Example 5.1-7 Let the distribution of X be binomial with parameters n and p. Since X has a discrete distribution, $Y = u(X)$ will also have a discrete distribution, with the same probabilities as those in the support of X. For example, with $n = 3, p = 1/4$, and $Y = X^2$, we have

$$g(y) = \binom{3}{\sqrt{y}}\left(\frac{1}{4}\right)^{\sqrt{y}}\left(\frac{3}{4}\right)^{3-\sqrt{y}}, \qquad y = 0, 1, 4, 9.$$

Exercises

5.1-1. Let X have the pdf $f(x) = 4x^3$, $0 < x < 1$. Find the pdf of $Y = X^2$.

5.1-2. Let X have the pdf $f(x) = xe^{-x^2/2}$, $0 < x < \infty$. Find the pdf of $Y = X^2$.

5.1-3. Let X have a gamma distribution with $\alpha = 3$ and $\theta = 2$. Determine the pdf of $Y = \sqrt{X}$.

5.1-4. The pdf of X is $f(x) = 2x$, $0 < x < 1$.

(a) Find the cdf of X.

(b) Describe how an observation of X can be simulated.

(c) Simulate 10 observations of X.

5.1-5. The pdf of X is $f(x) = \theta x^{\theta-1}$, $0 < x < 1, 0 < \theta < \infty$. Let $Y = -2\theta \ln X$. How is Y distributed?

5.1-6. Let X have a **logistic distribution** with pdf

$$f(x) = \frac{e^{-x}}{(1+e^{-x})^2}, \qquad -\infty < x < \infty.$$

Show that

$$Y = \frac{1}{1+e^{-X}}$$

has a $U(0, 1)$ distribution.

5.1-7. A sum of \$50,000 is invested at a rate R, selected from a uniform distribution on the interval (0.03, 0.07). Once R is selected, the sum is compounded instantaneously for a year, so that $X = 50000\,e^R$ dollars is the amount at the end of that year.

(a) Find the cdf and pdf of X.

(b) Verify that $X = 50000\,e^R$ is defined correctly if the compounding is done instantaneously. HINT: Divide the year into n equal parts, calculate the value of the amount at the end of each part, and then take the limit as $n \to \infty$.

5.1-8. The lifetime (in years) of a manufactured product is $Y = 5X^{0.7}$, where X has an exponential distribution with mean 1. Find the cdf and pdf of Y.

5.1-9. Statisticians frequently use the **extreme value distribution** given by the cdf

$$F(x) = 1 - \exp\left[-e^{(x-\theta_1)/\theta_2}\right], \qquad -\infty < x < \infty.$$

A simple case is when $\theta_1 = 0$ and $\theta_2 = 1$, giving

$$F(x) = 1 - \exp\left[-e^{x}\right], \qquad -\infty < x < \infty.$$

Let $Y = e^X$ or $X = \ln Y$; then the support of Y is $0 < y < \infty$.

(a) Show that the distribution of Y is exponential when $\theta_1 = 0$ and $\theta_2 = 1$.

(b) Find the cdf and the pdf of Y when $\theta_1 \neq 0$ and $\theta_2 > 0$.

(c) Let $\theta_1 = \ln\beta$ and $\theta_2 = 1/\alpha$ in the cdf and pdf of Y. What is this distribution?

(d) As suggested by its name, the extreme value distribution can be used to model the longest home run, the deepest mine, the greatest flood, and so on. Suppose the length X (in feet) of the maximum of someone's home runs was modeled by an extreme value distribution with $\theta_1 = 550$ and $\theta_2 = 25$. What is the probability that X exceeds 500 feet?

5.1-10. Let X have the uniform distribution $U(-1, 3)$. Find the pdf of $Y = X^2$.

5.1-11. Let X have a Cauchy distribution. Find

(a) $P(X > 1)$.

(b) $P(X > 5)$.

(c) $P(X > 10)$.

5.1-12. Let $f(x) = 1/[\pi(1+x^2)]$, $-\infty < x < \infty$, be the pdf of the Cauchy random variable X. Show that $E(X)$ does not exist.

5.1-13. If the distribution of X is $N(\mu, \sigma^2)$, then $M(t) = E(e^{tX}) = \exp(\mu t + \sigma^2 t^2/2)$. We then say that $Y = e^X$ has a **lognormal distribution** because $X = \ln Y$.

