4.117. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a distribution with mean μ and variance σ^2. Show that $E(S^2) = (n-1)\sigma^2/n$, where S^2 is the variance of the random sample.

Hint: Write $S^2 = (1/n)\sum_1^n (X_i - \mu)^2 - (\bar{X} - \mu)^2$.

4.118. Let X_1 and X_2 be independent random variables with nonzero variances. Find the correlation coefficient of $Y = X_1X_2$ and X_1 in terms of the means and variances of X_1 and X_2.

4.119. Let X_1 and X_2 have a joint distribution with parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$, and ρ. Find the correlation coefficient of the linear functions $Y = a_1X_1 + a_2X_2$ and $Z = b_1X_1 + b_2X_2$ in terms of the real constants a_1, a_2, b_1, b_2, and the parameters of the distribution.

4.120. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a distribution which has mean μ and variance σ^2. Use Chebyshev's inequality to show, for every $\epsilon > 0$, that $\lim_{n\to\infty} \Pr(|\bar{X} - \mu| < \epsilon) = 1$; this is another form of the law of large numbers.

4.121. Let X_1, X_2, and X_3 be random variables with equal variances but with correlation coefficients $\rho_{12} = 0.3$, $\rho_{13} = 0.5$, and $\rho_{23} = 0.2$. Find the correlation coefficient of the linear functions $Y = X_1 + X_2$ and $Z = X_2 + X_3$.

4.122. Find the variance of the sum of 10 random variables if each has variance 5 and if each pair has correlation coefficient 0.5.

4.123. Let X and Y have the parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$, and ρ. Show that the correlation coefficient of X and $[Y - \rho(\sigma_2/\sigma_1)X]$ is zero.

4.124. Let X_1 and X_2 have a bivariate normal distribution with parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$, and ρ. Compute the means, the variances, and the correlation coefficient of $Y_1 = \exp(X_1)$ and $Y_2 = \exp(X_2)$.

Hint: Various moments of Y_1 and Y_2 can be found by assigning appropriate values to t_1 and t_2 in $E[\exp(t_1X_1 + t_2X_2)]$.

4.125. Let X be $N(\mu, \sigma^2)$ and consider the transformation $X = \ln Y$ or, equivalently, $Y = e^X$.

(a) Find the mean and the variance of Y by first determining $E(e^X)$ and $E[(e^X)^2]$.
Hint: Use the m.g.f. of X.

(b) Find the p.d.f. of Y. This is the p.d.f. of the *lognormal distribution.*

4.126. Let X_1 and X_2 have a trinomial distribution with parameters n, p_1, p_2.

(a) What is the distribution of $Y = X_1 + X_2$?

(b) From the equality $\sigma_Y^2 = \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2$, once again determine the correlation coefficient ρ of X_1 and X_2.

4.127. Let $Y_1 = X_1 + X_2$ and $Y_2 = X_2 + X_3$, where X_1, X_2, and X_3 are three independent random variables. Find the joint m.g.f. and the correlation coefficient of Y_1 and Y_2 provided that:

(a) X_i has a Poisson distribution with mean μ_i, $i = 1, 2, 3$.

(b) X_i is $N(\mu_i, \sigma_i^2)$, $i = 1, 2, 3$.

4.128. Let $X_1, \ldots, X_n$ be random variables that have means $\mu_1, \ldots, \mu_n$ and variances $\sigma_1^2, \ldots, \sigma_n^2$. Let ρ_{ij}, $i \neq j$, denote the correlation coefficient of X_i and X_j. Let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be real constants. Show that the covariance of $Y = \sum_{i=1}^n a_iX_i$ and $Z = \sum_{j=1}^n b_jX_j$ is $\sum_{j=1}^n \sum_{i=1}^n a_ib_j\sigma_i\sigma_j\rho_{ij}$, where $\rho_{ii} = 1$, $i = 1, 2, \ldots, n$.

*4.10 The Multivariate Normal Distribution

We have studied in some detail normal distributions of one random variable. In this section we investigate a joint distribution of n random variables that will be called a *multivariate normal* distribution. This investigation assumes that the student is familiar with elementary matrix algebra, with real symmetric quadratic forms, and with orthogonal transformations. Henceforth, the expression *quadratic form* means a quadratic form in a prescribed number of variables whose matrix is real and symmetric. All symbols that represent matrices will be set in boldface type.

Let $\mathbf{A}$ denote an $n \times n$ real symmetric matrix which is positive definite. Let $\boldsymbol{\mu}$ denote the $n \times 1$ matrix such that $\boldsymbol{\mu}'$, the transpose of $\boldsymbol{\mu}$, is $\boldsymbol{\mu}' = [\mu_1, \mu_2, \ldots, \mu_n]$, where each μ_i is a real constant. Finally, let $\mathbf{x}$ denote the $n \times 1$ matrix such that $\mathbf{x}' = [x_1, x_2, \ldots, x_n]$. We shall show that if C is an appropriately chosen positive constant, the nonnegative function

$$f(x_1, x_2, \ldots, x_n) = C \exp\left[-\frac{(\mathbf{x} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{x} - \boldsymbol{\mu})}{2}\right],$$

$$-\infty < x_i < \infty, \quad i = 1, 2, \ldots, n,$$

is a joint p.d.f. of n random variables $X_1, X_2, \ldots, X_n$ that are of the continuous type. Thus we need to show that

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \cdots dx_n = 1. \tag{1}$$

Let $\mathbf{t}$ denote the $n \times 1$ matrix such that $\mathbf{t}' = [t_1, t_2, \ldots, t_n]$, where $t_1, t_2, \ldots, t_n$ are arbitrary real numbers. We shall evaluate the integral

$$C \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[\mathbf{t}'\mathbf{x} - \frac{(\mathbf{x}-\boldsymbol{\mu})'\mathbf{A}(\mathbf{x}-\boldsymbol{\mu})}{2}\right] dx_1 \cdots dx_n, \qquad (2)$$

and then we shall subsequently set $t_1 = t_2 = \cdots = t_n = 0$, and thus establish Equation (1). First, we change the variables of integration in integral (2) from $x_1, x_2, \ldots, x_n$ to $y_1, y_2, \ldots, y_n$ by writing $\mathbf{x} - \boldsymbol{\mu} = \mathbf{y}$, where $\mathbf{y}' = [y_1, y_2, \ldots, y_n]$. The Jacobian of the transformation is one and the n-dimensional x-space is mapped onto an n-dimensional y-space, so that integral (2) may be written as

$$C \exp(\mathbf{t}'\boldsymbol{\mu}) \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(\mathbf{t}'\mathbf{y} - \frac{\mathbf{y}'\mathbf{A}\mathbf{y}}{2}\right) dy_1 \cdots dy_n. \qquad (3)$$

Because the real symmetric matrix $\mathbf{A}$ is positive definite, the n characteristic numbers (proper values, latent roots, or eigenvalues) $a_1, a_2, \ldots, a_n$ of $\mathbf{A}$ are positive. There exists an appropriately chosen $n \times n$ real orthogonal matrix $\mathbf{L}$ ($\mathbf{L}' = \mathbf{L}^{-1}$, where $\mathbf{L}^{-1}$ is the inverse of $\mathbf{L}$) such that

$$\mathbf{L}'\mathbf{A}\mathbf{L} = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix},$$

for a suitable ordering of $a_1, a_2, \ldots, a_n$. We shall sometimes write $\mathbf{L}'\mathbf{A}\mathbf{L} = \operatorname{diag}[a_1, a_2, \ldots, a_n]$. In integral (3), we shall change the variables of integration from $y_1, y_2, \ldots, y_n$ to $z_1, z_2, \ldots, z_n$ by writing $\mathbf{y} = \mathbf{L}\mathbf{z}$, where $\mathbf{z}' = [z_1, z_2, \ldots, z_n]$. The Jacobian of the transformation is the determinant of the orthogonal matrix $\mathbf{L}$. Since $\mathbf{L}'\mathbf{L} = \mathbf{I}_n$, where $\mathbf{I}_n$ is the unit matrix of order n, we have the determinant $|\mathbf{L}'\mathbf{L}| = 1$ and $|\mathbf{L}|^2 = 1$. Thus the absolute value of the Jacobian is one. Moreover, the n-dimensional y-space is mapped onto an n-dimensional z-space. The integral (3) becomes

$$C \exp(\mathbf{t}'\boldsymbol{\mu}) \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[\mathbf{t}'\mathbf{L}\mathbf{z} - \frac{\mathbf{z}'(\mathbf{L}'\mathbf{A}\mathbf{L})\mathbf{z}}{2}\right] dz_1 \cdots dz_n. \qquad (4)$$

It is computationally convenient to write, momentarily, $\mathbf{t}'\mathbf{L} = \mathbf{w}'$, where $\mathbf{w}' = [w_1, w_2, \ldots, w_n]$. Then

$$\exp[\mathbf{t}'\mathbf{L}\mathbf{z}] = \exp[\mathbf{w}'\mathbf{z}] = \exp\left(\sum_1^n w_i z_i\right).$$

Moreover,

$$\exp\left[-\frac{\mathbf{z}'(\mathbf{L}'\mathbf{A}\mathbf{L})\mathbf{z}}{2}\right] = \exp\left[-\frac{\sum_1^n a_i z_i^2}{2}\right].$$

