Functions of Random Variables

1 Introduction

In this chapter we shall concern ourselves with the problem of finding the probability distributions or densities of **functions of one or more random variables**. That is, given a set of random variables $X_1, X_2, \ldots, X_n$ and their joint probability distribution or density, we shall be interested in finding the probability distribution or density of some random variable $Y = u(X_1, X_2, \ldots, X_n)$. This means that the values of Y are related to those of the X's by means of the equation

$$y = u(x_1, x_2, \ldots, x_n)$$

Several methods are available for solving this kind of problem. The ones we shall discuss in the next four sections are called the **distribution function technique**, the **transformation technique**, and the **moment-generating function technique**. Although all three methods can be used in some situations, in most problems one technique will be preferable (easier to use than the others). This is true, for example, in some instances where the function in question is linear in the random variables $X_1, X_2, \ldots, X_n$, and the moment-generating function technique yields the simplest derivations.

2 Distribution Function Technique

A straightforward method of obtaining the probability density of a function of continuous random variables consists of first finding its distribution function and then its probability density by differentiation. Thus, if $X_1, X_2, \ldots, X_n$ are continuous random variables with a given joint probability density, the probability density of $Y = u(X_1, X_2, \ldots, X_n)$ is obtained by first determining an expression for the probability

$$F(y) = P(Y \leqq y) = P[u(X_1, X_2, \ldots, X_n) \leqq y]$$

and then differentiating to get

$$f(y) = \frac{dF(y)}{dy}$$

EXAMPLE 1

If the probability density of X is given by

$$f(x) = \begin{cases} 6x(1-x) & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

find the probability density of $Y = X^3$.

Solution

Letting $G(y)$ denote the value of the distribution function of Y at y, we can write

$$\begin{aligned} G(y) &= P(Y \leqq y) \\ &= P(X^3 \leqq y) \\ &= P(X \leqq y^{1/3}) \\ &= \int_0^{y^{1/3}} 6x(1-x)\,dx \\ &= 3y^{2/3} - 2y \end{aligned}$$

and hence

$$g(y) = 2(y^{-1/3} - 1)$$

for $0 < y < 1$; elsewhere, $g(y) = 0$. In Exercise 15 the reader will be asked to verify this result by a different technique.

EXAMPLE 2

If $Y = |X|$, show that

$$g(y) = \begin{cases} f(y) + f(-y) & \text{for } y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

where $f(x)$ is the value of the probability density of X at x and $g(y)$ is the value of the probability density of Y at y. Also, use this result to find the probability density of $Y = |X|$ when X has the standard normal distribution.

Solution

For $y > 0$ we have

$$\begin{aligned} G(y) &= P(Y \leqq y) \\ &= P(|X| \leqq y) \\ &= P(-y \leqq X \leqq y) \\ &= F(y) - F(-y) \end{aligned}$$

and, upon differentiation,

$$g(y) = f(y) + f(-y)$$

Also, since $|x|$ cannot be negative, $g(y) = 0$ for $y < 0$. Arbitrarily letting $g(0) = 0$, we can thus write

$$g(y) = \begin{cases} f(y) + f(-y) & \text{for } y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

If X has the standard normal distribution and $Y = |X|$, it follows that

$$\begin{aligned} g(y) &= n(y; 0, 1) + n(-y; 0, 1) \\ &= 2n(y; 0, 1) \end{aligned}$$

for $y > 0$ and $g(y) = 0$ elsewhere. An important application of this result may be found in Example 9.

EXAMPLE 3

If the joint density of X_1 and X_2 is given by

$$f(x_1, x_2) = \begin{cases} 6e^{-3x_1 - 2x_2} & \text{for } x_1 > 0, x_2 > 0 \\ 0 & \text{elsewhere} \end{cases}$$

find the probability density of $Y = X_1 + X_2$.