(a) Show that the pdf of Y is

$$g(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp[-(\ln y - \mu)^2/2\sigma^2], \quad 0 < y < \infty.$$

(b) Using $M(t)$, find **(i)** $E(Y) = E(e^X) = M(1)$, **(ii)** $E(Y^2) = E(e^{2X}) = M(2)$, and **(iii)** $\mathrm{Var}(Y)$.

5.1-14. Let X be $N(0,1)$. Find the pdf of $Y = |X|$, a distribution that is often called the **half-normal**. HINT: Here $y \in S_y = \{y : 0 < y < \infty\}$. Consider the two transformations $x_1 = -y, -\infty < x_1 < 0$, and $x_2 = y, 0 < y < \infty$.

5.1-15. Let $Y = X^2$.

(a) Find the pdf of Y when the distribution of X is $N(0,1)$.

(b) Find the pdf of Y when the pdf of X is $f(x) = (3/2)x^2, \ -1 < x < 1$.

5.2 TRANSFORMATIONS OF TWO RANDOM VARIABLES

In Section 5.1, we considered the transformation of one random variable X with pdf $f(x)$. In particular, in the continuous case, if $Y = u(X)$ was an increasing or decreasing function of X, with inverse $X = v(Y)$, then the pdf of Y was

$$g(y) = |v'(y)|\, f[v(y)], \qquad c < y < d,$$

where the support $c < y < d$ corresponds to the support of X, say, $a < x < b$, through the transformation $x = v(y)$.

There is one note of warning here: If the function $Y = u(X)$ does not have a single-valued inverse, the determination of the distribution of Y will not be as simple. As a matter of fact, we did consider two examples in Section 5.1 in which there were two inverse functions, and we exercised special care in those examples. Here, we will not consider problems with many inverses; however, such a warning is nonetheless appropriate.

When two random variables are involved, many interesting problems can result. In the case of a single-valued inverse, the rule is about the same as that in the one-variable case, with the derivative being replaced by the Jacobian. That is, if X_1 and X_2 are two continuous-type random variables with joint pdf $f(x_1, x_2)$, and if $Y_1 = u_1(X_1, X_2)$, $Y_2 = u_2(X_1, X_2)$ has the single-valued inverse $X_1 = v_1(Y_1, Y_2)$, $X_2 = v_2(Y_1, Y_2)$, then the joint pdf of Y_1 and Y_2 is

$$g(y_1, y_2) = |J|\, f\left[v_1(y_1, y_2), v_2(y_1, y_2)\right], \qquad (y_1, y_2) \in S_Y,$$

where the Jacobian J is the determinant

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\[2ex] \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix}.$$

Of course, we find the support S_Y of Y_1, Y_2 by considering the mapping of the support S_X of X_1, X_2 under the transformation $y_1 = u_1(x_1, x_2)$, $y_2 = u_2(x_1, x_2)$. This method of finding the distribution of Y_1 and Y_2 is called the **change-of-variables technique**.

It is often the mapping of the support S_X of X_1, X_2 into that (say, S_Y) of Y_1, Y_2 which causes the biggest challenge. That is, in most cases, it is easy to solve for x_1 and x_2 in terms of y_1 and y_2, say,

$$x_1 = v_1(y_1, y_2), \qquad x_2 = v_2(y_1, y_2),$$

it follows that the joint pdf of Z_1 and Z_2 is

$$g(z_1, z_2) = \left|\frac{-1}{2\pi}\, e^{-q/2}\right| (1)$$

$$= \frac{1}{2\pi} \exp\left(-\frac{z_1^2 + z_2^2}{2}\right), \qquad -\infty < z_1 < \infty, \quad -\infty < z_2 < \infty.$$

Note that there is some difficulty with the definition of this transformation, particularly when $z_1 = 0$. However, these difficulties occur at events with probability zero and hence cause no problems. (See Exercise 5.2-15.) Summarizing, from two independent $U(0,1)$ random variables we have generated two independent $N(0,1)$ random variables through this **Box–Muller transformation**. ■

Exercises

5.2-1. Let X_1, X_2 denote two independent random variables, each with a $\chi^2(2)$ distribution. Find the joint pdf of $Y_1 = X_1$ and $Y_2 = X_2 + X_1$. Note that the support of Y_1, Y_2 is $0 < y_1 < y_2 < \infty$. Also, find the marginal pdf of each of Y_1 and Y_2. Are Y_1 and Y_2 independent?