Then integral (4) may be written as the product of n integrals in the following manner:

$$C \exp(\mathbf{w}'\mathbf{L}'\boldsymbol{\mu}) \prod_{i=1}^{n} \left[\int_{-\infty}^{\infty} \exp\left(w_i z_i - \frac{a_i z_i^2}{2}\right) dz_i\right]$$

$$= C \exp(\mathbf{w}'\mathbf{L}'\boldsymbol{\mu}) \prod_{i=1}^{n} \left[\sqrt{\frac{2\pi}{a_i}} \int_{-\infty}^{\infty} \frac{\exp\left(w_i z_i - \frac{a_i z_i^2}{2}\right)}{\sqrt{2\pi/a_i}} dz_i\right]. \qquad (5)$$

The integral that involves z_i can be treated as the m.g.f., with the more familiar symbol t replaced by w_i, of a distribution which is $N(0, 1/a_i)$. Thus the right-hand member of Equation (5) is equal to

$$C \exp(\mathbf{w}'\mathbf{L}'\boldsymbol{\mu}) \prod_{i=1}^{n} \left[\sqrt{\frac{2\pi}{a_i}} \exp\left(\frac{w_i^2}{2a_i}\right)\right]$$

$$= C \exp(\mathbf{w}'\mathbf{L}'\boldsymbol{\mu}) \sqrt{\frac{(2\pi)^n}{a_1 a_2 \cdots a_n}} \exp\left(\sum_1^n \frac{w_i^2}{2a_i}\right). \qquad (6)$$

Now, because $\mathbf{L}^{-1} = \mathbf{L}'$, we have

$$(\mathbf{L}'\mathbf{A}\mathbf{L})^{-1} = \mathbf{L}'\mathbf{A}^{-1}\mathbf{L} = \operatorname{diag}\left[\frac{1}{a_1}, \frac{1}{a_2}, \ldots, \frac{1}{a_n}\right].$$

Thus

$$\sum_1^n \frac{w_i^2}{a_i} = \mathbf{w}'(\mathbf{L}'\mathbf{A}^{-1}\mathbf{L})\mathbf{w} = (\mathbf{L}\mathbf{w})'\mathbf{A}^{-1}(\mathbf{L}\mathbf{w}) = \mathbf{t}'\mathbf{A}^{-1}\mathbf{t}.$$

Moreover, the determinant $|\mathbf{A}^{-1}|$ of $\mathbf{A}^{-1}$ is

$$|\mathbf{A}^{-1}| = |\mathbf{L}'\mathbf{A}^{-1}\mathbf{L}| = \frac{1}{a_1 a_2 \cdots a_n}.$$

Accordingly, the right-hand member of Equation (6), which is equal to integral (2), may be written as

$$C e^{\mathbf{t}'\boldsymbol{\mu}} \sqrt{(2\pi)^n |\mathbf{A}^{-1}|} \exp\left(\frac{\mathbf{t}'\mathbf{A}^{-1}\mathbf{t}}{2}\right). \qquad (7)$$

If, in this function, we set $t_1 = t_2 = \cdots = t_n = 0$, we have the value of the left-hand member of Equation (1). Thus we have

$$C\sqrt{(2\pi)^n|\mathbf{A}^{-1}|} = 1.$$

Accordingly, the function

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{A}^{-1}|}} \exp\left[-\frac{(\mathbf{x}-\boldsymbol{\mu})'\mathbf{A}(\mathbf{x}-\boldsymbol{\mu})}{2}\right],$$

$-\infty < x_i < \infty$, $i = 1, 2, \ldots, n$, is a joint p.d.f. of n random variables $X_1, X_2, \ldots, X_n$ that are of the continuous type. Such a p.d.f. is called a *nonsingular multivariate normal* p.d.f.

We have now proved that $f(x_1, x_2, \ldots, x_n)$ is a p.d.f. However, we have proved more than that. Because $f(x_1, x_2, \ldots, x_n)$ is a p.d.f., integral (2) is the m.g.f. $M(t_1, t_2, \ldots, t_n)$ of this joint distribution of probability. Since integral (2) is equal to function (7), the m.g.f. of the multivariate normal distribution is given by

$$M(t_1, t_2, \ldots, t_n) = \exp\left(\mathbf{t}'\boldsymbol{\mu} + \frac{\mathbf{t}'\mathbf{A}^{-1}\mathbf{t}}{2}\right).$$

Let the elements of the real, symmetric, and positive definite matrix $\mathbf{A}^{-1}$ be denoted by σ_{ij}, $i, j = 1, 2, \ldots, n$. Then

$$M(0, \ldots, 0, t_i, 0, \ldots, 0) = \exp\left(t_i\mu_i + \frac{\sigma_{ii}t_i^2}{2}\right)$$

is the m.g.f. of X_i, $i = 1, 2, \ldots, n$. Thus X_i is $N(\mu_i, \sigma_{ii})$, $i = 1, 2, \ldots, n$. Moreover, with $i \neq j$, we see that $M(0, \ldots, 0, t_i, 0, \ldots, t_j, 0, \ldots, 0)$, the m.g.f. of X_i and X_j, is equal to

$$\exp\left(t_i\mu_i + t_j\mu_j + \frac{\sigma_{ii}t_i^2 + 2\sigma_{ij}t_it_j + \sigma_{jj}t_j^2}{2}\right),$$

which is the m.g.f. of a *bivariate normal distribution*. In Exercise 4.131 the reader is asked to show that σ_{ij} is the covariance of the random variables X_i and X_j. Thus the matrix $\boldsymbol{\mu}$, where $\boldsymbol{\mu}' = [\mu_1, \mu_2, \ldots, \mu_n]$, is the matrix of the means of the random variables $X_1, \ldots, X_n$. Moreover, the elements on the principal diagonal of $\mathbf{A}^{-1}$ are, respectively, the variances $\sigma_{ii} = \sigma_i^2$, $i = 1, 2, \ldots, n$, and the elements not on the principal diagonal of $\mathbf{A}^{-1}$ are, respectively, the covariances $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$, $i \neq j$, of the random variables $X_1, X_2, \ldots, X_n$. We call the matrix $\mathbf{A}^{-1}$, which is given by

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_{nn} \end{bmatrix},$$

the *covariance matrix* of the multivariate normal distribution and henceforth we shall denote this matrix by the symbol $\mathbf{V}$. In terms of the positive definite covariance matrix $\mathbf{V}$, the multivariate normal p.d.f. is written

$$\frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{V}|}} \exp\left[-\frac{(\mathbf{x}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2}\right], \qquad -\infty < x_i < \infty,$$

$i = 1, 2, \ldots, n$, and the m.g.f. of this distribution is given by

$$\exp\left(\mathbf{t}'\boldsymbol{\mu} + \frac{\mathbf{t}'\mathbf{V}\mathbf{t}}{2}\right)$$

for all real values of $\mathbf{t}$.

Note that this m.g.f. equals the product of n functions, where the first is a function of t_1 alone, the second is a function of t_2 alone, and so on, if and only if $\mathbf{V}$ is a diagonal matrix. This condition, $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j = 0$, means $\rho_{ij} = 0$, $i \neq j$. That is, the multivariate normal random variables are independent if and only if $\rho_{ij} = 0$ for all $i \neq j$.

Example 1. Let $X_1, X_2, \ldots, X_n$ have a multivariate normal distribution with matrix $\boldsymbol{\mu}$ of means and positive definite covariance matrix $\mathbf{V}$. If we let $\mathbf{X}' = [X_1, X_2, \ldots, X_n]$, then the m.g.f. $M(t_1, t_2, \ldots, t_n)$ of this joint distribution of probability is

$$E(e^{\mathbf{t}'\mathbf{X}}) = \exp\left(\mathbf{t}'\boldsymbol{\mu} + \frac{\mathbf{t}'\mathbf{V}\mathbf{t}}{2}\right). \tag{8}$$

Consider a linear function Y of $X_1, X_2, \ldots, X_n$ which is defined by $Y = \mathbf{c}'\mathbf{X} = \sum_1^n c_iX_i$, where $\mathbf{c}' = [c_1, c_2, \ldots, c_n]$ and the several c_i are real and not all zero. We wish to find the p.d.f. of Y. The m.g.f. $m(t)$ of the distribution of Y is given by

$$m(t) = E(e^{tY}) = E(e^{t\mathbf{c}'\mathbf{X}}).$$

Now the expectation (8) exists for all real values of $\mathbf{t}$. Thus we can replace $\mathbf{t}'$ in expectation (8) by $t\mathbf{c}'$ and obtain

$$m(t) = \exp\left(t\mathbf{c}'\boldsymbol{\mu} + \frac{\mathbf{c}'\mathbf{V}\mathbf{c}t^2}{2}\right).$$

Thus the random variable Y is $N(\mathbf{c}'\boldsymbol{\mu}, \mathbf{c}'\mathbf{V}\mathbf{c})$.

EXERCISES

4.129. Let $X_1, X_2, \ldots, X_n$ have a multivariate normal distribution with positive definite covariance matrix $\mathbf{V}$. Prove that these random variables are mutually independent if and only if $\mathbf{V}$ is a diagonal matrix.

4.130. Let $n = 2$ and take

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

Determine $|\mathbf{V}|$, $\mathbf{V}^{-1}$, and $(\mathbf{x} - \boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x} - \boldsymbol{\mu})$. Compare the bivariate normal p.d.f. of Section 3.5 with this multivariate normal p.d.f. when $n = 2$.

4.131. Let $m(t_i, t_j)$ represent the m.g.f. of X_i and X_j as given in the text. Show that

$$\frac{\partial^2 m(0, 0)}{\partial t_i \partial t_j} - \left[\frac{\partial m(0, 0)}{\partial t_i}\right]\left[\frac{\partial m(0, 0)}{\partial t_j}\right] = \sigma_{ij};$$

that is, prove that the covariance of X_i and X_j is σ_{ij}, which appears in that formula for $m(t_i, t_j)$.