Solution

Integrating the joint density over the shaded region of Figure 1, we get

$$\begin{aligned} F(y) &= \int_0^y \int_0^{y-x_2} 6e^{-3x_1 - 2x_2} dx_1\, dx_2 \\ &= 1 + 2e^{-3y} - 3e^{-2y} \end{aligned}$$

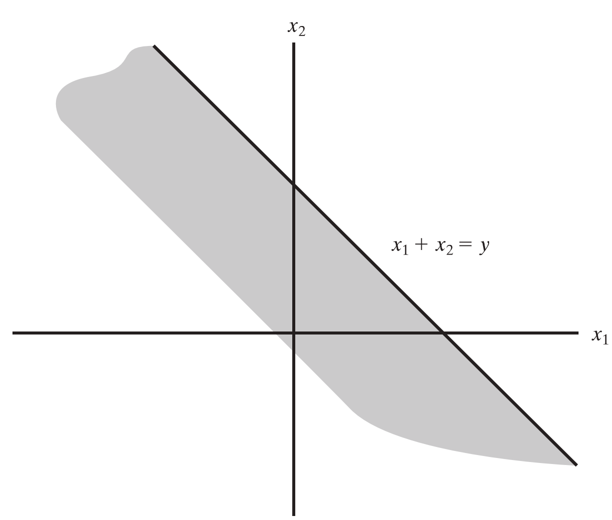

Figure 1. Diagram for Example 3.

and, differentiating with respect to y, we obtain

$$f(y) = 6(e^{-2y} - e^{-3y})$$

for $y > 0$; elsewhere, $f(y) = 0$.

Exercises

1. If X has an exponential distribution with the parameter θ, use the distribution function technique to find the probability density of the random variable $Y = \ln X$.

2. If the probability density of X is given by

$$f(x) = \begin{cases} 2xe^{-x^2} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

and $Y = X^2$, find
(a) the distribution function of Y;
(b) the probability density of Y.

3. If X has the uniform density with the parameters $\alpha = 0$ and $\beta = 1$, use the distribution function technique to find the probability density of the random variable $Y = \sqrt{X}$.

4. If the joint probability density of X and Y is given by

$$f(x, y) = \begin{cases} 4xye^{-(x^2+y^2)} & \text{for } x > 0, y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

and $Z = \sqrt{X^2 + Y^2}$, find
(a) the distribution function of Z;
(b) the probability density of Z.

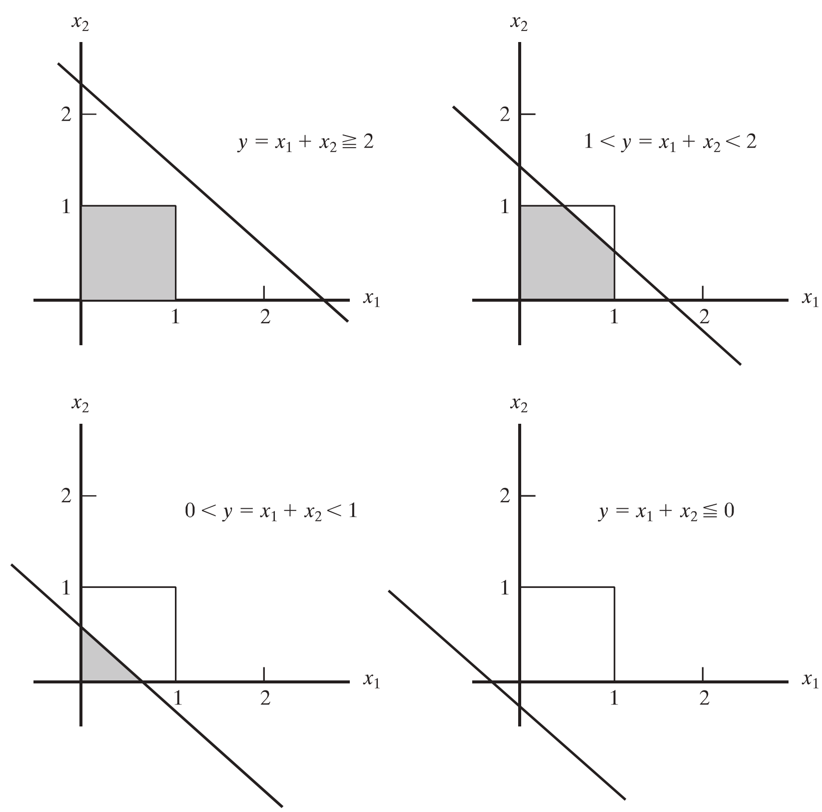

Figure 2. Diagram for Exercise 6.

