skew–normal approximation to the chi-square (3) distribution of R_3^2. The approximations are surprisingly good considering the small value of n.

radius r	1	2	3	4
probability $P(R_3 \le r)$	0.199	0.739	0.971	0.999
skew-normal approximation	0.233	0.741	0.966	1.000

Exercises 5.3

1. Continuing Example 1, calculate the following, where all distances are measured in standard units:

a) the probability of a shot falling inside a circle of radius 1/2;

b) the probability of a shot falling in the region of the positive quadrant between radii 1 and 2;

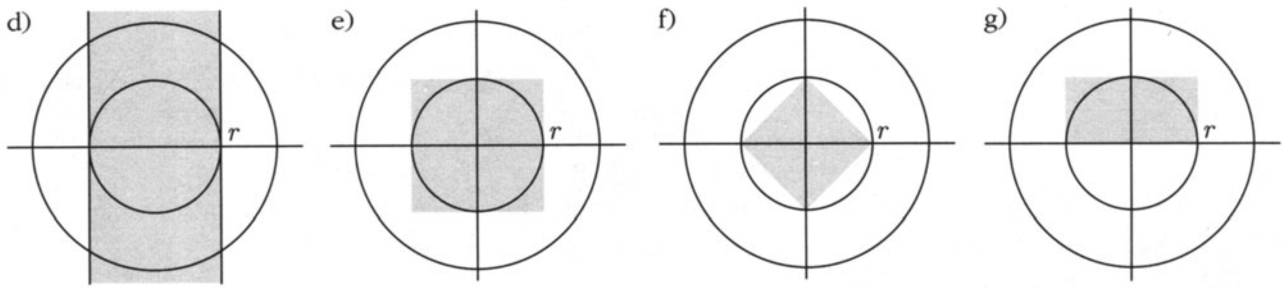
c) the approximate average absolute distance of the shots from the horizontal line through the center of the bull's eye;

d) the probability that a shot hit within distance r of the vertical axis through the center (r = radius of bull's eye in standard units);

e) the probability of hitting a square touching the outside of the bull's eye;

f) the probability of hitting a square touching the inside of the bull's eye;

g) the probability of hitting a rectangle of sides r and $2r$ positioned as shown relative to the bull's eye.



2. Let X and Y be independent random variables, with $E(X) = 1$, $E(Y) = 2$, $Var(X) = 3$, and $Var(Y) = 4$.

a) Find $E(10X^2 + 8Y^2 - XY + 8X + 5Y - 1)$.

b) Assuming all variables are normally distributed, find $P(2X > 3Y - 5)$.

3. W, X, Y and Z are independent standard normal random variables. Find (no integrations are necessary!)

a) $P(W + X > Y + Z + 1)$; b) $P(4X + 3Y < Z + W)$;

c) $E(4X + 3Y - 2Z^2 - W^2 + 8)$; d) $SD(3Z - 2X + Y + 15)$.

4. Suppose the true weight of a standard weight is 10 grams. It is weighed twice independently. Suppose that the first measurement is a normal random variable X with $E(X) = 10$ g and $SD(X) = 0.2$ g, and that the second measurement is a normal random variable Y with $E(Y) = 10$ g and $SD(Y) = 0.2$ g.

a) Compute the probability that the second measurement is closer to 10 g than the first measurement.

b) Compute the probability that the second measurement is smaller than the first, but not by more than 0.2 g.

5. Let X and Y be independent and normally distributed, X with mean 0 and variance 1, Y with mean 1. Suppose $P(X > Y) = 1/3$. Find the standard deviation of Y.

6. Let X and Y be independent standard normal variables. Find:

a) $P(3X + 2Y > 5)$; b) $P(\min(X, Y) < 1)$;

c) $P(|\min(X, Y)| < 1)$; d) $P(\min(X, Y) > \max(X, Y) - 1)$.

7. Suppose the AC Transit bus is scheduled to arrive at my corner at 8:10 A.M., but its actual arrival time is a normal random variable with mean 8:10 A.M., and standard deviation 40 seconds. Suppose I try to arrive at the corner at 8:09, but my arrival time is actually normally distributed with mean 8:09 A.M., and standard deviation 30 seconds.

a) What percentage of the time do I arrive at the corner before the bus is scheduled to arrive?

b) What percentage of the time do I arrive at the corner before the bus does?

c) If I arrive at the stop at 8:09 A.M. and the bus still hasn't come by 8:12 A.M., what is the probability that I have already missed it?

(State your assumptions carefully.)

8. Peter and Paul agree to meet at a restaurant at noon. Peter arrives at a time normally distributed with mean 12:00 noon, and standard deviation 5 minutes. Paul arrives at a time normally distributed with mean 12:02 P.M., and standard deviation 3 minutes. Assuming the two arrival times are independent, find the chance that

a) Peter arrives before Paul; b) both men arrive within 3 minutes of noon;

c) the two men arrive within 3 minutes of each other.

9. Suppose heights in a large population are approximately normally distributed with a mean of 5 feet 10 inches and an SD of 2 inches. Suppose a group of 100 people is picked at random from this population.

a) What is the probability that the tallest person in this group is over 6 feet 4 inches tall?

b) What is the probability that the average height of people in the group is over 5 feet 10.5 inches?

c) Suppose instead that the distribution of heights in the population was not normal, but some other distribution with the given mean and SD. To which of the problems a) and b) would the answer still be approximately the same? Explain carefully.

10. In a large corporation, people over age thirty have an annual income whose distribution can be approximated by a normal distribution with mean $60,000 and standard deviation $10,000. The incomes of those under age thirty are also approximately normal, but with mean $40,000 and standard deviation $10,000.

a) Two people are selected at random from those over age thirty. What is the chance that the average of their two incomes is over $65,000?

b) One person is selected at random from those over thirty, and independently, one person is selected at random from those under thirty. What is the chance that the younger's income exceeds the older's?

c) What is the chance that the smaller of the two incomes in b) exceeds $50,000?

11. **Einstein's model for Brownian motion.** Suppose that the X coordinate of a particle performing Brownian motion has normal distribution with mean 0 and variance σ^2 at time 1. Let X_t be the X displacement after time t. Assume the displacement over any time interval has a normal distribution with parameters depending only on the length of the interval, and that displacements over disjoint time intervals are independent.

a) Find the distribution of X_t.

b) Let (X_t, Y_t) represent the position at time t of a particle moving in two dimensions. Assume that X_t and Y_t are independent Brownian motions starting at 0 at time $t = 0$. Find the distribution of $R_t = \sqrt{X_t^2 + Y_t^2}$, and give the mean and standard deviation in terms of σ and t.

c) Suppose a particle performing Brownian motion (X_t, Y_t) as in b) has an X coordinate after one second which has mean 0 and standard deviation one millimeter (mm). Calculate the probability that the particle is more than 2 mm from the point $(0, 0)$ after one second.

12. Suppose two shots are fired at a target. Assume each shot hits with independent normally distributed coordinates, with the same means and equal unit variances.

a) Find the mean of the distance between the points where the two shots strike.

b) Find the variance of the same random variable.

13. Independence of radial and angular parts. Let X and Y be independent normal $(0, \sigma^2)$ random variables. Let (R, Θ) be (X, Y) in polar coordinates, so $X = R\cos\Theta$, $Y = R\sin\Theta$.

a) Show that R and Θ are independent, and that Θ has uniform $(0, 2\pi)$ distribution.

b) Let R and Θ now be arbitrary random variables such that R/σ has the Rayleigh distribution (c1), Θ has uniform $(0, 2\pi)$ distribution, and R and Θ are independent. Explain why the random variables $X = R\cos\Theta$ and $Y = R\sin\Theta$ must be independent normal $(0, \sigma^2)$.

c) Find functions h and k such that if U and V are independent uniform $(0, 1)$ random variables, then $X = \sigma h(U)\cos[k(V)]$ and $Y = \sigma h(U)\sin[k(V)]$ are independent normal $(0, \sigma^2)$. [This gives a means of simulating normal random variables using a computer random number generator. Try generating a random scatter of independent bivariate normally distributed pairs if you have random numbers available. It should look like the scatter in Example 1.]

14. Let X and Y be independent standard normal variables. Suppose they are transformed into polar coordinates, $X = R\cos\Theta$ and $Y = R\sin\Theta$ with $0 < \Theta < 2\pi$ and $0 < R < \infty$, as in Exercise 13.

a) Derive the distribution of $2\Theta \; mod \; 2\pi$. [The quantity $x \; mod \; a$ denotes the remainder when x is divided by a.]

b) Derive the joint distribution of $R\cos 2\Theta$ and $R\sin 2\Theta$.

c) Show that both

$$\frac{2XY}{\sqrt{X^2+Y^2}} \quad \text{and} \quad \frac{X^2-Y^2}{\sqrt{X^2+Y^2}}$$

have the standard normal distribution. Are they independent?