4.132. Let $X_1, X_2, \ldots, X_n$ have a multivariate normal distribution, where $\boldsymbol{\mu}$ is the matrix of the means and $\mathbf{V}$ is the positive definite covariance matrix. Let $Y = \mathbf{c}'\mathbf{X}$ and $Z = \mathbf{d}'\mathbf{X}$, where $\mathbf{X}' = [X_1, \ldots, X_n]$, $\mathbf{c}' = [c_1, \ldots, c_n]$, and $\mathbf{d}' = [d_1, \ldots, d_n]$ are real matrices.
(a) Find $m(t_1, t_2) = E(e^{t_1 Y + t_2 Z})$ to see that Y and Z have a bivariate normal distribution.
(b) Prove that Y and Z are independent if and only if $\mathbf{c}'\mathbf{V}\mathbf{d} = 0$.
(c) If $X_1, X_2, \ldots, X_n$ are independent random variables which have the same variance σ^2, show that the necessary and sufficient condition of part (b) becomes $\mathbf{c}'\mathbf{d} = 0$.

4.133. Let $\mathbf{X}' = [X_1, X_2, \ldots, X_n]$ have the multivariate normal distribution of Exercise 4.132. Consider the p linear functions of $X_1, \ldots, X_n$ defined by $\mathbf{W} = \mathbf{B}\mathbf{X}$, where $\mathbf{W}' = [W_1, \ldots, W_p]$, $p \le n$, and $\mathbf{B}$ is a $p \times n$ real matrix of rank p. Find $m(v_1, \ldots, v_p) = E(e^{\mathbf{v}'\mathbf{W}})$, where $\mathbf{v}'$ is a real matrix $[v_1, \ldots, v_p]$, to see that $W_1, \ldots, W_p$ have a p-variate normal distribution which has $\mathbf{B}\boldsymbol{\mu}$ for the means and $\mathbf{B}\mathbf{V}\mathbf{B}'$ for the covariance matrix.

4.134. Let $\mathbf{X}' = [X_1, X_2, \ldots, X_n]$ have the n-variate normal distribution of Exercise 4.132. Show that $X_1, X_2, \ldots, X_p$, $p < n$, have a p-variate normal distribution. What submatrix of $\mathbf{V}$ is the covariance matrix of $X_1, X_2, \ldots, X_p$?
Hint: In the m.g.f. $M(t_1, t_2, \ldots, t_n)$ of $X_1, X_2, \ldots, X_n$, let $t_{p+1} = \cdots = t_n = 0$.

ADDITIONAL EXERCISES

4.135. If X has the p.d.f. $f(x) = \frac{1}{3}$, $-1 < x < 2$, zero elsewhere, find the p.d.f. of $Y = X^4$.

4.136. The continuous random variable X has a p.d.f. given by $f(x) = 1$, $0 < x < 1$, zero elsewhere. The random variable Y is such that $Y = -2 \ln X$. What is the distribution of Y? What are the mean and the variance of Y?

4.137. Let X_1, X_2 be a random sample of size $n = 2$ from a Poisson distribution with mean μ. If $\Pr(X_1 + X_2 = 3) = (\frac{32}{3})e^{-4}$, compute $\Pr(X_1 = 2, X_2 = 4)$.

4.138. Let $X_1, X_2, \ldots, X_{25}$ be a random sample of size $n = 25$ from a distribution with p.d.f. $f(x) = 3/x^4$, $1 < x < \infty$, zero elsewhere. Let Y equal the number of these X values less than or equal to 2. What is the distribution of Y?

4.139. Find the probability that the range of a random sample of size 3 from the uniform distribution over the interval $(-5, 5)$ is less than 7.

4.140. Let $Y_1 < Y_2 < Y_3$ be the order statistics of a sample of size 3 from a distribution having p.d.f. $f(x) = \frac{1}{3}$, $-1 < x < 2$, zero elsewhere. Determine $\Pr[-\frac{1}{2} < Y_2 < \frac{1}{2}]$.

4.141. Let X and Y be random variables so that $Z = X - 2Y$ has variance equal to 28. If $\sigma_X^2 = 4$ and $\rho_{XY} = \frac{1}{2}$, find the variance σ_Y^2 of Y.

4.142. Let $Y_1 < Y_2 < Y_3 < Y_4$ be the order statistics of a random sample of size $n = 4$ from a distribution with p.d.f. $f(x) = 2(1 - x)$, $0 < x < 1$, zero elsewhere. Compute $\Pr(Y_1 < 0.1)$.

4.143. A certain job is completed in three steps in series. The means and standard deviations for the steps are (in hours):

Step	Mean	Standard Deviation
1	3	0.2
2	1	0.1
3	4	0.2

CHAPTER 10

Inferences About Normal Models

10.1 The Distributions of Certain Quadratic Forms

A homogeneous polynomial of degree 2 in n variables is called a *quadratic* form in those variables. If both the variables and the coefficients are real, the form is called a *real quadratic* form. Only real quadratic forms will be considered in this book. To illustrate, the form $X_1^2 + X_1X_2 + X_2^2$ is a quadratic form in the two variables X_1 and X_2; the form $X_1^2 + X_2^2 + X_3^2 - 2X_1X_2$ is a quadratic form in the three variables X_1, X_2, and X_3; but the form $(X_1 - 1)^2 + (X_2 - 2)^2 = X_1^2 + X_2^2 - 2X_1 - 4X_2 + 5$ is not a quadratic form in X_1 and X_2, although it is a quadratic form in the variables $X_1 - 1$ and $X_2 - 2$.

Let $\bar{X}$ and S^2 denote, respectively, the mean and the variance of a random sample $X_1, X_2, \ldots, X_n$ from an arbitrary distribution. Thus

$$nS^2 = \sum_1^n (X_i - \bar{X})^2 = \sum_1^n \left(X_i - \frac{X_1 + X_2 + \cdots + X_n}{n} \right)^2$$

is a quadratic form in the n variables $X_1, X_2, \ldots, X_n$. If the sample arises from a distribution that is $N(\mu, \sigma^2)$, we know that the random variable nS^2/σ^2 is $\chi^2(n-1)$ regardless of the value of μ. This fact proved useful in our search for a confidence interval for σ^2 when μ is unknown.

It has been seen that tests of certain statistical hypotheses require a statistic that is a quadratic form. For instance, Example 2, Section 9.2, made use of the statistic $\sum_1^n X_i^2$, which is a quadratic form in the variables $X_1, X_2, \ldots, X_n$. Later in this chapter, tests of other statistical hypotheses will be investigated, and it will be seen that functions of statistics that are quadratic forms will be needed to carry out the tests in an expeditious manner. But first we shall make a study of the distribution of certain quadratic forms in normal and independent random variables.

The following theorem will be proved in Section 10.9.

Theorem 1. *Let $Q = Q_1 + Q_2 + \cdots + Q_{k-1} + Q_k$, where $Q, Q_1, \ldots, Q_k$ are $k + 1$ random variables that are real quadratic forms in n independent random variables which are normally distributed with the means $\mu_1, \mu_2, \ldots, \mu_n$ and the same variance σ^2. Let Q/σ^2, $Q_1/\sigma^2, \ldots, Q_{k-1}/\sigma^2$ have chi-square distributions with degrees of freedom $r, r_1, \ldots, r_{k-1}$, respectively. Let Q_k be nonnegative. Then:*

(a) *$Q_1, \ldots, Q_k$ are independent, and hence*
(b) *Q_k/σ^2 has a chi-square distribution with $r - (r_1 + \cdots + r_{k-1}) = r_k$ degrees of freedom.*

Three examples illustrative of the theorem will follow. Each of these examples will deal with a distribution problem that is based on the remarks made in the subsequent paragraph.

Let the random variable X have a distribution that is $N(\mu, \sigma^2)$. Let a and b denote positive integers greater than 1 and let $n = ab$. Consider a random sample of size $n = ab$ from this normal distribution. The observations of the random sample will be denoted by the symbols

$$\begin{matrix} X_{11}, & X_{12}, & \ldots, & X_{1j}, & \ldots, & X_{1b} \\ X_{21}, & X_{22}, & \ldots, & X_{2j}, & \ldots, & X_{2b} \\ \vdots & & & & & \\ X_{i1}, & X_{i2}, & \ldots, & X_{ij}, & \ldots, & X_{ib} \\ \vdots & & & & & \\ X_{a1}, & X_{a2}, & \ldots, & X_{aj}, & \ldots, & X_{ab}. \end{matrix}$$

In this notation, the first subscript indicates the row, and the second subscript indicates the column in which the observation appears. Thus X_{ij} is in row i and column j, $i = 1, 2, \ldots, a$ and $j = 1, 2, \ldots, b$. By assumption these $n = ab$ random variables are independent, and each has the same normal distribution with mean μ and variance σ^2. Thus, if we wish, we may consider each row as being a random sample of size b from the given distribution; and we may consider each column as being a random sample of size a from the given distribution. We now define $a + b + 1$ statistics. They are

$$\bar{X}_{..} = \frac{X_{11} + \cdots + X_{1b} + \cdots + X_{a1} + \cdots + X_{ab}}{ab} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} X_{ij}}{ab},$$

$$\bar{X}_{i.} = \frac{X_{i1} + X_{i2} + \cdots + X_{ib}}{b} = \frac{\sum_{j=1}^{b} X_{ij}}{b}, \qquad i = 1, 2, \ldots, a,$$

and

$$\bar{X}_{.j} = \frac{X_{1j} + X_{2j} + \cdots + X_{aj}}{a} = \frac{\sum_{i=1}^{a} X_{ij}}{a}, \qquad j = 1, 2, \ldots, b.$$

The statistic $\bar{X}_{..}$ is the mean of the random sample of size $n = ab$; the statistics $\bar{X}_{1.}, \bar{X}_{2.}, \ldots, \bar{X}_{a.}$ are, respectively, the means of the rows; and the statistics $\bar{X}_{.1}, \bar{X}_{.2}, \ldots, \bar{X}_{.b}$ are, respectively, the means of the columns. The examples illustrative of the theorem follow.

