Ch4. Distribution of Quadratic Forms in y

1 definition

Definition 1.1 *If $\boldsymbol{A}$ is a symmetric matrix and $\boldsymbol{y}$ is a vector, the product*

$$\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \sum_i a_{ii}y_i^2 + \sum_{i \neq j} a_{ij}y_iy_j$$

is called a quadratic form. If $\boldsymbol{x}$ is $n \times 1$, $\boldsymbol{y}$ is $p \times 1$, and $\boldsymbol{A}$ is $n \times p$, the product

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{y} = \sum_{ij} a_{ij}x_iy_j$$

is called a bilinear form.

Note: $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{A}'\boldsymbol{y}$, so if let $\boldsymbol{B} = (\boldsymbol{A} + \boldsymbol{A}')/2$, then $B' = B$ (symmetric) and $\boldsymbol{y}'\boldsymbol{B}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$. Thus, to study a quadratic form $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$, we always assume that $\boldsymbol{A}$ is symmetric, $\boldsymbol{A}' = \boldsymbol{A}$.

2 Mean and Variance of Quadratic Forms

Theorem 2.1 *If $\boldsymbol{y}$ is a random vector with mean $\boldsymbol{\mu}$ and covariance matrix Σ and if $\boldsymbol{A}$ is a symmetric matrix of constants, then*

$$E(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) = tr(\boldsymbol{A}\Sigma) + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}. \tag{1}$$

PROOF: Since $\Sigma = E(\boldsymbol{y}-\boldsymbol{\mu})(\boldsymbol{y}-\boldsymbol{\mu})' = E(\boldsymbol{y}\boldsymbol{y}')-\boldsymbol{\mu}\boldsymbol{\mu}'$, i.e., $E(\boldsymbol{y}\boldsymbol{y}') = \Sigma+\boldsymbol{\mu}\boldsymbol{\mu}'$. Since $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is a scalar, it is equal to its trace. We have

$$\begin{aligned}
E(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) &= E[tr(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y})] \\
&= E[tr(\boldsymbol{A}\boldsymbol{y}\boldsymbol{y}')] \\
&= tr[E(\boldsymbol{A}\boldsymbol{y}\boldsymbol{y}')] \\
&= tr[\boldsymbol{A}(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}')] \\
&= tr(\boldsymbol{A}\Sigma) + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}.
\end{aligned}$$

□

Note that since $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is not a linear function of $\boldsymbol{y}$, $E(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) \neq E(\boldsymbol{y}')\boldsymbol{A}E(\boldsymbol{y})$.

Example 2.1 *Show that the sample variance*

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{2}$$

is an unbiased estimator of the population variance.

PROOF: Since the numerator of (2) can be written as

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \boldsymbol{y}'(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})\boldsymbol{y},$$

where $\boldsymbol{J} = \boldsymbol{j}\boldsymbol{j}'$ and $\boldsymbol{j} = (1, 1, \cdots, 1)'$. Thus for use in (1), we have $\boldsymbol{A} = \boldsymbol{I} - \frac{1}{n}\boldsymbol{J}$, $\Sigma = \sigma\boldsymbol{I}$, and $\boldsymbol{\mu} = \mu\boldsymbol{j}$. Hence,

$$\begin{aligned} E[\sum_{i=1}^{n}(y_i - \bar{y})^2] &= tr[(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})(\sigma^2\boldsymbol{I})] + \mu\boldsymbol{j}'(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})\mu\boldsymbol{j} \\ &= \sigma^2 tr(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}) + \mu^2(\boldsymbol{j}'\boldsymbol{j} - \frac{1}{n}\boldsymbol{j}'\boldsymbol{j}\boldsymbol{j}'\boldsymbol{j}) \\ &= \sigma(n - \frac{n}{n}) + \mu^2(n - \frac{1}{n}n^2) \\ &= (n-1)\sigma^2. \end{aligned}$$

