5 Distribution of Quadratic Forms in y

5.1 SUMS OF SQUARES

In Chapters 3 and 4, we discussed some properties of linear functions of the random vector $\mathbf{y}$. We now consider quadratic forms in $\mathbf{y}$. We will find it useful in later chapters to express a sum of squares encountered in regression or analysis of variance as a quadratic form $\mathbf{y}'\mathbf{A}\mathbf{y}$, where $\mathbf{y}$ is a random vector and $\mathbf{A}$ is a symmetric matrix of constants [see (2.33)]. In this format, we will be able to show that certain sums of squares have chi-square distributions and are independent, thereby leading to F tests.

Example 5.1. We express some simple sums of squares as quadratic forms in $\mathbf{y}$. Let $y_1, y_2, \ldots, y_n$ be a random sample from a population with mean μ and variance σ^2. In the following identity, the total sum of squares $\sum_{i=1}^n y_i^2$ is partitioned into a sum of squares about the sample mean $\bar{y} = \sum_{i=1}^n y_i/n$ and a sum of squares due to the mean:

$$\begin{aligned}\sum_{i=1}^n y_i^2 &= \left(\sum_{i=1}^n y_i^2 - n y^{-2}\right) + n\bar{y}^2 \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 + n\bar{y}^2. \qquad (5.1)\end{aligned}$$

Using (2.20), we can express $\sum_{i=1}^n y_i^2$ as a quadratic form

$$\sum_{i=1}^n y_i^2 = \mathbf{y}'\mathbf{y} = \mathbf{y}'\mathbf{I}\mathbf{y},$$

where $\mathbf{y}' = (y_1, y_2, \ldots, y_n)$. Using $\mathbf{j} = (1, 1, \ldots, 1)'$ as defined in (2.6), we can

Linear Models in Statistics, Second Edition, by Alvin C. Rencher and G. Bruce Schaalje

write $\bar{y}$ as

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\mathbf{j}'\mathbf{y}$$

$(1, \cdots, 1) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$

[see (2.24)]. Then $n\bar{y}^2$ becomes

$$\begin{aligned} n\bar{y}^2 &= n\left(\frac{1}{n}\mathbf{j}'\mathbf{y}\right)^2 = n\left(\frac{1}{n}\mathbf{j}'\mathbf{y}\right)\left(\frac{1}{n}\mathbf{j}'\mathbf{y}\right) \\ &= n\left(\frac{1}{n}\right)^2 \underset{1\times n}{\mathbf{y}'}\underbrace{\overset{n\times 1}{\mathbf{j}}\overset{1\times n}{\mathbf{j}'}}\underset{n\times 1}{\mathbf{y}} \quad \text{[by (2.18)]} \\ &= n\left(\frac{1}{n}\right)^2 \underset{1\times n}{\mathbf{y}'}\overset{n\times n}{\mathbf{J}}\underset{n\times 1}{\mathbf{y}} \quad \text{[by (2.23)]} \\ &= \mathbf{y}'\left(\frac{1}{n}\mathbf{J}\right)\mathbf{y}. \end{aligned}$$

$b'c = c'b$

$jj' = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} = J$

We can now write $\sum_{i=1}^{n}(y_i - \bar{y})^2$ as

$$\begin{aligned} \sum_{i=1}^{n}(y_i - \bar{y})^2 &= \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = \mathbf{y}'\mathbf{I}\mathbf{y} - \mathbf{y}'\left(\frac{1}{n}\mathbf{J}\right)\mathbf{y} \\ &= \mathbf{y}'\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{y}. \end{aligned} \tag{5.2}$$

Hence (5.1) can be written in terms of quadratic forms as

이런 이유로

$$\mathbf{y}'\mathbf{I}\mathbf{y} = \mathbf{y}'\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{y} + \mathbf{y}'\left(\frac{1}{n}\mathbf{J}\right)\mathbf{y}. \tag{5.3}$$

■

The matrices of the three quadratic forms in (5.3) have the following properties:

1. $\mathbf{I} = \left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right) + \frac{1}{n}\mathbf{J}$.
2. $\mathbf{I}, \mathbf{I} - \frac{1}{n}\mathbf{J}$, and $\frac{1}{n}\mathbf{J}$ are idempotent.