Example 1. Consider the variance S^2 of the random sample of size $n = ab$. We have the algebraic identity

$$\begin{aligned} abS^2 &= \sum_{i=1}^{a}\sum_{j=1}^{b} (X_{ij} - \bar{X}_{..})^2 \\ &= \sum_{i=1}^{a}\sum_{j=1}^{b} [(X_{ij} - \bar{X}_{i.}) + (\bar{X}_{i.} - \bar{X}_{..})]^2 \\ &= \sum_{i=1}^{a}\sum_{j=1}^{b} (X_{ij} - \bar{X}_{i.})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b} (\bar{X}_{i.} - \bar{X}_{..})^2 \\ &\quad + 2\sum_{i=1}^{a}\sum_{j=1}^{b} (X_{ij} - \bar{X}_{i.})(\bar{X}_{i.} - \bar{X}_{..}). \end{aligned}$$

The last term of the right-hand member of this identity may be written

$$2\sum_{i=1}^{a}\left[(\bar{X}_{i.} - \bar{X}_{..})\sum_{j=1}^{b}(X_{ij} - \bar{X}_{i.})\right] = 2\sum_{i=1}^{a}[(\bar{X}_{i.} - \bar{X}_{..})(b\bar{X}_{i.} - b\bar{X}_{i.})] = 0,$$

and the term

$$\sum_{i=1}^{a}\sum_{j=1}^{b} (\bar{X}_{i.} - \bar{X}_{..})^2$$

may be written

$$b\sum_{i=1}^{a} (\bar{X}_{i.} - \bar{X}_{..})^2.$$

Thus

$$abS^2 = \sum_{i=1}^{a}\sum_{j=1}^{b} (X_{ij} - \bar{X}_{i.})^2 + b\sum_{i=1}^{a} (\bar{X}_{i.} - \bar{X}_{..})^2,$$

or, for brevity,

$$Q = Q_1 + Q_2.$$

Clearly, Q, Q_1, and Q_2 are quadratic forms in the $n = ab$ variables X_{ij}. We shall use the theorem with $k = 2$ to show that Q_1 and Q_2 are independent. Since S^2 is the variance of a random sample of size $n = ab$ from the given normal distribution, then abS^2/σ^2 has a chi-square distribution with $ab - 1$ degrees of freedom. Now

$$\frac{Q_1}{\sigma^2} = \sum_{i=1}^{a}\left[\frac{\sum_{j=1}^{b} (X_{ij} - \bar{X}_{i.})^2}{\sigma^2}\right].$$

For each fixed value of i, $\sum_{j=1}^{b} (X_{ij} - \bar{X}_{i.})^2/b$ is the variance of a random sample of size b from the given normal distribution, and, accordingly, $\sum_{j=1}^{b} (X_{ij} - \bar{X}_{i.})^2/\sigma^2$ has a chi-square distribution with $b - 1$ degrees of freedom. Because the X_{ij} are independent, Q_1/σ^2 is the sum of a independent random variables, each having a chi-square distribution with $b - 1$ degrees of freedom. Hence Q_1/σ^2 has a chi-square distribution with $a(b - 1)$ degrees of freedom. Now $Q_2 = b \sum_{i=1}^{a} (\bar{X}_{i.} - \bar{X}_{..})^2 \geq 0$. In accordance with the theorem, Q_1 and Q_2 are independent, and Q_2/σ^2 has a chi-square distribution with $ab - 1 - a(b - 1) = a - 1$ degrees of freedom.

Example 2. In abS^2 replace $X_{ij} - \bar{X}_{..}$ by $(X_{ij} - \bar{X}_{.j}) + (\bar{X}_{.j} - \bar{X}_{..})$ to obtain

$$abS^2 = \sum_{j=1}^{b} \sum_{i=1}^{a} [(X_{ij} - \bar{X}_{.j}) + (\bar{X}_{.j} - \bar{X}_{..})]^2,$$

or

$$abS^2 = \sum_{j=1}^{b} \sum_{i=1}^{a} (X_{ij} - \bar{X}_{.j})^2 + a \sum_{j=1}^{b} (\bar{X}_{.j} - \bar{X}_{..})^2,$$

or, for brevity,

$$Q = Q_3 + Q_4.$$

It is easy to show (Exercise 10.1) that Q_3/σ^2 has a chi-square distribution with $b(a - 1)$ degrees of freedom. Since $Q_4 = a \sum_{j=1}^{b} (\bar{X}_{.j} - \bar{X}_{..})^2 \geq 0$, the theorem enables us to assert that Q_3 and Q_4 are independent and that Q_4/σ^2 has a chi-square distribution with $ab - 1 - b(a - 1) = b - 1$ degrees of freedom.

Example 3. In abS^2 replace $X_{ij} - \bar{X}_{..}$ by $(\bar{X}_{i.} - \bar{X}_{..}) + (\bar{X}_{.j} - \bar{X}_{..}) + (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})$ to obtain (Exercise 10.2)

$$abS^2 = b \sum_{i=1}^{a} (\bar{X}_{i.} - \bar{X}_{..})^2 + a \sum_{j=1}^{b} (\bar{X}_{.j} - \bar{X}_{..})^2$$

$$+ \sum_{j=1}^{b} \sum_{i=1}^{a} (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2,$$

or, for brevity,

$$Q = Q_2 + Q_4 + Q_5,$$

where Q_2 and Q_4 are as defined in Examples 1 and 2. From Examples 1 and 2, Q/σ^2, Q_2/σ^2, and Q_4/σ^2 have chi-square distributions with $ab - 1$, $a - 1$, and $b - 1$ degrees of freedom, respectively. Since $Q_5 \geq 0$, the theorem asserts that Q_2, Q_4, and Q_5 are independent and that Q_5/σ^2 has a chi-square distribution with $ab - 1 - (a - 1) - (b - 1) = (a - 1)(b - 1)$ degrees of freedom.

Once these quadratic form statistics have been shown to be independent, a multiplicity of F-statistics can be defined. For instance,

$$\frac{Q_4/[\sigma^2(b-1)]}{Q_3/[\sigma^2 b(a-1)]} = \frac{Q_4/(b-1)}{Q_3/[b(a-1)]}$$

has an F-distribution with $b - 1$ and $b(a - 1)$ degrees of freedom; and

$$\frac{Q_4/[\sigma^2(b-1)]}{Q_5/[\sigma^2(a-1)(b-1)]} = \frac{Q_4/(b-1)}{Q_5/(a-1)(b-1)}$$

has an F-distribution with $b - 1$ and $(a - 1)(b - 1)$ degrees of freedom. In the subsequent sections it will be seen that some likelihood ratio tests of certain statistical hypotheses can be based on these F-statistics.

EXERCISES

10.1. In Example 2 verify that $Q = Q_3 + Q_4$ and that Q_3/σ^2 has a chi-square distribution with $b(a - 1)$ degrees of freedom.

10.2. In Example 3 verify that $Q = Q_2 + Q_4 + Q_5$.

10.3. Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution $N(\mu, \sigma^2)$. Show that

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=2}^{n} (X_i - \bar{X}')^2 + \frac{n-1}{n}(X_1 - \bar{X}')^2,$$

where $\bar{X} = \sum_{i=1}^{n} X_i/n$ and $\bar{X}' = \sum_{i=2}^{n} X_i/(n-1)$.

Hint: Replace $X_i - \bar{X}$ by $(X_i - \bar{X}') - (X_1 - \bar{X}')/n$. Show that $\sum_{i=2}^{n} (X_i - \bar{X}')^2/\sigma^2$ has a chi-square distribution with $n - 2$ degrees of freedom. Prove that the two terms in the right-hand member are independent. What then is the distribution of

$$\frac{[(n-1)/n](X_1 - \bar{X}')^2}{\sigma^2}?$$

10.4. Let X_{ijk}, $i = 1, \ldots, a$; $j = 1, \ldots, b$; $k = 1, \ldots, c$, be a random sample of size $n = abc$ from a normal distribution $N(\mu, \sigma^2)$. Let $\bar{X}_{...} = \sum_{k=1}^{c} \sum_{j=1}^{b} \sum_{i=1}^{a} X_{ijk}/n$ and $\bar{X}_{i..} = \sum_{k=1}^{c} \sum_{j=1}^{b} X_{ijk}/bc$. Show that

$$\sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} (X_{ijk} - \bar{X}_{...})^2 = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} (X_{ijk} - \bar{X}_{i..})^2 + bc \sum_{i=1}^{a} (\bar{X}_{i..} - \bar{X}_{...})^2.$$

$R\sqrt{n-2}/\sqrt{1-R^2}$; because $g(t)$ does not depend upon $x_1, x_2, \ldots, x_n$ it is obvious that this marginal p.d.f. is $g(t)$, the conditional p.d.f. of $R_c\sqrt{n-2}/\sqrt{1-R_c^2}$. The change-of-variable technique can now be used to find the p.d.f. of R.