Therefore

$$E(s^2) = \sigma^2.$$

□

Theorem 2.2 *If $\boldsymbol{y}$ us $N_p(\boldsymbol{\mu}, \Sigma)$, then*

$$var(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) = 2tr[(\boldsymbol{A}\Sigma)^2] + 4\boldsymbol{\mu}'\boldsymbol{A}\Sigma\boldsymbol{A}\boldsymbol{\mu}. \tag{3}$$

PROOF: See Searle (1971, P. 57). □

Theorem 2.3 *If $\boldsymbol{y}$ is $N_p(\boldsymbol{\mu}, \Sigma)$, then*

$$cov(\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) = 2\Sigma\boldsymbol{A}\boldsymbol{\mu}.$$

PROOF: See Rencher and Schaalje (2008, pp.110). □

Corollary 2.1 *Let $\boldsymbol{B}$ be a $k \times p$ matrix of constants, then*

$$cov(\boldsymbol{B}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) = 2\boldsymbol{B}\Sigma\boldsymbol{A}\boldsymbol{\mu}.$$

Theorem 2.4 *Let* $\boldsymbol{v} = \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{pmatrix}$ *be a partitioned random vector with mean vector and covariance matrix as follows,*

$$E\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \end{pmatrix}, \quad cov\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{pmatrix} = \begin{pmatrix} \Sigma_{yx} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{pmatrix},$$

where $\boldsymbol{y}$ *is* $p \times 1$, $\boldsymbol{x}$ *is* $q \times 1$, *and* Σ_{yx} *is* $p \times q$. *Let* $\boldsymbol{A}$ *be a* $q \times p$ *matrix of constants. Then*

$$E(\boldsymbol{x}'\boldsymbol{A}\boldsymbol{y}) = tr(\boldsymbol{A}\Sigma_{yx}) + \boldsymbol{\mu}_x'\boldsymbol{A}\boldsymbol{\mu}_y.$$

PROOF: Exercise of students. □

Example 2.2 *Show that the population covariance* $\sigma_{xy} = E[(x - \mu_x)(y - \mu_y)]$ *can be estimated by the sample covariance*

$$S_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}),$$

where $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ *is a bivariate random sample from a population with means* μ_x *and* μ_y, *and covariance* σ_{xy}.

3 Noncentral Chi-square Distribution

Definition 3.1 *Suppose that* $\boldsymbol{y}$ *is* $N_n(\boldsymbol{\mu}, I_n)$. *Then* $v = \boldsymbol{y}'\boldsymbol{y}$ *is distributed as a noncentral* χ^2 *distribution with* n *degrees of freedom and noncentrality parameter* $\lambda = \mu'\mu$, *i.e.,*

$$v \sim \chi^2(n, \lambda).$$

When $\lambda = 0$, v *is distributed as a central* χ^2 *distribution, i.e.,* $v \sim \chi^2(n)$.

Theorem 3.1 *If* v *is distributed as* $\chi^2(n, \lambda)$, *then*

$$\begin{aligned} E(v) &= n + 2\lambda, \\ var(v) &= 2n + 8\lambda, \\ M_v(t) &= \frac{1}{(1-2t)^{n/2}} e^{-\lambda[1-1/(1-2t)]}. \end{aligned}$$

PROOF: See Graybill (1976, p.126). □

The χ^2 distribution has an additive property:

Theorem 3.2 *If $v_1, v_2, \cdots, v_k$ are independently distributed as $\chi^2(n_i, \lambda_i)$, then $\sum_{i=1}^k v_i$ is distributed as $\chi^2(\sum_{i=1}^k n_i, \sum_{i=1}^k \lambda_i)$.*

4 Noncentral F Distribution

Definition 4.1 *If u is distributed as a noncentral chi-square, $\chi^2(p, \lambda)$, while v remains a central chi-square, $\chi^2(q)$, with u and v independent. Then*

$$z = \frac{u/p}{v/q} \; is \; distributed \; as \; F(p, q, \lambda),$$

the noncentral F-distribution with noncentrality parameter λ, where λ is the same non-centrality parameter as in the distribution of u.