   멱등성. $II' = I$
3. $\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\left(\frac{1}{n}\mathbf{J}\right) = \mathbf{O}$.

Using theorems given later in this chapter (and assuming normality of the y_i's), these three properties lead to the conclusion that $\sum_{i=1}^{n}(y_i - \bar{y})^2/\sigma^2$ and $n\bar{y}^2/\sigma^2$ have chi-square distributions and are independent.

이차형식의 평균과 분산

5.2 MEAN AND VARIANCE OF QUADRATIC FORMS

We first consider the mean of a quadratic form $\mathbf{y}'\mathbf{A}\mathbf{y}$.

Theorem 5.2a. If $\mathbf{y}$ is a random vector with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ and if $\mathbf{A}$ is a symmetric matrix of constants, then

대칭행렬

$$E(\mathbf{y}'\mathbf{A}\mathbf{y}) = \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}. \tag{5.4}$$

평균이 μ, 공분산 행렬이 Σ.

PROOF. By (3.25), $\Sigma = E(\mathbf{y}\mathbf{y})' - \boldsymbol{\mu}\boldsymbol{\mu}'$, which can be written as

$$E(\mathbf{y}\mathbf{y}') = \Sigma + \boldsymbol{\mu}\boldsymbol{\mu}'. \tag{5.5}$$

Since $\mathbf{y}'\mathbf{A}\mathbf{y}$ is a scalar, it is equal to its trace. We thus have

대각합.

$1 \times n \quad n \times n \quad n \times 1$

$$
\begin{aligned}
E(\mathbf{y}'\mathbf{A}\mathbf{y}) &= E[\mathrm{tr}(\mathbf{y}'\mathbf{A}\mathbf{y})] \\
&= E[\mathrm{tr}(\mathbf{A}\mathbf{y}\mathbf{y}')] && \text{[by (2.87)]} \quad tr(AB) = tr(BA) \\
&= \mathrm{tr}[E(\mathbf{A}\mathbf{y}\mathbf{y}')] && \text{[by (3.5)]} \quad E[u(y)+v(y)] = E[u(y)] + E[v(y)] \\
&= \mathrm{tr}[\mathbf{A}E(\mathbf{y}\mathbf{y}')] && \text{[by (3.40)]} \quad E(AXB) = AE(X)B \\
&= \mathrm{tr}[\mathbf{A}(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}')] && \text{[by (5.8)]} \\
&= \mathrm{tr}[\mathbf{A}\boldsymbol{\Sigma} + \mathbf{A}\boldsymbol{\mu}\boldsymbol{\mu}'] && \text{[by (2.15)]} \\
&= \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \mathrm{tr}(\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}) && \text{[by (2.86)]} \\
&= \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}
\end{aligned}
$$

기댓값 연산과 행렬 곱 연산의 교환이 항상 성립하지 않는다.

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

↳ y에 대한 선형 함수가 아니라서. $\mu'A\mu$

이차형식은 선형이 아님.

Example 5.2a. To illustrate Theorem 5.2a, consider the sample variance

표본 분산

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}. \tag{5.6}$$

By (5.2), the numerator of (5.6) can be written as

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

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

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

where $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$. If the y's are assumed to be independently distributed with mean μ and variance σ^2, then $E(\mathbf{y}) = (\mu, \mu, \ldots, \mu)' = \mu\mathbf{j}$ and $\text{cov}(\mathbf{y}) = \sigma^2\mathbf{I}$. Thus for use in (5.4) we have $\mathbf{A} = \mathbf{I} - (1/n)\mathbf{J}$, $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$, and $\boldsymbol{\mu} = \mu\mathbf{j}$; hence

$\to E(y'Ay) = tr(A\Sigma) + \mu' A \mu$

$$
\begin{aligned}
E\left[\sum_{i=1}^{n}(y_i - \bar{y})^2\right] &= \text{tr}\left[\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)(\sigma^2\mathbf{I})\right] + \mu\mathbf{j}'\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mu\mathbf{j} \\
&= \sigma^2\text{tr}\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right) + \mu^2(\mathbf{j}'\mathbf{j} - \mathbf{j}'\mathbf{j}\mathbf{j}'\mathbf{j}) \qquad \text{[by (2.23)]} \\
&= \sigma^2\left(n - \frac{n}{n}\right) + \mu^2\left(n - \frac{1}{n}n^2\right) \qquad \text{[by (2.23)]} \\
&= \sigma^2(n-1) + 0.
\end{aligned}
$$