Remarks. Since R has, when $\rho = 0$, a conditional distribution that does not depend upon $x_1, x_2, \ldots, x_n$ (and hence that conditional distribution is, in fact, the marginal distribution of R), we have the remarkable fact that R is independent of $X_1, X_2, \ldots, X_n$. It follows that R is independent of *every function* of $X_1, X_2, \ldots, X_n$ alone, that is, a function that does not depend upon any Y_i. In like manner, R is independent of every function of $Y_1, Y_2, \ldots, Y_n$ alone. Moreover, a careful review of the argument reveals that nowhere did we use the fact that X has a normal marginal distribution. Thus, if X and Y are independent, and if Y has a normal distribution, then R has the same conditional distribution whatever be the distribution of X, subject to the condition $\sum_1^n (x_i - \bar{x})^2 > 0$. Moreover, if $\Pr\left[\sum_1^n (X_i - \bar{X})^2 > 0\right] = 1$, then R has the same marginal distribution whatever be the distribution of X.

If we write $T = R\sqrt{n-2}/\sqrt{1-R^2}$, where T has a t-distribution with $n - 2 > 0$ degrees of freedom, it is easy to show, by the change-of-variable technique (Exercise 10.34), that the p.d.f. of R is given by

$$g(r) = \frac{\Gamma[(n-1)/2]}{\Gamma(\frac{1}{2})\Gamma[(n-2)/2]}(1-r^2)^{(n-4)/2}, \qquad -1 < r < 1, \tag{2}$$
$$= 0 \qquad \text{elsewhere.}$$

We have now solved the problem of the distribution of R, when $\rho = 0$ and $n > 2$, or, perhaps more conveniently, that of $R\sqrt{n-2}/\sqrt{1-R^2}$. The likelihood ratio test of the hypothesis $H_0 : \rho = 0$ against all alternatives $H_1 : \rho \neq 0$ may be based either on the statistic R or on the statistic $R\sqrt{n-2}/\sqrt{1-R^2} = T$, although the latter is easier to use. In either case the significance level of the test is

$$\alpha = \Pr(|R| \geq c_1; H_0) = \Pr(|T| \geq c_2; H_0),$$

where the constants c_1 and c_2 are chosen so as to give the desired value of α.

Remark. It is also possible to obtain an approximate test of size α by using the fact that

$$W = \frac{1}{2}\ln\left(\frac{1+R}{1-R}\right)$$

has an approximate normal distribution with mean $\frac{1}{2}\ln[(1+\rho)/(1-\rho)]$ and variance $1/(n-3)$. We accept this statement without proof. Thus a test of $H_0 : \rho = 0$ can be based on the statistic

$$Z = \frac{\frac{1}{2}\ln[(1+R)/(1-R)] - \frac{1}{2}\ln[(1+\rho)/(1-\rho)]}{\sqrt{1/(n-3)}},$$

with $\rho = 0$ so that $\frac{1}{2}\ln[(1+\rho)/(1-\rho)] = 0$. However, using W, we can also test hypotheses like $H_0 : \rho = \rho_0$ against $H_1 : \rho \neq \rho_0$, where ρ_0 is not necessarily zero. In that case the hypothesized mean of W is

$$\frac{1}{2}\ln\left(\frac{1+\rho_0}{1-\rho_0}\right).$$

EXERCISES

10.31. Show that

$$R = \frac{\sum_1^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_1^n (X_i - \bar{X})^2 \sum_1^n (Y_i - \bar{Y})^2}} = \frac{\sum_1^n X_i Y_i - n\bar{X}\bar{Y}}{\sqrt{\left(\sum_1^n X_i^2 - n\bar{X}^2\right)\left(\sum_1^n Y_i^2 - n\bar{Y}^2\right)}}.$$

10.32. A random sample of size $n = 6$ from a bivariate normal distribution yields a value of the correlation coefficient of 0.89. Would we accept or reject, at the 5 percent signficance level, the hypothesis that $\rho = 0$?

10.33. Verify Equation (1) of this section.

10.34. Verify the p.d.f. (2) of this section.

10.8 The Distributions of Certain Quadratic Forms

Remark. It is essential that the reader have the background of the multivariate normal distribution as given in Section 4.10 to understand Sections 10.8 and 10.9.

Let X_i, $i = 1, 2, \ldots, n$, denote independent random variables which are $N(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots, n$, respectively. Then $Q = \sum_1^n (X_i - \mu_i)^2/\sigma_i^2$ is $\chi^2(n)$. Now Q is a quadratic form in the $X_i - \mu_i$ and Q is seen to be, apart from the coefficient $-\frac{1}{2}$, the random variable which is defined by the exponent on the number e in the joint p.d.f. of $X_1, X_2, \ldots, X_n$. We shall now show that this result can be generalized.

Let $X_1, X_2, \ldots, X_n$ have a multivariate normal distribution with p.d.f.

$$\frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{V}|}} \exp\left[-\frac{(\mathbf{x}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2}\right],$$

where, as usual, the covariance matrix $\mathbf{V}$ is positive definite. We shall show that the random variable Q (a quadratic form in the $X_i - \mu_i$), which is defined by $(\mathbf{x}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})$, is $\chi^2(n)$. We have for the m.g.f. $M(t)$ of Q the integral

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{V}|}}$$

$$\times \exp\left[t(\mathbf{x}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu}) - \frac{(\mathbf{x}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{2}\right] dx_1 \cdots dx_n$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{V}|}}$$

$$\times \exp\left[-\frac{(\mathbf{x}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})(1-2t)}{2}\right] dx_1 \cdots dx_n.$$

With $\mathbf{V}^{-1}$ positive definite, the integral is seen to exist for all real values of $t < \frac{1}{2}$. Moreover, $(1-2t)\mathbf{V}^{-1}$, $t < \frac{1}{2}$, is a positive definite matrix and, since $|(1-2t)\mathbf{V}^{-1}| = (1-2t)^n|\mathbf{V}^{-1}|$, it follows that

$$\frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{V}|/(1-2t)^n}} \exp\left[-\frac{(\mathbf{x}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})(1-2t)}{2}\right]$$

can be treated as a multivariate normal p.d.f. If we multiply our integrand by $(1-2t)^{n/2}$, we have this multivariate p.d.f. Thus the m.g.f. of Q is given by

$$M(t) = \frac{1}{(1-2t)^{n/2}}, \qquad t < \tfrac{1}{2},$$

and Q is $\chi^2(n)$, as we wished to show. This fact is the basis of the chi-square tests that were discussed in Chapter 6.

The remarkable fact that the random variable which is defined by $(\mathbf{x}-\boldsymbol{\mu})'\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})$ is $\chi^2(n)$ stimulates a number of questions about quadratic forms in normally distributed variables. We would like to treat this problem in complete generality, but limitations of space forbid this, and we find it necessary to restrict ourselves to some special cases.

Let $X_1, X_2, \ldots, X_n$ denote a random sample of size n from a distribution which is $N(0, \sigma^2)$, $\sigma^2 > 0$. Let $\mathbf{X}' = [X_1, X_2, \ldots, X_n]$ and let $\mathbf{A}$ denote an arbitrary $n \times n$ real symmetric matrix. We shall investigate the distribution of the quadratic form $\mathbf{X}'\mathbf{A}\mathbf{X}$. For instance, we know that $\mathbf{X}'\mathbf{I}_n\mathbf{X}/\sigma^2 = \mathbf{X}'\mathbf{X}/\sigma^2 = \sum_1^n X_i^2/\sigma^2$ is $\chi^2(n)$. First we shall find the m.g.f. of $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$. Then we shall investigate the conditions that must be imposed upon the real symmetric matrix $\mathbf{A}$ if $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ is to have a chi-square distribution. This m.g.f. is given by

$$M(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(\frac{t\mathbf{x}'\mathbf{A}\mathbf{x}}{\sigma^2} - \frac{\mathbf{x}'\mathbf{x}}{2\sigma^2}\right) dx_1 \cdots dx_n$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left[-\frac{\mathbf{x}'(\mathbf{I}-2t\mathbf{A})\mathbf{x}}{2\sigma^2}\right] dx_1 \cdots dx_n,$$

where $\mathbf{I} = \mathbf{I}_n$. The matrix $\mathbf{I} - 2t\mathbf{A}$ is positive definite if we take $|t|$ sufficiently small, say $|t| < h$, $h > 0$. Moreover, we can treat

$$\frac{1}{(2\pi)^{n/2}\sqrt{|(\mathbf{I}-2t\mathbf{A})^{-1}\sigma^2|}} \exp\left[-\frac{\mathbf{x}'(\mathbf{I}-2t\mathbf{A})\mathbf{x}}{2\sigma^2}\right]$$

as a multivariate normal p.d.f. Now $|(\mathbf{I}-2t\mathbf{A})^{-1}\sigma^2|^{1/2} = \sigma^n/|\mathbf{I}-2t\mathbf{A}|^{1/2}$. If we multiply our integrand by $|\mathbf{I}-2t\mathbf{A}|^{1/2}$, we have this multivariate p.d.f. Hence the m.g.f. of $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ is given by

$$M(t) = |\mathbf{I}-2t\mathbf{A}|^{-1/2}, \qquad |t| < h. \tag{1}$$

It proves useful to express this m.g.f. in a different form. To do this, let $a_1, a_2, \ldots, a_n$ denote the characteristic numbers of $\mathbf{A}$ and let $\mathbf{L}$ denote an $n \times n$ orthogonal matrix such that $\mathbf{L}'\mathbf{A}\mathbf{L} = \text{diag}[a_1, a_2, \ldots, a_n]$. Thus

$$\mathbf{L}'(\mathbf{I}-2t\mathbf{A})\mathbf{L} = \begin{bmatrix} 1-2ta_1 & 0 & \cdots & 0 \\ 0 & 1-2ta_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1-2ta_n \end{bmatrix}.$$