The mean of the distribution is

$$E(z) = \frac{q}{q-2}(1 + \frac{2\lambda}{p}).$$

5 Noncentral t Distribution

Definition 5.1 *If y is $N(\mu, 1)$, u is $\chi^2(p)$, and y and u are independent, then*

$$t = \frac{y}{\sqrt{u/p}} \, is \, distributed \, as \, t(p, \mu),$$

the noncentral t-distribution with p degrees of freedom and noncentrality parameter μ.

Note if y is $N(\mu, \sigma^2)$, the noncentrality parameter becomes μ/σ, since y/σ is distributed as $N(\mu/sigma, 1)$.

6 Distribution of Quadratic Forms

Theorem 6.1 *Let $\boldsymbol{y}$ be distributed as $N_p(\boldsymbol{\mu}, \Sigma)$, let $\boldsymbol{A}$ be a $p \times p$ symmetric matrix of constants of rank r, and let $\lambda = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$. Then $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is $\chi^2(r, \lambda)$ if and only if $\boldsymbol{A}\Sigma$ is idempotent.*

PROOF: (sufficiency). If $\boldsymbol{A}\Sigma$ is idempotent, $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is $\chi^2(r, \lambda)$.

Since $\Sigma > 0$, we have $\Sigma = \boldsymbol{B}^2$ where $\boldsymbol{B} = \Sigma^{1/2}$ is nonsingular and symmetric.

Let $\boldsymbol{R} = \boldsymbol{BAB}$, $\boldsymbol{R}$ is symmetric and of $rank(\boldsymbol{R}) = rank(\boldsymbol{A}) = r$. (Multiplication by a nonsingular matrix does not change the rank.)

Since $\boldsymbol{A}\Sigma$ is idempotent, we have

$$\boldsymbol{R}^2\boldsymbol{B} = \boldsymbol{BABBABB} = \boldsymbol{BA}\Sigma\boldsymbol{A}\Sigma = \boldsymbol{BA}\Sigma = \boldsymbol{BABB} = \boldsymbol{RB},$$

thus, $\boldsymbol{R}^2\boldsymbol{BB}^{-1} = \boldsymbol{RBB}^{-1}$ and $\boldsymbol{R}^2 = \boldsymbol{R}$, $\boldsymbol{R}$ is idempotent. Therefore, there exists an orthogonal matrix $\boldsymbol{Q}$ such that

$$\boldsymbol{R} = \boldsymbol{Q}' \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} \boldsymbol{Q} = \boldsymbol{P}'\boldsymbol{P},$$

where $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \end{pmatrix} \boldsymbol{Q}$.

Let $\boldsymbol{x} = \boldsymbol{PB}^{-1}\boldsymbol{y}$ be a linear transformation of $\boldsymbol{y}$,

$$E(\boldsymbol{x}) = \boldsymbol{PB}^{-1}\boldsymbol{\mu} = \boldsymbol{\theta},$$

$$cov(\boldsymbol{x}) = \boldsymbol{PB}^{-1}cov(\boldsymbol{y})\boldsymbol{B}^{-1}\boldsymbol{P}' = \boldsymbol{PB}^{-1}\boldsymbol{BBB}^{-1}\boldsymbol{P}' = \boldsymbol{PP}' = \boldsymbol{I}_r.$$

Hence, $\boldsymbol{x} \sim N_r(\boldsymbol{\theta}, \boldsymbol{I}_r)$. By definition, $\boldsymbol{x}'\boldsymbol{x} \sim \chi^2(r, \lambda)$ with the noncentrality parameter

$$\lambda = \frac{1}{2}\boldsymbol{\theta}'\boldsymbol{\theta} = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{B}^{-1}\boldsymbol{P}'\boldsymbol{P}\boldsymbol{B}^{-1}\boldsymbol{\mu} = \boldsymbol{\mu}'\boldsymbol{B}^{-1}\boldsymbol{B}\boldsymbol{A}\boldsymbol{B}\boldsymbol{B}^{-1}\boldsymbol{\mu} = \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}.$$

Since

$$\boldsymbol{x}'\boldsymbol{x} = \boldsymbol{y}'\boldsymbol{B}^{-1}\boldsymbol{P}'\boldsymbol{P}\boldsymbol{B}^{-1}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{A}\boldsymbol{y},$$

the proof is completed.