15. Chi-square distributions. These are the special case of half-integer gamma distributions which come from sums of squares of independent standard normal variables. Show:

a) If Z has standard normal distribution, then Z^2 has gamma $(1/2, 1/2)$ distribution, and $\Gamma(1/2) = \sqrt{\pi}$.

b) If n is an odd integer, then $\quad \Gamma(n/2) = \dfrac{\sqrt{\pi}(n-1)!}{2^{n-1}(\frac{n-1}{2})!}$

c) If X has normal $(0, \sigma^2)$ distribution, then X^2 has gamma $(1/2, 1/2\sigma^2)$ distribution.

d) If $Z_1, \dots, Z_n$ are independent standard normal random variables, then ${Z_1}^2 + \cdots + {Z_n}^2$ has gamma $(n/2, 1/2)$ distribution, also known as the *chi-square distribution with n degrees of freedom*, or chi-square (n) distribution.

e) If $Y_1, \dots, Y_n$ are independent chi-square random variables with $k_1, \dots, k_n$ degrees of freedom, respectively, then $Y_1 + \cdots + Y_n$ has chi-square $(k_1 + \cdots + k_n)$ distribution.

f) The mean, variance and skewness of the chi-square (n) distribution are as stated on page 366.

16. Poisson formula for the chi-square (2m) c.d.f. For $m = 1, 2, \dots$ let R_{2m}^2 have chi-square $(2m)$ distribution. Use the connection between the gamma distribution and the Poisson process to find formulae in terms of appropriate Poisson probabilities for:

a) the c.d.f. of R_{2m}^2; b) the c.d.f. of R_{2m}.

c) Check that your formulae agree with the formulae in the text for $m = 1$. Now make a table of $P(R_4 \leq r)$ for $r = 1, \dots 5$.

17. Skew-normal approximation to the chi-square distribution. Let R_n^2 have chi-square (n) distribution.

a) Find the approximation to $P(R_4 \leq r)$ for $r = 1, \dots 5$ obtained from the skew-normal approximation to the distribution of R_4^2. Compare to the exact results found in Exercise 16.

b) Find both the plain normal approximation and the skew-normal approximation to $P(R_{10}^2 \leq 9.34) = 0.500$. Which approximation is better?

18. Suppose a large number n identical molecules are distributed independently at random in a box with sides of 1 centimeter. Let X, Y, Z be the coordinates in centimeters of the center of mass of the n molecules at a particular instant, relative to the center of the box. Thus,

$$X = (X_1 + \cdots + X_n)/n$$

and so on, where (X_i, Y_i, Z_i) are the coordinates of the ith molecule in centimeters. Let $R = \sqrt{X^2 + Y^2 + Z^2}$ be the distance of the center of mass of the n molecules from the center of the box. Given that for the chi-square distribution with 3 degrees of freedom the 95th percentile is at 7.82, find approximately the value of r such that R is 95% sure to be smaller than r.

5.4 Operations (Optional)

Many applications require calculation of the distribution of some random variable Z which is a function of X and Y, where X and Y are random variables with some joint density $f(x, y)$. Here the function of X and Y might be, for example, $X + Y$, XY, X/Y, $\max(X, Y)$, $\min(X, Y)$, or $\sqrt{X^2 + Y^2}$. This kind of calculation has been done in special cases in previous sections. For example, maxima and minima in Section 4.5, sums and $\sqrt{X^2 + Y^2}$ for normal variables in Section 5.3. This section gives a general technique for computing such distributions by integration.

Calculating the whole distribution of a function of X and Y can sometimes be tedious. So keep in mind that for some purposes it may be enough to calculate an expectation. The expectation of a function of X and Y can always be expressed as an integral with respect to the density of (X, Y). For example, for the product XY,

$$\begin{aligned} E(XY) &= \iint xyf(x,y)dx\,dy \\ &= E(X)E(Y) \qquad \text{if } X \text{ and } Y \text{ are independent} \end{aligned}$$

despite the fact that there are very few examples where the whole distribution of a product of independent random variables can be found explicitly.

One method of finding the distribution of $Z = g(X, Y)$ is to find the c.d.f. $P(Z \leq z)$ by integration of $f(x, y)$ over the region in the (x, y) plane where $g(x, y) \leq z$. Provided this integral can be evaluated fairly explicitly, the density of Z can then be found by differentiation of the c.d.f. Usually a quicker method of finding the distribution of Z is to anticipate that Z will have a density function f_Z, and to find this density $f_Z(z) = P(Z \in dz)/dz$ by integrating the joint density of X and Y over the subset $(Z \in dz)$ in the (X, Y) plane. This technique gives integral formulae for the density for the sum $X + Y$, for other linear combinations like $X - Y$, and for the product XY, and ratio X/Y. The formulae for sums and ratios will now be worked out in detail. Results for other operations are similar and left as exercises.

Distribution of Sums

A good deal has already been said on this topic. Recall the addition rule for expectation

$$E(X + Y) = E(X) + E(Y) \qquad \text{whatever the joint distribution of } X \text{ and } Y$$

the addition rule for variances in the case of independence, and the central limit theorem governing the asymptotic distribution for the sum of a large number of independent and identically distributed terms. Also, the exact distribution of sums has been computed in special cases by a variety of methods. The following table reviews some important examples:

Distribution of terms	Distribution of sum	See Section
n independent Bernoulli (p)	binomial (n, p)	2.1
independent Poisson (μ_i)	Poisson $(\Sigma\mu_i)$	3.5
independent normal (μ_i, σ_i^2)	normal $(\Sigma\mu_i, \Sigma\sigma_i^2)$	5.3
r independent geometric (p)	negative binomial (r, p)	3.4
r independent exponential (λ)	gamma (r, λ)	4.2

In the discrete case the distribution of the sum of random variables is determined by the formula

$$P(X+Y=z) = \sum_{\text{all } x} P(X=x, Y=z-x)$$

found in Section 4.1. The following display gives the corresponding formula for densities:

Density of $X + Y$

If (X, Y) has density $f(x, y)$ in the plane, then $X + Y$ has density on the line

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x, z-x)dx$$

Density Convolution Formula

If X and Y are independent, then

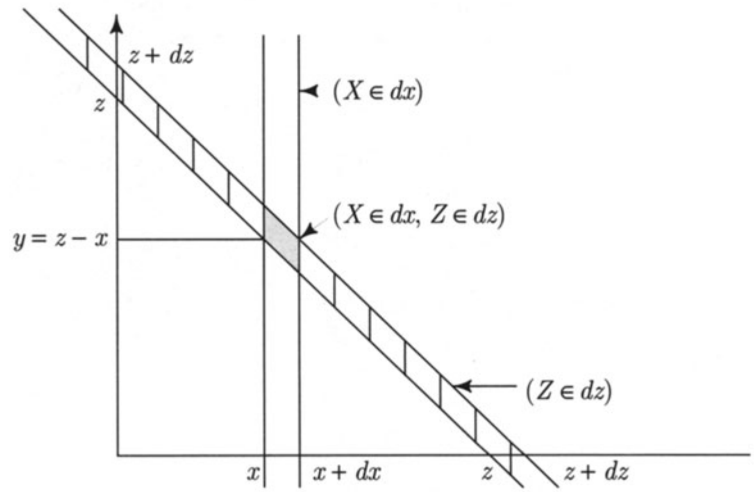
$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx$$

Note: If the random variables X and Y are non-negative, then the lower limit of integration in the convolution formula can be changed from $-\infty$ to 0, since $f_X(x) = 0$ for all $x < 0$, and the upper limit can be changed from ∞ to z, since $f_Y(z-x) = 0$ for $x > z$.

The convolution formula is the special case of the formula for the density of $X + Y$ when $f(x, y) = f_X(x) f_Y(y)$ by independence. This operation on probability density functions f_X and f_Y is called *convolution.* It leads to a new density, the density of the sum of random variables X and Y, assumed independent.

To avoid confusion about limits of integration in particular examples, sketch the subset of the plane where the joint density is strictly positive, and the line of integration corresponding to $X + Y = z$, as in examples below.

Derivation of the density of $X + Y$. Let $Z = X + Y$. The event $(Z \in dz)$ is shaded in the following diagram:



The event $(Z \in dz)$ can be broken up into vertical slices according to the values of X, as suggested by the vertical shading in the diagram. The heavily shaded parallelogram contained in the event $(Z \in dz)$ near the point $(x, z - x)$, represents the intersection of the events $(X \in dx)$ and $(Z \in dz)$, and has area $dx\,dz$. The probability density near this little parallelogram is $f(x, z - x)$, so

(a1) $$P(X \in dx, Z \in dz) = f(x, z - x)dx\,dz$$

This formula gives the joint density of X and Z. The marginal density of $Z = X + Y$ is therefore obtained by integrating out the x-variable

(a2) $$P(Z \in dz) = \left[\int_{-\infty}^{\infty} f(x, z - x)dx\right] dz$$

This gives the boxed formula for the density of $Z = X + Y$. Intuitively, you can think of (a2) as obtained by summing over infinitesimal parallelograms as in (a1).□

Example 1. Sums of independent exponential variables.