5. If X_1 and X_2 are independent random variables having exponential densities with the parameters θ_1 and θ_2, use the distribution function technique to find the probability density of $Y = X_1 + X_2$ when

(a) $\theta_1 \neq \theta_2$;
(b) $\theta_1 = \theta_2$.

(Example 3 is a special case of this with $\theta_1 = \frac{1}{3}$ and $\theta_2 = \frac{1}{2}$.)

6. Let X_1 and X_2 be independent random variables having the uniform density with $\alpha = 0$ and $\beta = 1$. Referring to Figure 2, find expressions for the distribution function of $Y = X_1 + X_2$ for

(a) $y \leqq 0$;
(b) $0 < y < 1$;
(c) $1 < y < 2$;
(d) $y \geqq 2$.

Also find the probability density of Y.

7. With reference to the two random variables of Exercise 5, show that if $\theta_1 = \theta_2 = 1$, the random variable

$$Z = \frac{X_1}{X_1 + X_2}$$

has the uniform density with $\alpha = 0$ and $\beta = 1$.

8. If the joint density of X and Y is given by

$$f(x, y) = \begin{cases} e^{-(x+y)} & \text{for } x > 0, y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

and $Z = \dfrac{X + Y}{2}$, find the probability density of Z by the distribution function technique.

3 Transformation Technique: One Variable

Let us show how the probability distribution or density of a function of a random variable can be determined without first getting its distribution function. In the discrete case there is no real problem as long as the relationship between the values of X and $Y = u(X)$ is one-to-one; all we have to do is make the appropriate substitution.

EXAMPLE 4

If X is the number of heads obtained in four tosses of a balanced coin, find the probability distribution of $Y = \dfrac{1}{1 + X}$.

Solution

Using the formula for the binomial distribution with $n = 4$ and $\theta = \frac{1}{2}$, we find that the probability distribution of X is given by

x	0	1	2	3	4
$f(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Then, using the relationship $y = \dfrac{1}{1 + x}$ to substitute values of Y for values of X, we find that the probability distribution of Y is given by

y	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$
$g(y)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

If we had wanted to make the substitution directly in the formula for the binomial distribution with $n = 4$ and $\theta = \frac{1}{2}$, we could have substituted $x = \dfrac{1}{y} - 1$ for x in

$$f(x) = \binom{4}{x}\left(\frac{1}{2}\right)^4 \quad \text{for } x = 0, 1, 2, 3, 4$$

getting

$$g(y) = f\left(\frac{1}{y} - 1\right) = \binom{4}{\frac{1}{y} - 1}\left(\frac{1}{2}\right)^4 \quad \text{for } y = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}$$

Note that in the preceding example the probabilities remained unchanged; the only difference is that in the result they are associated with the various values of Y instead of the corresponding values of X. That is all there is to the **transformation** (or **change-of-variable**) **technique** in the discrete case as long as the relationship is one-to-one. If the relationship is not one-to-one, we may proceed as in the following example.

EXAMPLE 5

With reference to Example 4, find the probability distribution of the random variable $Z = (X - 2)^2$.

Solution

Calculating the probabilities $h(z)$ associated with the various values of Z, we get

$$h(0) = f(2) = \frac{6}{16}$$

$$h(1) = f(1) + f(3) = \frac{4}{16} + \frac{4}{16} = \frac{8}{16}$$

$$h(4) = f(0) + f(4) = \frac{1}{16} + \frac{1}{16} = \frac{2}{16}$$

and hence

z	0	1	4
$h(z)$	$\frac{3}{8}$	$\frac{4}{8}$	$\frac{1}{8}$

To perform a transformation of variable in the continuous case, we shall assume that the function given by $y = u(x)$ is differentiable and either increasing or decreasing for all values within the range of X for which $f(x) \neq 0$, so the inverse function, given by $x = w(y)$, exists for all the corresponding values of y and is differentiable except where $u'(x) = 0$.† Under these conditions, we can prove the following theorem.