The proof for the necessity can be found in Searle (1971, pp.57-58). □

Corollary 6.1 *If $\boldsymbol{y}$ is $N_p(\boldsymbol{O}, \boldsymbol{I})$, then $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ is $\chi^2(r)$ if and only if $\boldsymbol{A}$ is idempotent of rank r.*

Corollary 6.2 *If $\boldsymbol{y}$ is $N_p(\boldsymbol{\mu}, \sigma^2\boldsymbol{I})$, then $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}/\sigma^2$ is $\chi^2(r, \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/2\sigma^2)$ is and only if $\boldsymbol{A}$ is idempotent of rank r.*

Example 6.1 *Consider the sample variance*

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2,$$

where $y_i \sim N(\mu, \sigma^2)$. In the matrix form, it is can be written as

$$S^2 = \frac{1}{n-1}\boldsymbol{y}'(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})\boldsymbol{y},$$

where $\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}$ is idempotent and of rank $n-1$. We next find λ,

$$\lambda = \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/(2\sigma^2) = \mu^2[\boldsymbol{j}'\boldsymbol{j} - (1/n)\boldsymbol{j}'\boldsymbol{J}\boldsymbol{j}]/(2\sigma^2) = 0.$$

Therefore, $\boldsymbol{y}'(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})\boldsymbol{y}/\sigma^2 = (n-1)S^2/\sigma^2$ is distributed as $\chi^2(n-1)$.

7 Independence of Linear Forms and Quadratic Forms

Lemma 7.1 *A symmetric matrix $\boldsymbol{A}$, of order n and rank r, can be written as $\boldsymbol{L}\boldsymbol{L}'$ where $\boldsymbol{L}$ is $n \times r$ of rank r, i.e., L has full column rank.*

PROOF:

$$\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}' = \begin{pmatrix} \boldsymbol{D}_r^2 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} = \begin{pmatrix} \boldsymbol{D}_r \\ \boldsymbol{O} \end{pmatrix} \begin{pmatrix} \boldsymbol{D}_r & \boldsymbol{O} \end{pmatrix}$$

for some orthogonal $\boldsymbol{P}$, where $\boldsymbol{D}_r^2$ is diagonal of order r. Hence

$$\boldsymbol{A} = \boldsymbol{P}' \begin{pmatrix} \boldsymbol{D}_r \\ \boldsymbol{O} \end{pmatrix} \begin{pmatrix} \boldsymbol{D}_r & \boldsymbol{O} \end{pmatrix} \boldsymbol{P} = \boldsymbol{L}\boldsymbol{L}'$$

where $\boldsymbol{L}' = \begin{pmatrix} \boldsymbol{D}_r & \boldsymbol{O} \end{pmatrix} \boldsymbol{P}$ of order $r \times n$ and full row rank; i.e., $\boldsymbol{L}$ is of full column rank. Note also that although $\boldsymbol{L}\boldsymbol{L}' = \boldsymbol{A}$, $\boldsymbol{L}'\boldsymbol{L} = \boldsymbol{D}_r^2$. Also, $\boldsymbol{L}'$ is real only when $\boldsymbol{A}$ is non-negative definite, for only then are the non-zero elements of $\boldsymbol{D}_r^2$ positive. □

Theorem 7.1 *(linear and quadratic) Suppose $\boldsymbol{B}$ is a $k \times p$ matrix of constants, $\boldsymbol{A}$ is a $p \times p$ symmetric matrix of constants, and $\boldsymbol{y}$ is distributed as $N_p(\boldsymbol{\mu}, \Sigma)$. Then $\boldsymbol{B}\boldsymbol{y}$ and $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ are independent if and only if $\boldsymbol{B}\Sigma\boldsymbol{A} = \boldsymbol{O}$.*

PROOF: (sufficiency) $\boldsymbol{B}\Sigma\boldsymbol{A} = \boldsymbol{O}$ implies independence.