In Section 4.2 a Poisson process argument was used to show that the distribution of the sum of r independent exponential (λ) random variables is gamma (r, λ): If $f_{r,\lambda}(t)$ denotes the density of such a sum, then

$$f_{r,\lambda}(t) = \frac{1}{(r-1)!}\lambda^r t^{r-1} e^{-\lambda t} \qquad (t \geq 0)$$

This fact can also be derived using the convolution formula. Here is the calculation for $r = 2$.

Suppose T and U are independent, each exponentially distributed with rate λ. By independence, the joint density of T and U at (t, u) is

$$f(t,u) = f_T(t)f_U(u) = \lambda e^{-\lambda t}\lambda e^{-\lambda u} = \lambda^2 e^{-\lambda(t+u)} \qquad (t, u \geq 0)$$

Note how this joint density is a function of $t + u$. You can see the effect of this in Figure 1.

The density of $S = T + U$ at s is given by the convolution formula

$$\begin{aligned} f_S(s) &= \int_{-\infty}^{\infty} f_T(t)f_U(s-t)dt \\ &= \int_0^s f_T(t)f_U(s-t)dt \qquad \text{since } f_T(t) = 0 \text{ if } t < 0 \\ &\qquad\qquad\qquad\qquad\qquad \text{and} \quad f_U(s-t) = 0 \text{ if } t > s \\ &= \int_0^s \lambda e^{-\lambda t}\lambda e^{-\lambda(s-t)}dt \\ &= \int_0^s \lambda^2 e^{-\lambda s}dt \\ &= \lambda^2 s e^{-\lambda s} \qquad (s \geq 0) \end{aligned}$$

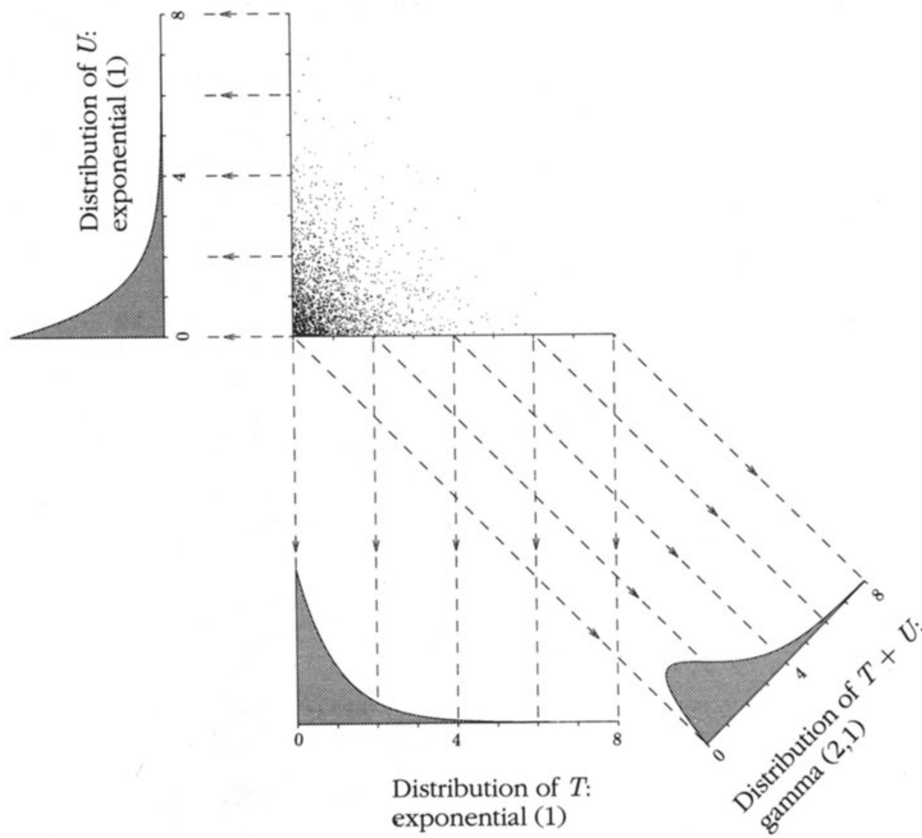
See Figure 1. For small s the factor of s makes the density grow linearly near zero. For large s the exponential factor $e^{-\lambda s}$ brings the density down to zero very rapidly.

Another way to derive this density is to argue infinitesimally: Let $s \geq 0$. The probability of $(S \in ds)$ is the integral of the joint density over the infinite strip $((t, u) : s \leq t + u \leq s + ds)$. We need only integrate over the (approximately) rectangular segment $((t, u) : s \leq t + u \leq s + ds,\ t \geq 0, u \geq 0)$, where the joint density is nonzero. This segment has length $\sqrt{2}s$ and width $ds/\sqrt{2}$, and the joint density has nearly constant value $\lambda^2 e^{-\lambda(t+u)} = \lambda^2 e^{-\lambda s}$ for points (t, u) in this segment; so the desired probability is

$$P(S \in ds) = \sqrt{2}s \cdot ds/\sqrt{2} \cdot \lambda^2 e^{-\lambda s} = \lambda^2 s e^{-\lambda s} ds \qquad (s \geq 0)$$

The fact that the sum of r independent exponential (λ) variables has gamma (r, λ) distribution can be derived from the convolution formula by mathematical induction on r.

FIGURE 1. **Distribution of the sum of two independent exponential variables.** Here is a random scatter of points suggesting the joint density of independent exponential variables T and U, along with graphs of the densities of T, U, and $S = T + U$.



Example 2. Sums of independent gamma variables.

Recall from Section 4.2 that the gamma (r, λ) distribution is defined for every real $r > 0$ by the density

$$f_{r,\lambda}(t) = \begin{cases} [\Gamma(r)]^{-1}\lambda^r t^{r-1}e^{-\lambda t} & t > 0 \\ 0 & t \le 0 \end{cases} \qquad \text{where} \quad \Gamma(r) = \int_0^\infty t^{r-1}e^{-t}dt$$

> If T_r and T_s are independent random variables with gamma (r, λ) and gamma (s, λ) distributions, respectively, then $T_r + T_s$ has gamma $(r + s, \lambda)$ distribution.

Proof for positive integers r and s. This case follows from the representation of a gamma variable as the sum of independent exponential variables. To see how, note first that the density of an independent sum $T_r + T_s$ is determined by the densities of T_r and T_s, by the convolution formula. So it is enough to derive the result for any convenient pair of independent random variables with gamma (r, λ) and gamma (s, λ) distributions. But the conclusion is obvious if we consider

$$T_r = W_1 + \cdots + W_r \quad \text{and} \quad T_s' = W_1' + \cdots + W_s'$$

defined by $r+s$ independent exponentials $W_1 \ldots, W_r, W_1', \ldots, W_s'$. Because then

$$T_r + T_s' = W_1 + \cdots + W_r + W_1' + \cdots + W_s'$$

is the sum of $r+s$ independent exponentials, with gamma $(r+s, \lambda)$ distribution. □

Proof for positive half-integers r and s. The case $r = n/2$ and $s = m/2$ for positive integers n and m can be derived almost the same way, using the result found in Section 5.3 that the gamma $(n/2, 1/2)$ distribution is the chi-square distribution of the sum of squares of n independent standard normal variables. Adding the sum of squares of n variables to the sum of squares of m variables gives the sum of squares of $n+m$ variables. Changing the rate parameter $1/2$ to a general λ is just a matter of multiplying of the chi-square variables by $1/(2\lambda)$. (See Exercises 5.3.15 and 4.4.2). □

Proof for general positive r and s. For $r > 0$, $s > 0$, let T_r and T_s be independent, with gamma (r, λ) and gamma (s, λ) distributions, and let $Z = T_r + T_s$. Then by the convolution formula

$$\begin{aligned}
f_Z(z) &= \int_0^z f_{T_r}(x) f_{T_s}(z-x)\,dx \\
&= \int_0^z \frac{1}{\Gamma(r)} \cdot \lambda^r x^{r-1} e^{-\lambda x} \cdot \frac{1}{\Gamma(s)} \lambda^s (z-x)^{s-1} e^{-\lambda(z-x)}\,dx \\
&= \int_0^z \frac{1}{\Gamma(r)\Gamma(s)} \lambda^{r+s} x^{r-1} (z-x)^{s-1} e^{-\lambda z}\,dx \\
&= \int_0^1 \frac{1}{\Gamma(r)\Gamma(s)} \lambda^{r+s} (zu)^{r-1} (z-zu)^{s-1} e^{-\lambda z} z\,du \qquad (x = zu, dx = z\,du) \\
&= \frac{1}{\Gamma(r+s)} \lambda^{r+s} z^{r+s-1} e^{-\lambda z} \int_0^1 \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} u^{r-1} (1-u)^{s-1}\,du \\
&= f_{r+s,\lambda}(z) \int_0^1 \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} u^{r-1} (1-u)^{s-1}\,du
\end{aligned}$$

where $f_{r+s,\lambda}(z)$ is the gamma $(r+s, \lambda)$ density. The integral on the right is a constant which does not depend on z. Since both $f_Z(z)$ and $f_{r+s,\lambda}(z)$ are probability densities on $(0, \infty)$, integrating both sides with respect to z from 0 to ∞ gives

$$1 = 1 \times \int_0^1 \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} u^{r-1} (1-u)^{s-1}\,du$$