Then

$$\prod_{i=1}^{n} (1-2ta_i) = |\mathbf{L}'(\mathbf{I}-2t\mathbf{A})\mathbf{L}| = |\mathbf{I}-2t\mathbf{A}|.$$

Accordingly, we can write $M(t)$, as given in Equation (1), in the form

$$M(t) = \left[\prod_{i=1}^{n} (1 - 2ta_i)\right]^{-1/2}, \qquad |t| < h. \tag{2}$$

Let r, $0 < r \le n$, denote the rank of the real symmetric matrix $\mathbf{A}$. Then exactly r of the real numbers $a_1, a_2, \ldots, a_n$, say $a_1, \ldots, a_r$, are not zero and exactly $n - r$ of these numbers, say $a_{r+1}, \ldots, a_n$, are zero. Thus we can write the m.g.f. of $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ as

$$M(t) = [(1 - 2ta_1)(1 - 2ta_2) \cdots (1 - 2ta_r)]^{-1/2}.$$

Now that we have found, in suitable form, the m.g.f. of our random variable, let us turn to the question of the conditions that must be imposed if $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ is to have a chi-square distribution. Assume that $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ is $\chi^2(k)$. Then

$$M(t) = [(1 - 2ta_1)(1 - 2ta_2) \cdots (1 - 2ta_r)]^{-1/2} = (1 - 2t)^{-k/2},$$

or, equivalently,

$$(1 - 2ta_1)(1 - 2ta_2) \cdots (1 - 2ta_r) = (1 - 2t)^k, \qquad |t| < h.$$

Because the positive integers r and k are the degrees of these polynomials, and because these polynomials are equal for infinitely many values of t, we have $k = r$, the rank of $\mathbf{A}$. Moreover, the uniqueness of the factorization of a polynomial implies that $a_1 = a_2 = \cdots = a_r = 1$. If each of the nonzero characteristic numbers of a real symmetric matrix is one, the matrix is idempotent, that is, $\mathbf{A}^2 = \mathbf{A}$, and conversely (see Exercise 10.38). Accordingly, if $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ has a chi-square distribution, then $\mathbf{A}^2 = \mathbf{A}$ and the random variable is $\chi^2(r)$, where r is the rank of $\mathbf{A}$. Conversely, if $\mathbf{A}$ is of rank r, $0 < r \le n$, and if $\mathbf{A}^2 = \mathbf{A}$, then $\mathbf{A}$ has exactly r characteristic numbers that are equal to one, and the remaining $n - r$ characteristic numbers are equal to zero. Thus the m.g.f. of $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ is given by $(1 - 2t)^{-r/2}$, $t < \frac{1}{2}$, and $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ is $\chi^2(r)$. This establishes the following theorem.

Theorem 2. *Let Q denote a random variable which is a quadratic form in the observations of a random sample of size n from a distribution which is $N(0, \sigma^2)$. Let $\mathbf{A}$ denote the symmetric matrix of Q and let r, $0 < r \le n$, denote the rank of $\mathbf{A}$. Then Q/σ^2 is $\chi^2(r)$ if and only if $\mathbf{A}^2 = \mathbf{A}$.*

Remark. If the normal distribution in Theorem 2 is $N(\mu, \sigma^2)$, the condition $\mathbf{A}^2 = \mathbf{A}$ remains a necessary and sufficient condition that Q/σ^2 have a chi-square distribution. In general, however, Q/σ^2 is not $\chi^2(r)$ but, instead, Q/σ^2 has a noncentral chi-square distribution if $\mathbf{A}^2 = \mathbf{A}$. The number of degrees of freedom is r, the rank of $\mathbf{A}$, and the noncentrality parameter is $\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}/\sigma^2$, where $\boldsymbol{\mu}' = [\mu, \mu, \ldots, \mu]$. Since $\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = \mu^2 \sum_{i,j} a_{ij}$, where $\mathbf{A} = [a_{ij}]$, then, if $\mu \ne 0$, the conditions $\mathbf{A}^2 = \mathbf{A}$ and $\sum_{i,j} a_{ij} = 0$ are necessary and sufficient conditions that Q/σ^2 be central $\chi^2(r)$. Moreover, the theorem may be extended to a quadratic form in random variables which have a multivariate normal distribution with positive definite covariance matrix $\mathbf{V}$; here the necessary and sufficient condition that Q have a chi-square distribution is $\mathbf{AVA} = \mathbf{A}$.

EXERCISES

10.35. Let $Q = X_1X_2 - X_3X_4$, where X_1, X_2, X_3, X_4 is a random sample of size 4 from a distribution which is $N(0, \sigma^2)$. Show that Q/σ^2 does not have a chi-square distribution. Find the m.g.f. of Q/σ^2.

10.36. Let $\mathbf{X}' = [X_1, X_2]$ be bivariate normal with matrix of means $\boldsymbol{\mu}' = [\mu_1, \mu_2]$ and positive definite covariance matrix $\mathbf{V}$. Let

$$Q_1 = \frac{X_1^2}{\sigma_1^2(1 - \rho^2)} - 2\rho \frac{X_1X_2}{\sigma_1\sigma_2(1 - \rho^2)} + \frac{X_2^2}{\sigma_2^2(1 - \rho^2)}.$$

Show that Q_1 is $\chi^2(r, \theta)$ and find r and θ. When and only when does Q_1 have a central chi-square distribution?

10.37. Let $\mathbf{X}' = [X_1, X_2, X_3]$ denote a random sample of size 3 from a distribution that is $N(4, 8)$ and let

$$\mathbf{A} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

Justify the assertion that $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ is $\chi^2(2, 6)$.

10.38. Let $\mathbf{A}$ be a real symmetric matrix. Prove that each of the nonzero characteristic numbers of $\mathbf{A}$ is equal to 1 if and only if $\mathbf{A}^2 = \mathbf{A}$.

Hint: Let $\mathbf{L}$ be an orthogonal matrix such that $\mathbf{L}'\mathbf{A}\mathbf{L} = \text{diag}\,[a_1, a_2, \ldots, a_n]$ and note that $\mathbf{A}$ is idempotent if and only if $\mathbf{L}'\mathbf{A}\mathbf{L}$ is idempotent.

10.39. The sum of the elements on the principal diagonal of a square matrix $\mathbf{A}$ is called the trace of $\mathbf{A}$ and is denoted by tr $\mathbf{A}$.

(a) If $\mathbf{B}$ is $n \times m$ and $\mathbf{C}$ is $m \times n$, prove that tr $(\mathbf{BC})$ = tr $(\mathbf{CB})$.

(b) If $\mathbf{A}$ is a square matrix and if $\mathbf{L}$ is an orthogonal matrix, use the result of part (a) to show that tr $(\mathbf{L}'\mathbf{A}\mathbf{L})$ = tr $\mathbf{A}$.

(c) If $\mathbf{A}$ is a real symmetric idempotent matrix, use the result of part (b) to prove that the rank of $\mathbf{A}$ is equal to tr $\mathbf{A}$.

10.40. Let $\mathbf{A} = [a_{ij}]$ be a real symmetric matrix. Prove that $\sum_j \sum_i a_{ij}^2$ is equal to the sum of the squares of the characteristic numbers of $\mathbf{A}$.

Hint: If $\mathbf{L}$ is an orthogonal matrix, show that $\sum_j \sum_i a_{ij}^2 = \text{tr}(\mathbf{A}^2) = \text{tr}(\mathbf{L}'\mathbf{A}^2\mathbf{L}) = \text{tr}[(\mathbf{L}'\mathbf{A}\mathbf{L})(\mathbf{L}'\mathbf{A}\mathbf{L})]$.

10.41. Let $\bar{X}$ and S^2 denote, respectively, the mean and the variance of a random sample of size n from a distribution which is $N(0, \sigma^2)$.

(a) If $\mathbf{A}$ denotes the symmetric matrix of $n\bar{X}^2$, show that $\mathbf{A} = (1/n)\mathbf{P}$, where $\mathbf{P}$ is the $n \times n$ matrix, each of whose elements is equal to one.

(b) Demonstrate that $\mathbf{A}$ is idempotent and that the tr $\mathbf{A} = 1$. Thus $n\bar{X}^2/\sigma^2$ is $\chi^2(1)$.

(c) Show that the symmetric matrix $\mathbf{B}$ of nS^2 is $\mathbf{I} - (1/n)\mathbf{P}$.

(d) Demonstrate that $\mathbf{B}$ is idempotent and that tr $\mathbf{B} = n - 1$. Thus nS^2/σ^2 is $\chi^2(n-1)$, as previously proved otherwise.

(e) Show that the product matrix $\mathbf{AB}$ is the zero matrix.

10.9 The Independence of Certain Quadratic Forms

We have previously investigated the independence of linear functions of normally distributed variables (see Exercise 4.132). In this section we shall prove some theorems about the independence of quadratic forms. As we remarked on p. 483, we shall confine our attention to normally distributed variables that constitute a random sample of size n from a distribution that is $N(0, \sigma^2)$.