5.2-2. Let X_1 and X_2 be independent chi-square random variables with r_1 and r_2 degrees of freedom, respectively. Let $Y_1 = (X_1/r_1)/(X_2/r_2)$ and $Y_2 = X_2$.

(a) Find the joint pdf of Y_1 and Y_2.

(b) Determine the marginal pdf of Y_1 and show that Y_1 has an F distribution. (This is another, but equivalent, way of finding the pdf of F.)

5.2-3. Find the mean and the variance of an F random variable with r_1 and r_2 degrees of freedom by first finding $E(U)$, $E(1/V)$, $E(U^2)$, and $E(1/V^2)$.

5.2-4. Let the distribution of W be $F(9, 24)$. Find the following:

(a) $F_{0.05}(9, 24)$.

(b) $F_{0.95}(9, 24)$.

(c) $P(0.277 \le W \le 2.70)$.

5.2-5. Let the distribution of W be $F(8, 4)$. Find the following:

(a) $F_{0.01}(8, 4)$.

(b) $F_{0.99}(8, 4)$.

(c) $P(0.198 \le W \le 8.98)$.

5.2-6. Let X_1 and X_2 have independent gamma distributions with parameters α, θ and β, θ, respectively. Let $W = X_1/(X_1 + X_2)$. Use a method similar to that given in the derivation of the F distribution (Example 5.2-4) to show that the pdf of W is

$$g(w) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, w^{\alpha-1}(1-w)^{\beta-1}, \qquad 0 < w < 1.$$

We say that W has a beta distribution with parameters α and β. (See Example 5.2-3.)

5.2-7. Let X_1 and X_2 be independent chi-square random variables with r_1 and r_2 degrees of freedom, respectively. Show that

(a) $U = X_1/(X_1 + X_2)$ has a beta distribution with $\alpha = r_1/2$ and $\beta = r_2/2$.

(b) $V = X_2/(X_1 + X_2)$ has a beta distribution with $\alpha = r_2/2$ and $\beta = r_1/2$.

5.2-8. Let X have a beta distribution with parameters α and β. (See Example 5.2-3.)

(a) Show that the mean and variance of X are, respectively,

$$\mu = \frac{\alpha}{\alpha+\beta} \qquad \text{and} \qquad \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}.$$

(b) Show that when $\alpha > 1$ and $\beta > 1$, the mode is at $x = (\alpha - 1)/(\alpha + \beta - 2)$.

5.2-9. Determine the constant c such that $f(x) = cx^3(1-x)^6$, $0 < x < 1$, is a pdf.

5.2-10. When α and β are integers and $0 < p < 1$, we have

$$\int_0^p \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, y^{\alpha-1}(1-y)^{\beta-1}\, dy = \sum_{y=\alpha}^{n} \binom{n}{y} p^y (1-p)^{n-y},$$

where $n = \alpha + \beta - 1$. Verify this formula when $\alpha = 4$ and $\beta = 3$. Hint: Integrate the left member by parts several times.

5.2-11. Evaluate

$$\int_0^{0.4} \frac{\Gamma(7)}{\Gamma(4)\Gamma(3)}\, y^3(1-y)^2\, dy$$

(a) Using integration.

(b) Using the result of Exercise 5.2-10.

sample observations, $X_1, X_2, \ldots, X_n$, that does not have any unknown parameters is called a **statistic**, so here $\overline{X}$ is a statistic and also an **estimator** of the distribution mean μ. Another important statistic is the **sample variance**

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

and later we find that S^2 is an estimator of σ^2.

Exercises

5.3-1. Let X_1 and X_2 be independent Poisson random variables with respective means $\lambda_1 = 2$ and $\lambda_2 = 3$. Find

(a) $P(X_1 = 3, X_2 = 5)$.

(b) $P(X_1 + X_2 = 1)$.
HINT. Note that this event can occur if and only if $\{X_1 = 1, X_2 = 0\}$ or $\{X_1 = 0, X_2 = 1\}$.