So the integral must equal 1. Therefore Z has the gamma $(r+s, \lambda)$ density. □

The last line of the previous argument evaluates an important integral:

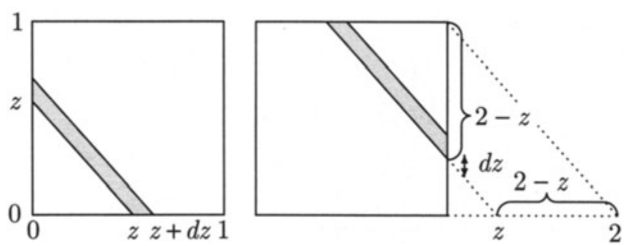
The Beta Integral

$$B(r,s) = \int_0^1 u^{r-1}(1-u)^{s-1}du = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)} \qquad (r > 0, s > 0)$$

This evaluation of $B(r,s)$ in terms of the gamma function agrees with the evaluation in Section 4.6 for integer r and s because $\Gamma(r) = (r-1)!$ for positive integers r.

Example 3. Sums of independent uniform variables.

Two terms. Suppose X and Y are independent, each with uniform $(0,1)$ distribution. To find the density of $X+Y$ it is simpler to work directly with a diagram than to use the convolution formula. Here (X,Y) has uniform distribution on the unit square. See Figure 2 on page 380.

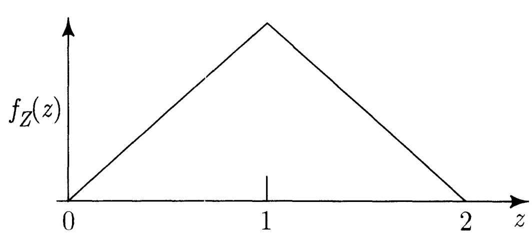


For $0 < z < 1$, the event $(X+Y \in dz)$ is represented as in the diagram by a shape of area $z\,dz + \frac{1}{2}(dz)^2$, by splitting the area into a parallelogram with altitude z perpendicular to sides of length dz, plus half a square of side dz. Ignoring the $(dz)^2$ as negligible in comparison to dz, gives simply

$$P(Z \in dz) = z\,dz$$

since the total area is 1. Similarly, for $1 \le z < 2$,

$$P(Z \in dz) = (2-z)dz$$



Thus $Z = X + Y$ has a tent-shaped density,

$$f_Z(z) = \begin{cases} z & 0 < z < 1 \\ 2-z & 1 \le z < 2 \\ 0 & \text{otherwise} \end{cases}$$

Three terms. Consider now $T = X + Y + W$ where X, Y, and W are independent uniform $(0,1)$. The joint distribution of (X,Y,W) is now uniform on a unit cube, and the density of T is proportional to the areas of slices through the cube perpendicular to an axis passing through the long diagonal. As you can convince yourself by handling a real cube, there are now several cases depending on which faces of the cube cut the slicing plane. This 3-dimensional geometry is tricky, but it reduces to two simpler two-dimensional problems.

To compute the density of $T = X + Y + W$ where X, Y, and W are independent uniform $(0, 1)$, write $T = X+Y+W = Z+W$, say, where the density of $Z = X+Y$ was found before. The convolution formula gives the density of $T = Z + W$ as an integral

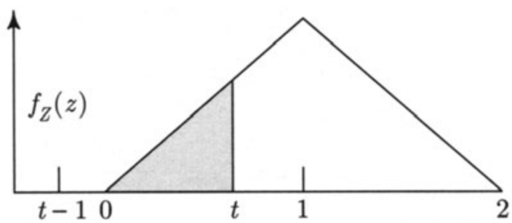
$$\begin{aligned} f_T(t) &= \int_{-\infty}^{\infty} f_Z(z) f_W(t-z)dz \\ &= \int_{t-1}^{t} f_Z(z)dz \qquad \text{since} \quad f_W(t-z) = 1 \text{ for } t-1 < z < t, \quad 0 \text{ else} \\ &= P(t-1 < Z < t) \qquad \text{by definition of } f_Z \end{aligned}$$

So the probability density of T at t turns out in this case to be a probability defined in terms of Z and t. This probability is represented by the shaded areas under the density $f_Z(z)$ in the diagrams that follow. There are 3 cases to consider.

Case 1. $0 < t < 1$. Then $t - 1 < 0$, so

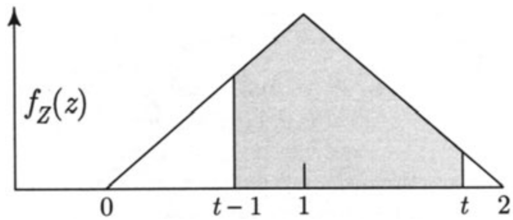
$$f_T(t) = P(t-1 < Z < t) = \frac{1}{2}t^2$$

by the formula for area of a triangle.



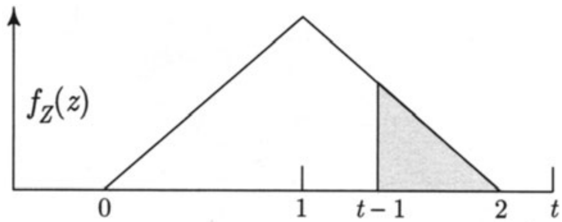
Case 2. $1 < t < 2$. Then $0 < t - 1 < 1$. The relevant area is a unit square less two triangles, hence

$$\begin{aligned} f_T(t) &= P(t-1 < Z < t) \\ &= 1 - \frac{1}{2}(2-t)^2 - \frac{1}{2}(t-1)^2 \\ &= -t^2 + 3t - \frac{3}{2} \end{aligned}$$



Case 3. $2 < t < 3$. Then $1 < t - 1 < 2$. The relevant area is now another triangle

$$f_T(t) = P(t - 1 < Z < t) = \frac{1}{2}(3 - t)^2$$



To summarize, the density of the sum $T = X + Y + W$ of three independent uniform $(0, 1)$ variables is $f_T(t)$, as defined above by quadratic functions of t, on each of the intervals $(0, 1)$, $(1, 2)$, and $(2, 3)$, and zero elsewhere. See Figures 2 and 3. Note the symmetric bell shape of the density of T.

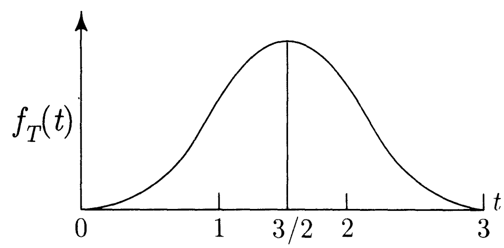


Illustration of Example 3 by numerical calculations. Let $T = X + Y + W$ where X, Y, and W are independent with uniform $(0, 1)$ distribution. Let us find:

a) $P(T < 3/2) = 1/2$ by symmetry of the density of T about $3/2$,

b) $P(1/2 < T < 3/2) = P(T < 3/2) - P(T \le 1/2)$

$$= 1/2 - \int_0^{1/2} \frac{t^2}{2} dt = \frac{23}{48} \quad \text{by a) and Case 1 on page 378}$$

c) $P(T > 5/2) = P(T \le 1/2) = 1/48$ by integral evaluated in b);

d) $E(T) = 3/2$ by symmetry;

e) $SD(T) = \sqrt{3} SD(W)$ where W is uniform $(0, 1)$, by the square root law

$= \sqrt{3} \cdot 1/\sqrt{12}$ by calculation done in Section 4.1

$= 1/2$

FIGURE 2. Distribution of the sum of two independent uniform $(0, 1)$ variables X and Y. The joint density of (X, Y) is suggested by a scatter, along with graphs of the densities of X, Y, and $X+Y$.

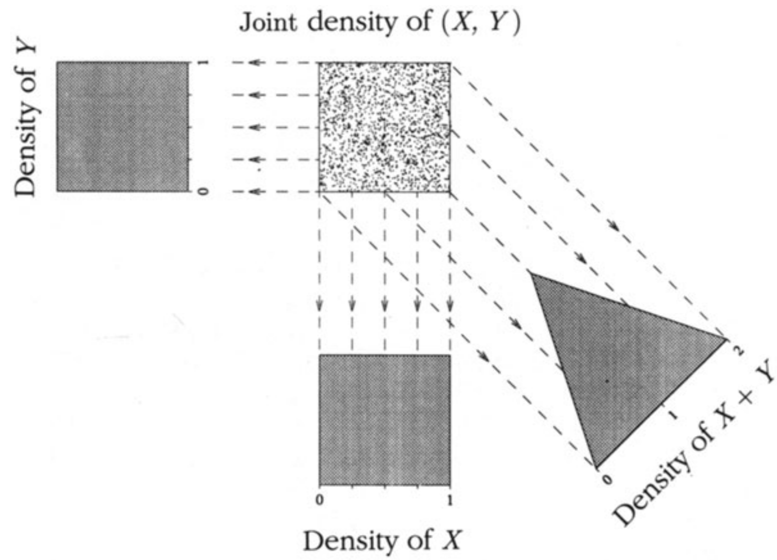


FIGURE 3. Distribution of the sum of three independent uniform $(0, 1)$ variables X, Y, and W. The joint density of $(X+Y, W)$ is suggested by a scatter, along with graphs of the densities of $X+Y$, W, and $X+Y+W$.