Let $X_1, X_2, \ldots, X_n$ denote a random sample of size n from a distribution which is $N(0, \sigma^2)$. Let $\mathbf{A}$ and $\mathbf{B}$ denote two real symmetric matrices, each of order n. Let $\mathbf{X}' = [X_1, X_2, \ldots, X_n]$ and consider the two quadratic forms $\mathbf{X}'\mathbf{A}\mathbf{X}$ and $\mathbf{X}'\mathbf{B}\mathbf{X}$. We wish to show that these quadratic forms are independent if and only if $\mathbf{AB} = \mathbf{0}$, the zero matrix. We shall first compute the m.g.f. $M(t_1, t_2)$ of $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ and $\mathbf{X}'\mathbf{B}\mathbf{X}/\sigma^2$. We have

$$M(t_1, t_2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(\frac{t_1 \mathbf{x}'\mathbf{A}\mathbf{x}}{\sigma^2} + \frac{t_2 \mathbf{x}'\mathbf{B}\mathbf{x}}{\sigma^2} - \frac{\mathbf{x}'\mathbf{x}}{2\sigma^2}\right) dx_1 \cdots dx_n$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(-\frac{\mathbf{x}'(\mathbf{I} - 2t_1\mathbf{A} - 2t_2\mathbf{B})\mathbf{x}}{2\sigma^2}\right) dx_1 \cdots dx_n.$$

The matrix $\mathbf{I} - 2t_1\mathbf{A} - 2t_2\mathbf{B}$ is positive definite if we take $|t_1|$ and $|t_2|$ sufficiently small, say $|t_1| < h_1$, $|t_2| < h_2$, where $h_1, h_2 > 0$. Then, as on p. 483, we have

$$M(t_1, t_2) = |\mathbf{I} - 2t_1\mathbf{A} - 2t_2\mathbf{B}|^{-1/2}, \qquad |t_1| < h_1, \quad |t_2| < h_2.$$

Let us assume that $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ and $\mathbf{X}'\mathbf{B}\mathbf{X}/\sigma^2$ are independent (so that likewise are $\mathbf{X}'\mathbf{A}\mathbf{X}$ and $\mathbf{X}'\mathbf{B}\mathbf{X}$) and prove that $\mathbf{AB} = \mathbf{0}$. Thus we assume that

$$M(t_1, t_2) = M(t_1, 0)M(0, t_2) \tag{1}$$

for all t_1 and t_2 for which $|t_i| < h_i$, $i = 1, 2$. Identity (1) is equivalent to the identity

$$|\mathbf{I} - 2t_1\mathbf{A} - 2t_2\mathbf{B}| = |\mathbf{I} - 2t_1\mathbf{A}||\mathbf{I} - 2t_2\mathbf{B}|, \qquad |t_i| < h_i, \quad i = 1, 2. \tag{2}$$

Let $r > 0$ denote the rank of $\mathbf{A}$ and let $a_1, a_2, \ldots, a_r$ denote the r nonzero characteristic numbers of $\mathbf{A}$. There exists an orthogonal matrix $\mathbf{L}$ such that

$$\mathbf{L}'\mathbf{A}\mathbf{L} = \left[\begin{array}{cccc|c} a_1 & 0 & \cdots & 0 & \\ 0 & a_2 & \cdots & 0 & \mathbf{0} \\ \vdots & \vdots & & \vdots & \\ 0 & 0 & \cdots & a_r & \\ \hline & & \mathbf{0} & & \mathbf{0} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{C}_{11} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] = \mathbf{C}$$

for a suitable ordering of $a_1, a_2, \ldots, a_r$. Then $\mathbf{L}'\mathbf{B}\mathbf{L}$ may be written in the identically partitioned form

$$\mathbf{L}'\mathbf{B}\mathbf{L} = \left[\begin{array}{c|c} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \hline \mathbf{D}_{21} & \mathbf{D}_{22} \end{array}\right] = \mathbf{D}.$$

The identity (2) may be written as

$$|\mathbf{L}'||\mathbf{I} - 2t_1\mathbf{A} - 2t_2\mathbf{B}||\mathbf{L}| = |\mathbf{L}'||\mathbf{I} - 2t_1\mathbf{A}||\mathbf{L}||\mathbf{L}'||\mathbf{I} - 2t_2\mathbf{B}||\mathbf{L}|, \tag{2'}$$

or as

$$|\mathbf{I} - 2t_1\mathbf{C} - 2t_2\mathbf{D}| = |\mathbf{I} - 2t_1\mathbf{C}||\mathbf{I} - 2t_2\mathbf{D}|. \tag{3}$$

The coefficient of $(-2t_1)^r$ in the right-hand member of Equation (3) is seen by inspection to be $a_1a_2 \cdots a_r|\mathbf{I} - 2t_2\mathbf{D}|$. It is not so easy to find the coefficient of $(-2t_1)^r$ in the left-hand member of Equation (3).

Conceive of expanding this determinant in terms of minors of order r formed from the first r columns. One term in this expansion is the product of the minor of order r in the upper left-hand corner, namely, $|\mathbf{I}_r - 2t_1\mathbf{C}_{11} - 2t_2\mathbf{D}_{11}|$, and the minor of order $n - r$ in the lower right-hand corner, namely, $|\mathbf{I}_{n-r} - 2t_2\mathbf{D}_{22}|$. Moreover, this product is the only term in the expansion of the determinant that involves $(-2t_1)^r$. Thus the coefficient of $(-2t_1)^r$ in the left-hand member of Equation (3) is $a_1 a_2 \cdots a_r|\mathbf{I}_{n-r} - 2t_2\mathbf{D}_{22}|$. If we equate these coefficients of $(-2t_1)^r$, we have, for all t_2, $|t_2| < h_2$,

$$|\mathbf{I} - 2t_2\mathbf{D}| = |\mathbf{I}_{n-r} - 2t_2\mathbf{D}_{22}|. \tag{4}$$

Equation (4) implies that the nonzero characteristic numbers of the matrices $\mathbf{D}$ and $\mathbf{D}_{22}$ are the same (see Exercise 10.49). Recall that the sum of the squares of the characteristic numbers of a symmetric matrix is equal to the sum of the squares of the elements of that matrix (see Exercise 10.40). Thus the sum of the squares of the elements of matrix $\mathbf{D}$ is equal to the sum of the squares of the elements of $\mathbf{D}_{22}$. Since the elements of the matrix $\mathbf{D}$ are real, it follows that each of the elements of $\mathbf{D}_{11}$, $\mathbf{D}_{12}$, and $\mathbf{D}_{21}$ is zero. Accordingly, we can write $\mathbf{D}$ in the form

$$\mathbf{D} = \mathbf{L}'\mathbf{B}\mathbf{L} = \left[\begin{array}{c|c} \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{D}_{22} \end{array}\right].$$

Thus $\mathbf{CD} = \mathbf{L}'\mathbf{A}\mathbf{L}\mathbf{L}'\mathbf{B}\mathbf{L} = \mathbf{0}$ and $\mathbf{L}'\mathbf{A}\mathbf{B}\mathbf{L} = \mathbf{0}$ and $\mathbf{AB} = \mathbf{0}$, as we wished to prove.

To complete the proof of the theorem, we assume that $\mathbf{AB} = \mathbf{0}$. We are to show that $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ and $\mathbf{X}'\mathbf{B}\mathbf{X}/\sigma^2$ are independent. We have, for all real values of t_1 and t_2,

$$(\mathbf{I} - 2t_1\mathbf{A})(\mathbf{I} - 2t_2\mathbf{B}) = \mathbf{I} - 2t_1\mathbf{A} - 2t_2\mathbf{B},$$

since $\mathbf{AB} = \mathbf{0}$. Thus

$$|\mathbf{I} - 2t_1\mathbf{A} - 2t_2\mathbf{B}| = |\mathbf{I} - 2t_1\mathbf{A}||\mathbf{I} - 2t_2\mathbf{B}|.$$

Since the m.g.f. of $\mathbf{X}'\mathbf{A}\mathbf{X}/\sigma^2$ and $\mathbf{X}'\mathbf{B}\mathbf{X}/\sigma^2$ is given by

$$M(t_1, t_2) = |\mathbf{I} - 2t_1\mathbf{A} - 2t_2\mathbf{B}|^{-1/2}, \qquad |t_i| < h_i, \quad i = 1, 2,$$

we have

$$M(t_1, t_2) = M(t_1, 0)M(0, t_2),$$

and the proof of the following theorem is complete.

Theorem 3. *Let Q_1 and Q_2 denote random variables which are quadratic forms in the observations of a random sample of size n from a distribution which is $N(0, \sigma^2)$. Let $\mathbf{A}$ and $\mathbf{B}$ denote, respectively, the real symmetric matrices of Q_1 and Q_2. The random variables Q_1 and Q_2 are independent if and only if $\mathbf{AB} = \mathbf{0}$.*

Remark. Theorem 3 remains valid if the random sample is from a distribution which is $N(\mu, \sigma^2)$, whatever be the real value of μ. Moreover, Theorem 2 may be extended to quadratic forms in random variables that have a joint multivariate normal distribution with a positive definite covariance matrix $\mathbf{V}$. The necessary and sufficient condition for the independence of two such quadratic forms with symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ then becomes $\mathbf{AVB} = \mathbf{0}$. In our Theorem 2, we have $\mathbf{V} = \sigma^2\mathbf{I}$, so that $\mathbf{AVB} = \mathbf{A}\sigma^2\mathbf{I}\mathbf{B} = \sigma^2\mathbf{AB} = \mathbf{0}$.

We shall next prove Theorem 1 that was stated in Section 10.1.