5.3-2. Let X_1 and X_2 be independent random variables with respective binomial distributions $b(3, 1/2)$ and $b(5, 1/2)$. Determine

(a) $P(X_1 = 2, X_2 = 4)$.

(b) $P(X_1 + X_2 = 7)$.

5.3-3. Let X_1 and X_2 be independent random variables with probability density functions $f_1(x_1) = 2x_1$, $0 < x_1 < 1$, and $f_2(x_2) = 4x_2^3$, $0 < x_2 < 1$, respectively. Compute

(a) $P(0.5 < X_1 < 1 \text{ and } 0.4 < X_2 < 0.8)$.

(b) $E\left(X_1^2 X_2^3\right)$.

5.3-4. Let X_1 and X_2 be a random sample of size $n = 2$ from the exponential distribution with pdf $f(x) = 2e^{-2x}$, $0 < x < \infty$. Find

(a) $P(0.5 < X_1 < 1.0, 0.7 < X_2 < 1.2)$.

(b) $E[X_1(X_2 - 0.5)^2]$.

5.3-5. Let X_1 and X_2 be observations of a random sample of size $n = 2$ from a distribution with pmf $f(x) = x/6$, $x = 1, 2, 3$. Then find the pmf of $Y = X_1 + X_2$. Determine the mean and the variance of the sum in two ways.

5.3-6. Let X_1 and X_2 be a random sample of size $n = 2$ from a distribution with pdf $f(x) = 6x(1 - x)$, $0 < x < 1$. Find the mean and the variance of $Y = X_1 + X_2$.

5.3-7. The distributions of incomes in two cities follow the two Pareto-type pdfs

$$f(x) = \frac{2}{x^3}, \quad 1 < x < \infty, \quad \text{and} \quad g(y) = \frac{3}{y^4}, \quad 1 < y < \infty,$$

respectively. Here one unit represents \$20,000. One person with income is selected at random from each city. Let X and Y be their respective incomes. Compute $P(X < Y)$.

5.3-8. Suppose two independent claims are made on two insured homes, where each claim has pdf

$$f(x) = \frac{4}{x^5}, \qquad 1 < x < \infty,$$

in which the unit is \$1000. Find the expected value of the larger claim. HINT: If X_1 and X_2 are the two independent claims and $Y = \max(X_1, X_2)$, then

$$G(y) = P(Y \le y) = P(X_1 \le y)P(X_2 \le y) = [P(X \le y)]^2.$$

Find $g(y) = G'(y)$ and $E(Y)$.

5.3-9. Let X_1, X_2 be a random sample of size $n = 2$ from a distribution with pdf $f(x) = 3x^2$, $0 < x < 1$. Determine

(a) $P(\max X_i < 3/4) = P(X_1 < 3/4, X_2 < 3/4)$.

(b) The mean and the variance of $Y = X_1 + X_2$.

5.3-10. Let X_1, X_2, X_3 denote a random sample of size $n = 3$ from a distribution with the geometric pmf

$$f(x) = \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{x-1}, \qquad x = 1, 2, 3, \ldots.$$

(a) Compute $P(X_1 = 1, X_2 = 3, X_3 = 1)$.

(b) Determine $P(X_1 + X_2 + X_3 = 5)$.

(c) If Y equals the maximum of X_1, X_2, X_3, find

$$P(Y \le 2) = P(X_1 \le 2)P(X_2 \le 2)P(X_3 \le 2).$$

5.3-11. Let X_1, X_2, X_3 be three independent random variables with binomial distributions $b(4, 1/2)$, $b(6, 1/3)$, and $b(12, 1/6)$, respectively. Find

(a) $P(X_1 = 2, X_2 = 2, X_3 = 5)$.

(b) $E(X_1 X_2 X_3)$.

(c) The mean and the variance of $Y = X_1 + X_2 + X_3$.

5.3-12. Let X_1, X_2, X_3 be a random sample of size $n = 3$ from the exponential distribution with pdf $f(x) = e^{-x}$, $0 < x < \infty$. Find

$$P(1 < \min X_i) = P(1 < X_1, 1 < X_2, 1 < X_3).$$