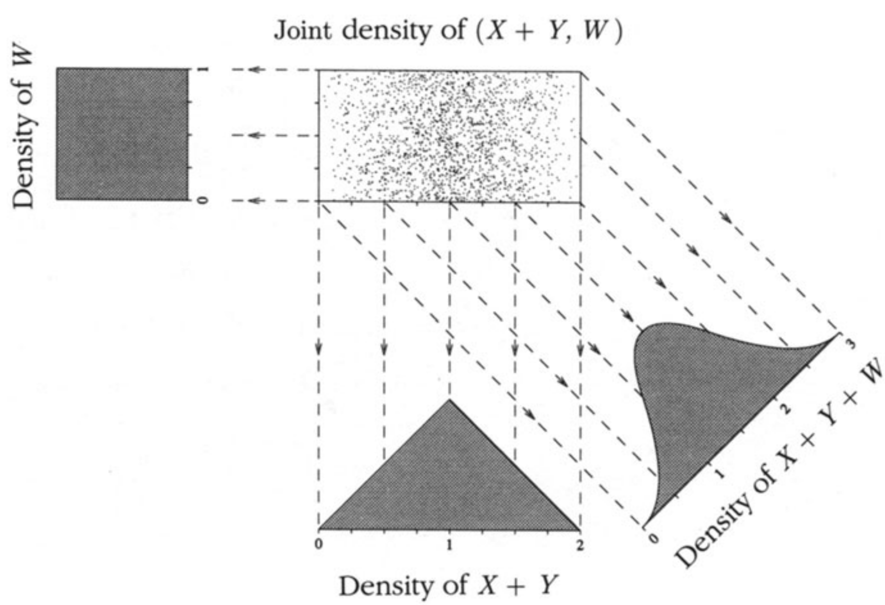
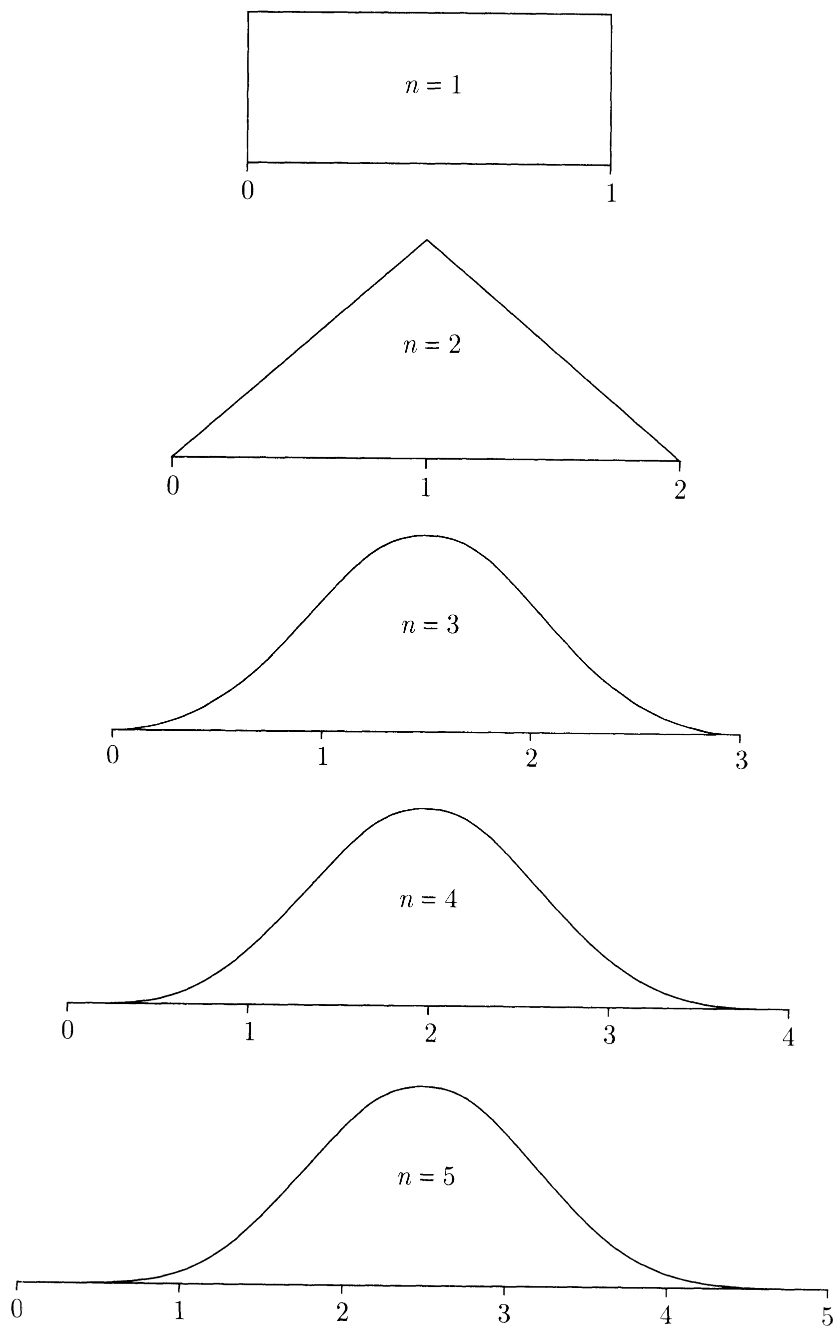


FIGURE 4. Density of the sum of n independent uniform $(0, 1)$ variables. The graphs are all centered at the mean with a constant horizontal distance on the page representing one standard unit in each graph. This shows how rapidly the shape of the distribution becomes normal as n increases.



Example 4. Roundoff errors.

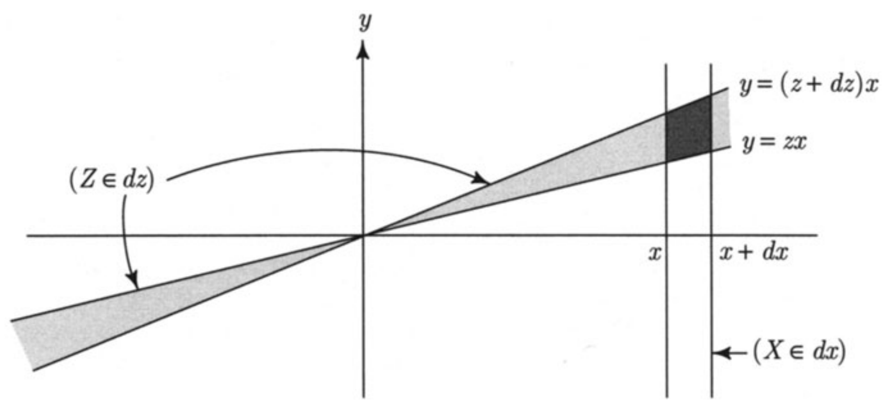
Problem. Suppose three numbers are computed, each with a roundoff error known to be smaller than 10^{-6} in absolute value. If the roundoff errors are assumed independent and uniformly distributed, what is the probability that the sum of the rounded numbers differs from the true sum of the numbers by more than 2×10^{-6}?

Solution. Let X_i be the error in the ith number in multiples of 10^{-6}, so the X_i are independent uniform $(-1, 1)$. To reduce to previous calculations, let $U_i = (X_i + 1)/2$, so the U_i are independent uniform $(0, 1)$. The problem is to find

$$\begin{aligned} P(|X_1 + X_2 + X_3| > 2) &= 2P(X_1 + X_2 + X_3 > 2) \quad \text{by symmetry} \\ &= 2P(2U_1 - 1 + 2U_2 - 1 + 2U_3 - 1 > 2) \\ &= 2P(U_1 + U_2 + U_3 > 5/2) \\ &= 2/48 \quad \text{by numerical calculation c) of Example 3 above.} \end{aligned}$$

Distribution of Ratios

Let $Z = Y/X$. The event $(Z \in dz)$ is shaded in the following diagram, for $z > 0$.