†To avoid points where $u'(x)$ might be 0, we generally do not include the endpoints of the intervals for which probability densities are nonzero. This is the practice that we follow throughout this chapter.

THEOREM 1. Let $f(x)$ be the value of the probability density of the continuous random variable X at x. If the function given by $y = u(x)$ is differentiable and either increasing or decreasing for all values within the range of X for which $f(x) \neq 0$, then, for these values of x, the equation $y = u(x)$ can be uniquely solved for x to give $x = w(y)$, and for the corresponding values of y the probability density of $Y = u(X)$ is given by

$$g(y) = f[w(y)] \cdot |w'(y)| \quad \text{provided } u'(x) \neq 0$$

Elsewhere, $g(y) = 0$.

Proof First, let us prove the case where the function given by $y = u(x)$ is increasing. As can be seen from Figure 3, X must take on a value between $w(a)$ and $w(b)$ when Y takes on a value between a and b. Hence,

$$\begin{aligned} P(a < Y < b) &= P[w(a) < X < w(b)] \\ &= \int_{w(a)}^{w(b)} f(x)\, dx \\ &= \int_a^b f[w(y)]w'(y)\, dy \end{aligned}$$

where we performed the change of variable $y = u(x)$, or equivalently $x = w(y)$, in the integral. The integrand gives the probability density of Y as long as $w'(y)$ exists, and we can write

$$g(y) = f[w(y)]w'(y)$$

When the function given by $y = u(x)$ is decreasing, it can be seen from Figure 3 that X must take on a value between $w(b)$ and $w(a)$ when Y takes on a value between a and b. Hence,

$$\begin{aligned} P(a < Y < b) &= P[w(b) < X < w(a)] \\ &= \int_{w(b)}^{w(a)} f(x)\, dx \\ &= \int_b^a f[w(y)]w'(y)\, dy \\ &= -\int_a^b f[w(y)]w'(y)\, dy \end{aligned}$$

where we performed the same change of variable as before, and it follows that

$$g(y) = -f[w(y)]w'(y)$$

Since $w'(y) = \dfrac{dx}{dy} = \dfrac{1}{\frac{dy}{dx}}$ is positive when the function given by $y = u(x)$ is increasing, and $-w'(y)$ is positive when the function given by $y = u(x)$ is decreasing, we can combine the two cases by writing

$$g(y) = f[w(y)] \cdot |w'(y)|$$

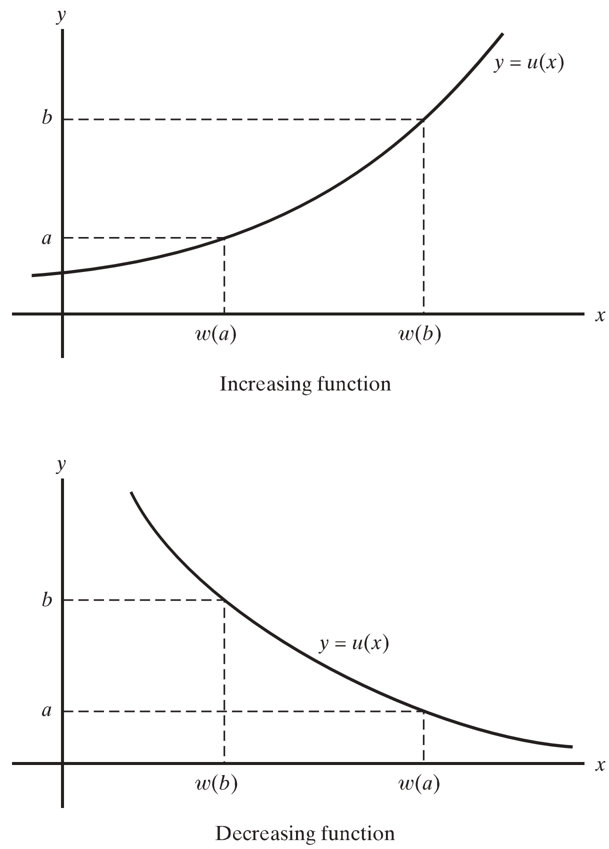

Figure 3. Diagrams for proof of Theorem 1.