From the lemma, $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}'$, where $\boldsymbol{L}$ is of full-column-rank.

$$\boldsymbol{B}\Sigma\boldsymbol{A} = \boldsymbol{O} \; implies \; \boldsymbol{B}\Sigma\boldsymbol{L}\boldsymbol{L}'\boldsymbol{L}(\boldsymbol{L}'\boldsymbol{L})^{-1} = \boldsymbol{O} \; i.e. \; \boldsymbol{B}\Sigma\boldsymbol{L} = \boldsymbol{O}.$$

Therefore $cov(\boldsymbol{B}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{L}) = \boldsymbol{B}\Sigma\boldsymbol{L} = \boldsymbol{O}$. Hence, because $\boldsymbol{y}$ is normal, $\boldsymbol{B}\boldsymbol{y}$ and $\boldsymbol{y}'\boldsymbol{L}$ are independent. Consequently $\boldsymbol{B}\boldsymbol{y}$ and $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ are independent.

(necessity): the independence of $\boldsymbol{By}$ and $\boldsymbol{y}'\boldsymbol{Ay}$ implies $\boldsymbol{B}\Sigma\boldsymbol{A} = \boldsymbol{O}$.

The independence property gives $cov(\boldsymbol{By}, \boldsymbol{y}'\boldsymbol{Ay}) = 2\boldsymbol{B}\Sigma\boldsymbol{A}\boldsymbol{\mu} = 0$. Hence $2\boldsymbol{B}\Sigma\boldsymbol{A}\boldsymbol{\mu} = 0$, and since this is true for all $\boldsymbol{\mu}$, $\boldsymbol{B}\Sigma\boldsymbol{A} = \boldsymbol{O}$, and so the proof is complete. □

Example 7.1 *Consider* $s^2 = \frac{1}{n-1}\boldsymbol{y}'(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})\boldsymbol{y}$ *and* $\bar{y} = \frac{1}{n}\boldsymbol{j}'\boldsymbol{y}$. *By theorem 7.1,* $\bar{y}$ *is independent of* s^2 *since* $\boldsymbol{j}'(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}) = \boldsymbol{O}$.

Theorem 7.2 *(quadratic and quadratic) Let* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *be symmetric matrices of constants. If* $\boldsymbol{y}$ *is* $N_p(\boldsymbol{\mu}, \Sigma)$*, then* $\boldsymbol{y}'\boldsymbol{Ay}$ *and* $\boldsymbol{y}'bB\boldsymbol{y}$ *are independent if and only if* $\boldsymbol{A}\Sigma\boldsymbol{B} = \boldsymbol{O}$.

PROOF: (sufficiency) $\boldsymbol{A}\Sigma\boldsymbol{B} = \boldsymbol{O}$ implies independence.

By the lemma, we have $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}'$ and $\boldsymbol{B} = \boldsymbol{M}\boldsymbol{M}'$, where each of $\boldsymbol{L}$ and $\boldsymbol{M}$ have full column rank. Therefore, if $\boldsymbol{A}\Sigma\boldsymbol{B} = \boldsymbol{O}$, $\boldsymbol{L}\boldsymbol{L}'\Sigma\boldsymbol{M}\boldsymbol{M}' = \boldsymbol{O}$, and because $(\boldsymbol{L}'\boldsymbol{L})^-1$ and $(\boldsymbol{M}'\boldsymbol{M})^{-1}$ exist this means $\boldsymbol{L}'\Sigma\boldsymbol{M} = \boldsymbol{O}$. Therefore

$$cov(\boldsymbol{L}'\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{M}) = \boldsymbol{L}'\Sigma\boldsymbol{M} = \boldsymbol{O}.$$

Hence, because $\boldsymbol{y}$ is normal, $\boldsymbol{L}'\boldsymbol{y}$ and $\boldsymbol{y}'\boldsymbol{M}$ are independent. Consequently $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{L}\boldsymbol{L}'\boldsymbol{y}$ and $\boldsymbol{B}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{M}\boldsymbol{M}'\boldsymbol{y}$ are independent.