The event $(Z \in dz)$ is broken up into vertical slices according to values of X. The heavily shaded region, near (x, xz), represents the event $(X \in dx, Z \in dz)$. For small dx and dz this region is approximately a parallelogram. The left side has length $|x|dz$, and there is distance dx between the two vertical sides, so the area of the parallelogram is approximately $|x|dz\,dx = |x|dx\,dz$. The probability density over the small parallelogram can be taken to be $f(x, xz)$, so as dx and dz tend to zero we obtain the formula

$$P(X \in dx, Y/X \in dz) = f(x, xz)|x|dx\,dz$$

This works just as well for $z < 0$, though the picture looks a little different. Integrating out x yields

(f) $$f_{Y/X}(z) = \int_{-\infty}^{\infty} |x| f(x, xz) dx$$

As a special case, if X and Y are independent positive random variables, (f) reduces to $f_{Y/X}(z) = 0$ for $z \leq 0$, and

(g) $$f_{Y/X}(z) = \int_{0}^{\infty} x f_X(x) f_Y(xz) dx \qquad (0 < z < \infty)$$

Example 5. **Ratio of independent normal variables.**

Suppose that X and Y are independent and normally distributed with mean 0 and variance σ^2.

Problem. Find the distribution of X/Y.

Solution. We may assume $\sigma = 1$, since

$$\frac{X}{Y} = \frac{X/\sigma}{Y/\sigma} \quad \text{and both} \quad X/\sigma \quad \text{and} \quad Y/\sigma \qquad \text{are standard normal.}$$

By symmetry between X and Y and (f) above

$$\begin{aligned} f_{X/Y}(z) = f_{Y/X}(z) &= \int_{-\infty}^{\infty} |x| f_{X,Y}(x, xz) dx \\ &= \int_{-\infty}^{\infty} |x| \cdot \frac{1}{2\pi} e^{-\frac{x^2+x^2z^2}{2}} dx \\ &= \int_{0}^{\infty} \frac{1}{\pi} x e^{-\frac{x^2(z^2+1)}{2}} dx \\ &= \frac{1}{\pi} \cdot \frac{(-1)}{z^2+1} \cdot e^{-\frac{x^2(z^2+1)}{2}} \Big|_0^{\infty} \\ &= \frac{1}{\pi(z^2+1)} \end{aligned}$$

That is, X/Y has Cauchy distribution (see Exercise 4.4.6).

Remark. This calculation illustates the general method, but is a bit heavy handed. In fact the distribution of X/Y is Cauchy whenever the joint distribution of X and Y is symmetric under rotations. See Exercise 14 below.

Exercises 5.4

1. Let X_1 be uniform $(0,1)$ independent of X_2, that is, uniform $(0,2)$. Find:

a) $P(X_1 + X_2 \leq 2)$; b) the density of $X_1 + X_2$; c) the c.d.f. of $X_1 + X_2$.

2. Let S_n be the sum of n independent uniform $(0,1)$ random variables. Find

a) $P(S_2 \leq 1.5)$; b) $P(S_3 \leq 1.5)$; c) $P(S_3 \leq 1.1)$;

d) $P(1.0 \leq S_3 \leq 1.001)$ approximately.

3. A computer job must pass through two queues before it is processed. Suppose the waiting time in the first queue is exponential with rate α, and the waiting time in the second queue is exponential with rate β, independent of the first.

a) Find the density of the total time the job spends waiting in the two queues. Sketch the density in case $\alpha = 1$ and $\beta = 2$.

b) Find the expected total waiting time in terms of α and β.

c) Find the SD of the total waiting time in terms of α and β.

4. A system consists of two components. Suppose each component is subject to failure at constant rate λ, independently of the other, up to when the first component fails. After that moment the remaining component is subject to additional load and to failure at constant rate 2λ.

a) Find the distribution of the time until both components have failed.

b) What are the mean and variance of this distribution?

c) Find the 90th percentile of this distribution.

5. Let X be the number on a die roll, between 1 and 6. Let Y be a random number which is uniformly distributed on $[0, 1]$, independent of X. Let $Z = 10X + 10Y$.

a) What is the distribution of Z? Explain.

b) Find $P(29 \leq Z \leq 58)$.

6. Suppose $X_1, X_2, \ldots, X_n$ are independent and X_i has gamma (r_i, λ) distribution. What is the distribution of $X_1 + X_2 + \cdots + X_n$? Explain.

7. Let X and Y have joint density $f(x, y)$. Find formulae for the densities of each of the following random variables: a) XY; b) $X - Y$; c) $X + 2Y$.

8. Let X and Y be independent exponential variables with rates α and β. Find the c.d.f. of X/Y.

9. Find the density of $X = UV$ for independent uniform $(0,1)$ variables U and V.

10. Find the density of $Y = U/V$ for independent uniform $(0,1)$ variables U and V.

11. Find the distribution of $\min(U, V)/\max(U, V)$ for independent uniform$(0,1)$ variables U and V.

12. Let U, V be independent random variables, each uniform on $(0,1)$.

(a) Find the density of $X = -\log\{U(1 - V)\}$. b) Compute $E(X)$ and $Var(X)$.

13. Find the density of $Z = X - Y$ for independent exponential (λ) variables X and Y.

14. Let X and Y have a joint distribution which is symmetric under rotations (e.g., uniform on a circle around 0, or uniform on a disc centered at 0). By changing to polar coordinates, show that

a) the distribution of X/Y is Cauchy (see Exercise 4.4.6);

b) the distribution of $X^2/(X^2 + Y^2)$ is arcsine (see Exercise 4.4.8).

15. Let $Z = \min(X, Y)/\max(X, Y)$ for independent exponential (λ) variables X and Y.

a) Explain with little calculation why the distribution of Z does not depend on λ

b) Let $0 < z < 1$. Identify the set $(Z \le z)$ as a subset of the (x, y) plane, and calculate $P(Z \le z)$ by integration of the joint density over this subset.

c) Find the density of Z at z for $0 < z < 1$.

16. Consider the c.d.f. of T with gamma (r, λ) distribution, $F(r, \lambda, t) = P(T \le t)$. Section 4.2 gives a formula for $F(r, \lambda, t)$ for integer r, but for r not an integer there is no simple formula for $F(r, \lambda, t)$.

a) Show that for fixed r and t, $F(r, \lambda, t)$ is an increasing function of λ. [*Hint*: Rescale to the gamma $(r, 1)$ distribution.]

b) Show that for fixed λ and t, $F(r, \lambda, t)$ is a decreasing function of r. [*Hint*: Use sums of independent gamma variables.]

17. Take a unit cube in three dimensions. Cut the cube by a plane perpendicular to the line from its corners $(0, 0, 0)$ and $(1, 1, 1)$, that cuts this line at the point $(t/3, t/3, t/3)$.

a) What is the cross-sectional area of this slice through the cube?

b) Check your answer by describing geometrically the shape of the cross section in the case when $t \le 1$ and $t = 3/2$.

18. Let f_n be the density function and F_n the c.d.f. of the sum S_n of n independent uniform $(0, 1)$ random variables.

a) Show that $f_n(x) = F_{n-1}(x) - F_{n-1}(x - 1)$.

b) Show that on each of the n intervals $(i - 1, i)$ for $i = 1$ to n, f_n is equal to a polynomial of degree $n - 1$, and F_n is equal to a polynomial of degree n.

c) Find $f_n(x)$ and $F_n(x)$ for $0 \le x \le 1$.

d) Find $f_n(x)$ and $F_n(x)$ for $n - 1 \le x \le n$.

Find also: e) $P(0 \le S_4 \le 1)$; f) $P(1 \le S_4 \le 2)$; g) $P(1.5 \le S_4 \le 2)$.

19. Let X and Y be independent variables with gamma (r, λ) and gamma (s, λ) distribution, respectively. Show that $X/(X + Y)$ has beta (r, s) distribution, independently of $X + Y$.

Continuous Joint Distributions: Summary

Differential Formula for Joint Density

$$P(X \in dx, Y \in dy) = f(x,y)dx\,dy$$

The density $f(x,y)$ is the probability per unit area for values near (x,y). See pages 348 and 349 of Section 5.2 for properties of joint densities, and comparison with joint distributions in the discrete case.

Central Limit Theorem

Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables, each with mean μ and variance σ^2. Let $S_n = X_1 + \cdots + X_n$. Provided $\sigma^2 < \infty$, the limit distribution, as $n \to \infty$, of the standardized sum $Z_n = [S_n - n\mu]/(\sqrt{n}\sigma)$ is the standard normal distribution.

Formula for Density of $X + Y$

If (X,Y) has density $f(x,y)$ in the plane, then $X+Y$ has density on the line

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x, z-x)dx.$$

Convolution Formula

If X and Y are independent, then

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx.$$

Exact distribution of various functions of particular variables. See distribution summaries.

The Rayleigh Distribution

If X and Y are independent standard normal variables, then $R = \sqrt{X^2+Y^2}$ has the Rayleigh distribution, with density

$$f_R(r) = re^{-\frac{1}{2}r^2}, \qquad r > 0,$$

and distribution function

$$F_R(r) = 1 - e^{-\frac{1}{2}r^2}, \qquad r > 0.$$

The variable R represents the distance from the origin of the random point (X,Y).