EXAMPLE 6

If X has the exponential distribution given by

$$f(x) = \begin{cases} e^{-x} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

find the probability density of the random variable $Y = \sqrt{X}$.

Solution

The equation $y = \sqrt{x}$, relating the values of X and Y, has the unique inverse $x = y^2$, which yields $w'(y) = \dfrac{dx}{dy} = 2y$. Therefore,

$$g(y) = e^{-y^2}|2y| = 2ye^{-y^2}$$

for $y > 0$ in accordance with Theorem 1. Since the probability of getting a value of Y less than or equal to 0, like the probability of getting a value of X less than or equal to 0, is zero, it follows that the probability density of Y is given by

$$g(y) = \begin{cases} 2ye^{-y^2} & \text{for } y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

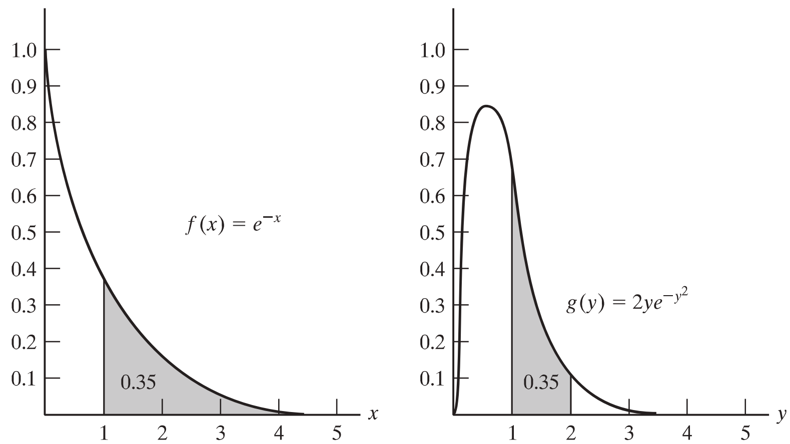

Figure 4. Diagrams for Example 6.

The two diagrams of Figure 4 illustrate what happened in this example when we transformed from X to Y. As in the discrete case (for instance, Example 4), the probabilities remain the same, but they pertain to different values (intervals of values) of the respective random variables. In the diagram on the left, the 0.35 probability pertains to the event that X will take on a value on the interval from 1 to 4, and in the diagram on the right, the 0.35 probability pertains to the event that Y will take on a value on the interval from 1 to 2.

EXAMPLE 7

If the double arrow of Figure 5 is spun so that the random variable Θ has the uniform density

$$f(\theta) = \begin{cases} \frac{1}{\pi} & \text{for } -\frac{\pi}{2} < \theta < \frac{\pi}{2} \\ 0 & \text{elsewhere} \end{cases}$$

determine the probability density of X, the abscissa of the point on the x-axis to which the arrow will point.

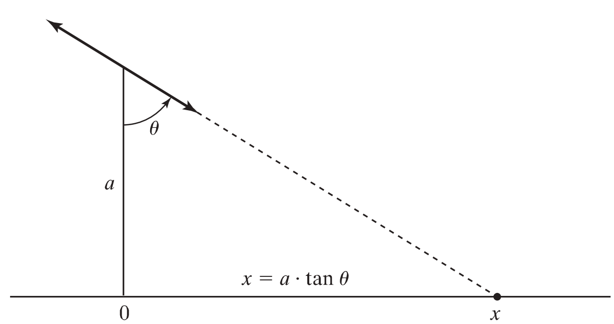

Figure 5. Diagram for Example 7.