Theorem 4. *Let $Q = Q_1 + \cdots + Q_{k-1} + Q_k$, where Q, $Q_1, \ldots, Q_{k-1}, Q_k$ are $k + 1$ random variables that are quadratic forms in the observations of a random sample of size n from a distribution which is $N(0, \sigma^2)$. Let Q/σ^2 be $\chi^2(r)$, let Q_i/σ^2 be $\chi^2(r_i)$, $i = 1, 2, \ldots, k - 1$, and let Q_k be nonnegative. Then the random variables $Q_1, Q_2, \ldots, Q_k$ are independent and, hence, Q_k/σ^2 is $\chi^2(r_k = r - r_1 - \cdots - r_{k-1})$.*

Proof. Take first the case of $k = 2$ and let the real symmetric matrices of Q, Q_1, and Q_2 be denoted, respectively, by $\mathbf{A}$, $\mathbf{A}_1$, $\mathbf{A}_2$. We are given that $Q = Q_1 + Q_2$ or, equivalently, that $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2$. We are also given that Q/σ^2 is $\chi^2(r)$ and that Q_1/σ^2 is $\chi^2(r_1)$. In accordance with Theorem 2, p. 484, we have $\mathbf{A}^2 = \mathbf{A}$ and $\mathbf{A}_1^2 = \mathbf{A}_1$. Since $Q_2 \geq 0$, each of the matrices $\mathbf{A}$, $\mathbf{A}_1$, and $\mathbf{A}_2$ is positive semidefinite. Because $\mathbf{A}^2 = \mathbf{A}$, we can find an orthogonal matrix $\mathbf{L}$ such that

$$\mathbf{L}'\mathbf{A}\mathbf{L} = \left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right].$$

If then we multiply both members of $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2$ on the left by $\mathbf{L}'$ and on the right by $\mathbf{L}$, we have

$$\left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] = \mathbf{L}'\mathbf{A}_1\mathbf{L} + \mathbf{L}'\mathbf{A}_2\mathbf{L}.$$

Now each of $\mathbf{A}_1$ and $\mathbf{A}_2$, and hence each of $\mathbf{L}'\mathbf{A}_1\mathbf{L}$ and $\mathbf{L}'\mathbf{A}_2\mathbf{L}$ is positive semidefinite. Recall that, if a real symmetric matrix is positive semidefinite, each element on the principal diagonal is positive or zero. Moreover, if an element on the principal diagonal is zero, then all elements in that row and all elements in that column are zero. Thus $\mathbf{L}'\mathbf{A}\mathbf{L} = \mathbf{L}'\mathbf{A}_1\mathbf{L} + \mathbf{L}'\mathbf{A}_2\mathbf{L}$ can be written as

$$\left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{G}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] + \left[\begin{array}{c|c} \mathbf{H}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right]. \tag{5}$$

Since $\mathbf{A}_1^2 = \mathbf{A}_1$, we have

$$(\mathbf{L}'\mathbf{A}_1\mathbf{L})^2 = \mathbf{L}'\mathbf{A}_1\mathbf{L} = \left[\begin{array}{c|c} \mathbf{G}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right].$$

If we multiply both members of Equation (5) on the left by the matrix $\mathbf{L}'\mathbf{A}_1\mathbf{L}$, we see that

$$\left[\begin{array}{c|c} \mathbf{G}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{G}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] + \left[\begin{array}{c|c} \mathbf{G}_r\mathbf{H}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right],$$

or, equivalently, $\mathbf{L}'\mathbf{A}_1\mathbf{L} = \mathbf{L}'\mathbf{A}_1\mathbf{L} + (\mathbf{L}'\mathbf{A}_1\mathbf{L})(\mathbf{L}'\mathbf{A}_2\mathbf{L})$. Thus $(\mathbf{L}'\mathbf{A}_1\mathbf{L}) \times (\mathbf{L}'\mathbf{A}_2\mathbf{L}) = \mathbf{0}$ and $\mathbf{A}_1\mathbf{A}_2 = \mathbf{0}$. In accordance with Theorem 3, Q_1 and Q_2 are independent. This independence immediately implies that Q_2/σ^2 is $\chi^2(r_2 = r - r_1)$. This completes the proof when $k = 2$. For $k > 2$, the proof may be made by induction. We shall merely indicate how this can be done by using $k = 3$. Take $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2 + \mathbf{A}_3$, where $\mathbf{A}^2 = \mathbf{A}$, $\mathbf{A}_1^2 = \mathbf{A}_1$, $\mathbf{A}_2^2 = \mathbf{A}_2$, and $\mathbf{A}_3$ is positive semidefinite. Write $\mathbf{A} = \mathbf{A}_1 + (\mathbf{A}_2 + \mathbf{A}_3) = \mathbf{A}_1 + \mathbf{B}_1$, say. Now $\mathbf{A}^2 = \mathbf{A}$, $\mathbf{A}_1^2 = \mathbf{A}_1$, and $\mathbf{B}_1$ is positive semidefinite. In accordance with the case of $k = 2$, we have $\mathbf{A}_1\mathbf{B}_1 = \mathbf{0}$, so that $\mathbf{B}_1^2 = \mathbf{B}_1$. With $\mathbf{B}_1 = \mathbf{A}_2 + \mathbf{A}_3$, where $\mathbf{B}_1^2 = \mathbf{B}_1$, $\mathbf{A}_2^2 = \mathbf{A}_2$, it follows from the case of $k = 2$ that $\mathbf{A}_2\mathbf{A}_3 = \mathbf{0}$ and $\mathbf{A}_3^2 = \mathbf{A}_3$. If we regroup by writing $\mathbf{A} = \mathbf{A}_2 + (\mathbf{A}_1 + \mathbf{A}_3)$, we obtain $\mathbf{A}_1\mathbf{A}_3 = \mathbf{0}$, and so on.

Remark. In our statement of Theorem 4 we took $X_1, X_2, \ldots, X_n$ to be observations of a random sample from a distribution which is $N(0, \sigma^2)$. We did this because our proof of Theorem 3 was restricted to that case. In fact, if $Q', Q'_1, \ldots, Q'_k$ are quadratic forms in any normal variables (including multivariate normal variables), if $Q' = Q'_1 + \cdots + Q'_k$, if $Q', Q'_1, \ldots, Q'_{k-1}$ are central or noncentral chi-square, and if Q'_k is nonnegative, then $Q'_1, \ldots, Q'_k$ are independent and Q'_k is either central or noncentral chi-square.

This section will conclude with a proof of a frequently quoted theorem due to Cochran.

Theorem 5. *Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution which is $N(0, \sigma^2)$. Let the sum of the squares of these observations be written in the form*

$$\sum_1^n X_i^2 = Q_1 + Q_2 + \cdots + Q_k,$$

where Q_j is a quadratic form in $X_1, X_2, \ldots, X_n$, with matrix $\mathbf{A}_j$ which has rank r_j, $j = 1, 2, \ldots, k$. The random variables $Q_1, Q_2, \ldots, Q_k$ are independent and Q_j/σ^2 is $\chi^2(r_j)$, $j = 1, 2, \ldots, k$, if and only if $\sum_1^k r_j = n$.

Proof. First assume the two conditions $\sum_1^k r_j = n$ and $\sum_1^n X_i^2 = \sum_1^k Q_j$ to be satisfied. The latter equation implies that $\mathbf{I} = \mathbf{A}_1 + \mathbf{A}_2 + \cdots + \mathbf{A}_k$. Let $\mathbf{B}_i = \mathbf{I} - \mathbf{A}_i$. That is, $\mathbf{B}_i$ is the sum of the matrices $\mathbf{A}_1, \ldots, \mathbf{A}_k$ exclusive of $\mathbf{A}_i$. Let R_i denote the rank of $\mathbf{B}_i$. Since the rank of the sum of several matrices is less than or equal to the sum of the ranks, we have $R_i \le \sum_1^k r_j - r_i = n - r_i$. However, $\mathbf{I} = \mathbf{A}_i + \mathbf{B}_i$, so that $n \le r_i + R_i$ and $n - r_i \le R_i$. Hence $R_i = n - r_i$. The characteristic numbers of $\mathbf{B}_i$ are the roots of the equation $|\mathbf{B}_i - \lambda\mathbf{I}| = 0$. Since $\mathbf{B}_i = \mathbf{I} - \mathbf{A}_i$, this equation can be written as $|\mathbf{I} - \mathbf{A}_i - \lambda\mathbf{I}| = 0$. Thus we have $|\mathbf{A}_i - (1 - \lambda)\mathbf{I}| = 0$. But each root of the last equation is one minus a characteristic number of $\mathbf{A}_i$. Since $\mathbf{B}_i$ has exactly $n - R_i = r_i$ characteristic numbers that are zero, then $\mathbf{A}_i$ has exactly r_i characteristic numbers that are equal to 1. However, r_i is the rank of $\mathbf{A}_i$. Thus each of the r_i nonzero characteristic numbers of $\mathbf{A}_i$ is 1. That is, $\mathbf{A}_i^2 = \mathbf{A}_i$ and thus Q_i/σ^2 is $\chi^2(r_i)$, $i = 1, 2, \ldots, k$. In accordance with Theorem 4, the random variables $Q_1, Q_2, \ldots, Q_k$ are independent.

To complete the proof of Theorem 5, take

$$\sum_1^n X_i^2 = Q_1 + Q_2 + \cdots + Q_k,$$

let $Q_1, Q_2, \ldots, Q_k$ be independent, and let Q_j/σ^2 be $\chi^2(r_j)$, $j = 1, 2, \ldots, k$. Then $\sum_1^k Q_j/\sigma^2$ is $\chi^2\left(\sum_1^k r_j\right)$. But $\sum_1^k Q_j/\sigma^2 = \sum_1^n X_i^2/\sigma^2$ is $\chi^2(n)$. Thus $\sum_1^k r_j = n$ and the proof is complete.

EXERCISES

10.42. Let X_1, X_2, X_3 be a random sample from the normal distribution $N(0, \sigma^2)$. Are the quadratic forms $X_1^2 + 3X_1X_2 + X_2^2 + X_1X_3 + X_3^2$ and $X_1^2 - 2X_1X_2 + \frac{2}{3}X_2^2 - 2X_1X_3 - X_3^2$ independent or dependent?