Review Exercises

1. For X and Y independent and uniform $(0, 1)$, find $P(Y \geq 1/2 \,|\, Y \geq X^2)$.

2. For X and Y independent and both uniform $(-1, 1)$, find

a) $P(|X + Y| \leq 1)$; b) $E(|X + Y|)$.

3. Coin in a can. A coin of diameter 1 inch is tossed in the air and caught in an empty soup can of bottom radius 3 inches. The coin lies flat on the bottom.

a) What is the chance that the coin covers the center point of the bottom of the can?

Suppose that instead of the soup can, the coin is dropped into a box whose bottom is a square with sides of length 5 inches.

b) What is the chance that the coin covers the center point of the bottom of the box?

c) Consider one of the main diagonals of the bottom of the box. What is the probability that part of the coin crosses that diagonal line?

State any assumptions you make.

4. Let X and Y be independent with uniform $(-1, 1)$ distribution. Find

a) $P(X^2 + Y^2) \leq r^2$; b) the c.d.f. of $R^2 = X^2 + Y^2$; c) the density of R^2.

5. A point is chosen uniformly at random from a unit square. Let D be the distance of the point from the midpoint of one side of the square. Find a) $P(D \geq \frac{1}{2})$; b) $E(D^2)$.

6. For a particular kind of call, the phone company charges $1 for the first minute or any portion thereof, and one cent per second for time after the first minute. Calculate the approximate value of the long-run average charge per call assuming the distribution of call duration is:

a) exponential with mean 1 minute;

b) exponential with mean 2 minutes;

c) gamma with shape parameter 2 and mean 1 minute.

7. Suppose that $X_1, X_2, \ldots, X_{100}$ are independent random variables, with normal $(\mu, 1)$ distribution, representing 100 measurements whose average $\bar{X} = (X_1 + \cdots + X_{100})/100$ should be close to the number μ. Calculate the probability that $|\bar{X} - \mu| \geq 0.25$.

8. Suppose that $X_1, \ldots, X_{100}$ are independent random variables with common distribution with mean μ and variance 1, but not necessarily normally distributed. Repeat Exercise 7 with these assumptions. Explain why the answer will be approximately the same.

9. Let X be the number of heads in two fair coin tosses. Suppose U has uniform distribution on $(0, 1)$, independently of X.

a) Find the density of $X + U$ and sketch its graph.

b) Find an alternate distribution for U such that for any integer-valued random variable X independent of U, the graph of the density of $X + U$ is simply the usual histogram of the distribution of X.

10. Let X, Y be independent exponential random variables with parameters λ and μ.

a) Find the density function for $Z = \min(X, Y)$.

b) Calculate $P(X \geq Y)$.

c) Calculate $P(\frac{1}{2} < X/Y < 2)$, in the case $\lambda = \mu$. [*Hint*: Use the result of b).]

11. Let U and V be two independent uniform $(0, 1)$ random variables. Let $X = U/V$.

a) For $0 < x < 1$, calculate $P(X > x)$.

b) Find the c.d.f. F of the random variable X. Sketch the graph of F.

c) Find the density function f of X. Sketch the graph of f.

12. A marksman fires at the center of a target; he hits a random point (X, Y) (measured relative to the center of the target) such that X and Y are independent normal $(0, a^2)$ random variables. A second marksman fires, and hits at (X', Y') where X' and Y' are independent with normal $(0, b^2)$ distributions. What is the chance that the second marksman hits closer to the center of the target than the first marksman?

13. Suppose (X, Y) is uniformly distributed according to relative arc length on the circumference of the circle $\{(x, y) : x^2 + y^2 = 1\}$. Find the c.d.f. of

a) X; b) Y; c) $X + Y$.

14. Suppose U_1, U_2, U_3 are independent and uniform $(0, 1)$. Find: a) $P(U_1 < U_2 < U_3)$;

b) $E(U_1U_2U_3)$; c) $Var(U_1U_2U_3)$; d) $P(U_1U_2 > U_3)$; e) $P(\max(U_1, U_2) > U_3)$.

15. Repeat Exercise 14 for Z_i instead of U_i, where the Z_i are independent normal $(0, 1)$ random variables. Find also:

f) $P(Z_1^2 + Z_2^2 > 1)$ g) $P(Z_1 + Z_2 + Z_3 < 2)$;

h) $P(Z_1/Z_2 < 1)$; i) $P(3Z_1 - 2Z_2 < 4Z_3 + 1)$.

16. A point is picked randomly in space. Its three coordinates X, Y and Z are independent standard normal variables. Let $R = \sqrt{X^2 + Y^2 + Z^2}$ be the distance of the point from the origin. Find

a) the density of R^2; b) the density of R; c) $E(R)$; d) $Var(R)$.

17. Let $X_1, X_2, \ldots,$ be independent normally distributed random variables having mean 0 and variance 1. Use the normal approximation to find:

a) $P(X_1^2 + X_2^2 + \cdots + X_{100}^2 \geq 80)$;

b) a number c such that $P(100 - c \leq X_1^2 + \cdots + X_{100}^2 \leq 100 + c) \approx 0.95$.

18. For X and Y independent normal $(0, 1)$ variables, show that for $r > 0$

$$P(aX + bY \leq r\sqrt{a^2 + b^2} \text{ for all } a, b \geq 0) = \Phi(r) - \frac{1}{4}e^{-\frac{1}{2}r^2}$$

19. Independent Poisson processes. Suppose particles of d different kinds, labeled $k = 1, 2, \ldots, d$, arrive at a counter according to independent Poisson processes at rates λ_k. Let W_k be arrival time of the first particle of kind k. Let K_1 be the kind of the first particle to arrive, K_2 the kind of the second particle to arrive, and so on. So the K_n are discrete random variables with values in the set $\{1, \ldots, d\}$.

a) Express the event $(K_1 = k)$ in terms of the random variables $W_1, \ldots, W_d$.

b) Use this expression to find $p_k = P(K_1 = k), 1 \le k \le d$ in terms of $\lambda_1, \ldots, \lambda_d$.

c) Explain informally why $K_1, K_2, \ldots$ are independent with identical distribution.

d) Assuming the result of c), derive the formula for p_k in another way after filling in the blanks in the following statements e), f) and g): After a very long time T,

e) the number of arrivals of type k should be about ________.

f) the number of all arrivals of all types should be about ________.

g) the fraction of all arrivals that are of type k should be about ________.

20. Minimum of independent exponential variables. Let T_1 and T_2 be two independent exponential variables, with rates λ_1 and λ_2. Think of T_i as the lifetime of component i, $i = 1, 2$. Let T_{min} represent the lifetime of a system which fails whenever the first of the two components fails, so $T_{\text{min}} = \min(T_1, T_2)$. Let X_{min} designate which component failed first, so X_{min} has value 1 if $T_1 < T_2$ and value 2 if $T_2 < T_1$. Show:

a) that the distribution of T_{min} is exponential $(\lambda_1 + \lambda_2)$;

b) that the distribution of X_{min} is given by the formula $P(X_{\text{min}} = i) = \dfrac{\lambda_i}{\lambda_1 + \lambda_2}$ for $i = 1, 2$;

c) that the random variables T_{min} and X_{min} are independent;

d) how these results generalize simply to describe the minimum of n independent exponential random variables with rates $\lambda_1, \ldots, \lambda_n$.

21. Closest point. Consider a Poisson random scatter of points in a plane with mean intensity λ per unit area. Let R be the distance from 0 to the closest point of the scatter.

a) Find formulae for the c.d.f. and density of R, and sketch their graphs.

b) Show that $\sqrt{2\lambda\pi}R$ has the Rayleigh distribution described in Section 5.3

c) Use b) to find formulae for the mean and SD of R from results of Section 5.3.

d) Find the mode and the median of the distribution of R.

22. In Maxwell's model of a gas, molecules of mass m are assumed to have velocity components, V_x, V_y, V_z that are independent, with a joint distribution that is invariant under rotation of the three-dimensional coordinate system. Maxwell showed that V_x, V_y, V_z must have normal $(0, \sigma^2)$ distribution for some σ. Taking this result for granted:

a) find a formula for the density of the kinetic energy

$$K = \frac{1}{2}mV_x^2 + \frac{1}{2}mV_y^2 + \frac{1}{2}mV_z^2$$

b) find the mean and mode of the energy distribution.

23. Let Y be the minimum of three independent random variables with uniform distribution on $(0, 1)$, and let Z be their maximum. Find:

a) $P(Z \le \frac{2}{3} | Y \ge \frac{1}{3})$; b) $P(Z \le \frac{2}{3} | Y \le \frac{1}{3})$.

24. A coin of diameter d is tossed at random on a grid of squares of side s. Making appropriate assumptions, to be stated clearly, calculate:

a) the probability that the coin lands inside some square (i.e., not touching any line);

b) the probability that the coin lands heads inside some square.