(necessity) the independence implies $\boldsymbol{A}\Sigma\boldsymbol{B} = \boldsymbol{O}$.

When $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ and $\boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}$ are independent, $cov(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}) = 0$, so that

$$var(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} + \boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}) = var(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) + var(\boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}),$$

i.e.,

$$var(\boldsymbol{y}'(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{y}) = var(\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}) + var(\boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}).$$

Applying equation (3) to all three terms in this results leads, after a little simplification, to

$$tr(\Sigma\boldsymbol{A}\Sigma\boldsymbol{B}) + 2\boldsymbol{\mu}'\boldsymbol{A}\Sigma\boldsymbol{B}\boldsymbol{\mu} = 0.$$

This is true for all $\boldsymbol{\mu}$, including $\boldsymbol{\mu} = 0$, so $tr(\Sigma\boldsymbol{A}\Sigma\boldsymbol{B}) = 0$ and on substituting back gives $2\boldsymbol{\mu}'\boldsymbol{A}\Sigma\boldsymbol{B}\boldsymbol{\mu} = 0$. This in turn is true for all $\boldsymbol{\mu}$, and so $\boldsymbol{A}\Sigma\boldsymbol{B} = 0$. Thus theorem is proved. □

Example 7.2 *To illustrate theorem 7.2, consider the partitioning of* $\sum_i y_i^2 = \sum_{i=1}^n (y_i -$

$\bar{y})^2 + n\bar{y}^2$, *i.e.,*

$$\boldsymbol{y}'\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})\boldsymbol{y} + \boldsymbol{y}'(\frac{1}{n}\boldsymbol{J})\boldsymbol{y}.$$

If $\boldsymbol{y}$ *is* $N_n(\mu\boldsymbol{j}, \sigma^2\boldsymbol{I})$, *then by theorem 7.2,* $\boldsymbol{y}'(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})\boldsymbol{y}$ *and* $\boldsymbol{y}'(\frac{1}{n}\boldsymbol{J})\boldsymbol{y}$ *are independent if and only if* $(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})(\frac{1}{n}\boldsymbol{J}) = \boldsymbol{O}$.

Theorem 7.3 *(Several quadratic forms) Let* $\boldsymbol{y}$ *be* $N_n(\boldsymbol{\mu}, \sigma^2 I)$*, let* $\boldsymbol{A}_i$ *be symmetric of rank* r_i *for* $i = 1, 2, \cdots, k$*, and let* $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \sum_{i=1}^{k} \boldsymbol{y}'\boldsymbol{A}_i\boldsymbol{y}$*, where* $\boldsymbol{A} = \sum_{i=1}^{k} \boldsymbol{A}_i$ *is symmetric of rank* r*. Then*

(1) $\boldsymbol{y}'\boldsymbol{A}_i\boldsymbol{y}/\sigma^2$ *is* $\chi^2(r_i, \boldsymbol{\mu}'\boldsymbol{A}_i\boldsymbol{\mu}/2\sigma^2)$, $i = 1, 2, \cdots, k$*, and*

(2) $\boldsymbol{y}'\boldsymbol{A}_i\boldsymbol{y}$ *and* $\boldsymbol{y}'\boldsymbol{A}_j\boldsymbol{y}$ *are independent for all* $i \neq j$*, and*

(3) $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}/\sigma^2$ *is* $\chi^2(r, \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}/2\sigma^2)$

if and only if any two of the following three statements are true:

(a) *each* $\boldsymbol{A}_i$ *is idempotent,*

(b) $\boldsymbol{A}_i\boldsymbol{A}_j = \boldsymbol{O}$ *for all* $i \neq j$*.*

(c) $\boldsymbol{A} = \sum_{i=1}^{k} \boldsymbol{A}_i$ *is idempotent.*

or if and only if (c) and (d) are true, where (d) is the following statement:

(d) $r = \sum_{i=1}^{k} r_i$*.*