Suppose now that the coin is tossed four times. Let X be the number of times it lands inside a square, Y the number of heads. Assume $d = s/2$. Calculate:

c) $P(X = Y)$; d) $P(X < Y)$; e) $P(X > Y)$.

25. Joint distribution of order statistics. Let $V_1 < V_2 < \cdots < V_n$ be the order statistics of n independent uniform $(0, 1)$ variables. (Refer to Section 4.6.) Let $1 \leq k < m \leq n$.

a) Find the joint density of V_k and V_m.

Now show that each of the following variables has a beta distribution, and identify the parameters: b) $V_m - V_k$; c) V_k/V_m;

26. Averages of order statistics. Let $V_1, \ldots, V_n$ be the order statistics of n independent uniform $(0, 1)$ variables. Let
$A_{\text{all}} = (V_1 + \cdots + V_n)/n$
$A_{1n} = (V_1 + V_n)/2$
$A_{\text{mid}} = V_{(n+1)/2}$ the middle value, where you can assume n is odd.

a) Show that for sufficiently large n, each of these three variables is most likely very close to $1/2$.

b) For all large enough values of n, one of these variables can be expected to be very much closer to $1/2$ than either of the two others. Which one, and why?

c) Confirm your answer to b) for $n = 100$ by finding for each of the A's a good approximation to the probability that it is between 0.49 and 0.51.

27. A box contains n balls numbered $1, \ldots, n$. Balls are drawn at random until the first draw that produces a ball obtained on some previous draw. Let D_n be the random number of draws required. So the possible values of D_n are $2, \ldots, n + 1$.

a) Check that for $0 < x < \infty$,

$$\lim_{n \to \infty} P(D_n/\sqrt{n} > x) = e^{-x^2/2}$$

That is to say, the limit distribution of $D_n/\sqrt{n}$ is the Rayleigh distribution.

b) Assuming a switch in the order of the limit and integration can be justified (it can, but do not worry about that), deduce that

$$\lim_{n \to \infty} E(D_n/\sqrt{n}) = \sqrt{\pi/2}$$

c) There seems to be no simpler expression for $E(D_n)$ than a sum of n or $n + 1$ terms. But the terms can be arranged in some interesting ways. Show by writing $E(D_n)$ as the sum of the tail probabilities $P(D_n > k)$ in reverse order that

$$E(D_n) = P(X_n \leq n)\, n!\, n^{-n} e^n$$

where X_n is a Poisson random variable with mean n.

d) Deduce the limit of $P(X_n \leq n)$ as $n \to \infty$ from the central limit theorem, then combine b) and c) to give a derivation of Stirling's formula

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

e) Derive the following formula, which is surprisingly simple in view of c)

$$E(D_n^2 - D_n) = 2n$$

f) Transform the identity e) as in the calculation c) to derive the formula

$$E(|X_n - n|) = \frac{2n^{n+1}}{e^n n!}$$

and give yet another derivation of Stirling's formula, much as in d) above, this time using the central limit theorem instead of a).

28. Volumes in higher dimensions. Use the derivation of the chi-square distribution to derive part a), then use a) for the remaining parts:

a) Find the constant c_n such that the $(n-1)$-dimensional volume of the "surface" of a sphere of radius r in n-dimensional space is $c_n r^{n-1}$.

b) Find d_n so the n-dimensional volume inside a sphere of radius r is $d_n r^n$.

c) An n-dimensional sphere of radius r is packed inside an n-dimensional cube with sides of length $2r$. What proportion p_n of the volume of the cube is inside the sphere?

d) Use Stirling's formula $\Gamma(s) \sim \sqrt{2\pi} s^{s-1/2} e^{-s}$ as $s \to \infty$ to find a simple approximation for p_n for large n. What is the limit of p_n as $n \to \infty$?

e) Interpret p_n probabilistically in terms of n independent uniform $(-1, 1)$ variables.

29. A needle is tossed at random on a grid of equally spaced parallel lines. Assume the needle is so much longer than the spacing between the lines that the possibility of the needle not crossing any line can be neglected. Let X be the distance between the center of the needle and the closest point at which the needle crosses one of the lines. Find:

a) the distribution function of X;

b) the density function of X.

30. Random walk inside squares. Draw a square centered at $(0, 0)$ with sides of length 2 parallel to the axes, so the corners are at $(\pm 1, \pm 1)$. Let (X_1, Y_1) be picked uniformly at random from the area inside this square. Given (X_1, Y_1), draw a square centered at (X_1, Y_1), with sides of length 2 parallel to the axes, so the corners are at $(X_1 \pm 1, X_2 \pm 1)$. Let (X_2, Y_2), be picked uniformly at random from the area inside this square, and so on: Given $(X_1, Y_1), \ldots, (X_n, Y_n)$ let (X_{n+1}, Y_{n+1}) be picked uniformly at random from the area inside the square with corners at $(X_n \pm 1, Y_n \pm 1)$. For $n = 300$, use a normal approximation to find the following probabilities:

a) $P(|X_n| > 10)$; b) $P(|Y_n| > 10)$.

c) The probability that (X_n, Y_n) lies outside the square with corners at $(\pm 10, \pm 10)$.

d) The probability that (X_n, Y_n) lies outside the circle of radius 10 centered at $(0, 0)$.

31. Random walk inside circles. Fix $r > 0$. Draw a circle centered at $(0, 0)$ with radius r. Let (X_1, Y_1) be picked uniformly at random from the area inside this circle. Given (X_1, Y_1), draw a circle with radius r centered at (X_1, Y_1). Let (X_2, Y_2), be picked uniformly at random from the area inside this circle, and so on. Given $(X_1, Y_1), \ldots, (X_n, Y_n)$, let (X_{n+1}, Y_{n+1}) be picked uniformly at random from the area inside the circle around (X_n, Y_n) with radius r.

a) Find r so that for large n the distribution of X_n in this problem is nearly the same as in Exercise 30 for a square of side 2 instead of the circle of radius r. [*Hint*: Find $E[X_1^2]$ by considering $E[Y_1^2]$ as well.]

b) Are X_n and Y_n independent?

c) The point (X_n, Y_n) is projected onto the line rotated an angle θ from the X-axis at $X_n \cos\theta + Y_n \sin\theta$ measured from the origin along this line. Use the normal approximation for sums of independent random variables to show that with r as in part a), for every $\theta \in [0, 2\pi]$ and for large n, the distribution of $X_n \cos\theta + Y_n \sin\theta$ is nearly the same for both the circle of radius r and the square of side 2.

d) It is known that a joint distribution of (X, Y) in the plane is determined by the distributions of all the projections $X \cos\theta + Y \sin\theta$ as θ ranges over $[0, 2\pi]$. In particular if $X \cos\theta + Y \sin\theta$ has standard normal distribution for every θ then X and Y are independent standard normal variables. An approximate version of this result is also true: if $X \cos\theta + Y \sin\theta$ has approximately the standard normal distribution for every θ, then X and Y are approximately independent standard normal variables. Apply this result and part c) to approximate the probability that for r as in part a), and $n = 300$, the point (X_n, Y_n) defined using circles of radius r lies outside the circle of radius 10 centered at the origin.

32. Random walk on circles. Repeat Exercise 31 for the motion defined by picking points at random according to the uniform distribution on the *perimeter* of the circle of radius r, so each new point is at distance r from the previous one, in a random direction.

33. Mixture of discrete and continuous. Repeat Exercise 31 for the motion defined by repeatedly picking points at random according to the uniform distribution (proportional to length) on the perimeter of a square centered at the current point with sides of length $2r$. Note that the distribution of X_n in this case is neither discrete nor continuous but a mixture of the two kinds. The second moment of X_1 is defined by adding the discrete and continuous parts. It can be shown that the usual method of calculating the second moment of X_n is still valid, and that the normal approximation is still correct in the limit of large n. Following parts a) to d) as in Exercise 31,

e) Calculate and plot the graph of the distribution function of X_1.

f) Calculate and plot the graph of the distribution function of X_2.

g) What is the total probability in the discrete part of the distribution of X_n?

34. Ratios of sums of squares.

a) Use the result of Exercise 5.4.19 to show that if X, Y and Z are independent normal $(0, 1)$ random variables, then $X^2/(X^2 + Y^2 + Z^2)$ has beta $(1/2, 1)$ distribution, independent of $X^2 + Y^2 + Z^2$.

b) Suppose that (U, V, W) has uniform distribution on the surface of the unit sphere in three dimensions. Deduce from a) that $U^2/(U^2 + V^2 + W^2)$ has beta $(1/2, 1)$ distribution.

c) What is the distribution of $U^2/(U^2 + V^2 + W^2)$ if (U, V, W) has uniform distribution over the volume inside the unit sphere in three dimensions?

d) Suppose that $U_1, U_2, \ldots, U_n$ are independent uniform $(-1, 1)$ variables. For $1 \le k \le n$, let $S_k = U_1^2 + \cdots + U_k^2$. Find the the conditional distribution of S_k/S_n given that $S_n \le 1$.