$j' \frac{1}{n} j j' j$

$j'j = n,$

$jj' = J$

$\to 0$

Therefore

$$
E(s^2) = \frac{E\left[\sum_{i=1}^{n}(y_i - \bar{y})^2\right]}{n-1} = \frac{(n-1)\sigma^2}{n-1} = \sigma^2. \tag{5.7}
$$

∎

정규성

Note that normality of the y's is not assumed in Theorem 5.2a. However, normality is assumed in obtaining the moment generating function of $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\text{var}(\mathbf{y}'\mathbf{A}\mathbf{y})$ in the following theorems.

mgf

↳ 확률변수 X의 적률생성함수(mgf) : $M_X(t) = E(e^{tX})$

Theorem 5.2b. If $\mathbf{y}$ is $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then the moment generating function of $\mathbf{y}'\mathbf{A}\mathbf{y}$ is

$$
M_{\mathbf{y}'\mathbf{A}\mathbf{y}}(t) = |\mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma}|^{-1/2} e^{-\boldsymbol{\mu}'[\mathbf{I} - (\mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma})^{-1}]\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}/2}. \tag{5.8}
$$

$\nearrow E[u(y)] = \int_{-\infty}^{\infty} u(y) f(y)\, dy$

Proof. By the multivariate analog of (3.3), we obtain

y의 확률밀도함수 $P(y) = k_1 e^{-(y-\mu)'\Sigma^{-1}(y-\mu)/2}$, $k_1 = 1/(2\pi)^{P/2}|\Sigma|^{1/2}$

각 y_i에 대해 적분을 수행해야 하므로 다중적분.

$\nearrow -\frac{1}{2}\left((y-\mu)'\Sigma^{-1}(y-\mu) - 2ty'Ay\right)$

$\downarrow$ $y'\Sigma^{-1}y - 2\mu'\Sigma^{-1}y + \mu'\Sigma^{-1}\mu$

↳ $y'(\Sigma^{-1} - 2tA)y - 2\mu'\Sigma^{-1}y + \mu'\Sigma^{-1}\mu$

$$
\begin{aligned}
M_{\mathbf{y}'\mathbf{A}\mathbf{y}}(t) = E(e^{t\mathbf{y}'\mathbf{A}\mathbf{y}}) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{t\mathbf{y}'\mathbf{A}\mathbf{y}} k_1 e^{-(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})/2} d\mathbf{y} \\
&= k_1 \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-[\mathbf{y}'\underbrace{(\mathbf{I}-2t\mathbf{A}\boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}}_{B}\mathbf{y} - 2\underbrace{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}}_{C = \Sigma^{-1}\mu}\mathbf{y} + \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}]/2} d\mathbf{y},
\end{aligned}
$$

$\boxed{\begin{array}{l} \text{* 이차 형식의 완전 제곱 정리.} \\ y'By + 2c'y \\ = (y - B^{-1}c)'B(y - B^{-1}c) - c'B^{-1}c \end{array}}$

$B^{-1} = (\Sigma^{-1}(I - 2tA\Sigma)^{-1})$

[$(AB)^{-1} = B^{-1}A^{-1}$]성질 이용.

$\theta' = B^{-1}c = (\Sigma^{-1}(I-2tA\Sigma)^{-1}) \cdot \Sigma^{-1}\mu = \Sigma^{-1}((I-2tA\Sigma)^{-1}\Sigma^{-1})\mu = \mu'(I-2tA\Sigma)^{-1}$

where $k_1 = 1/[(\sqrt{2\pi})^p|\Sigma|^{1/2}]$ and $d\mathbf{y} = dy_1\, dy_2 \ldots dy_p$. For t sufficiently close to 0, $\mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma}$ is nonsingular. Letting $\boldsymbol{\theta}' = \boldsymbol{\mu}'(\mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma})^{-1}$ and $\mathbf{V}^{-1} = (\mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}$, we obtain

└ 새로운 평균 벡터

└ 새로운 공분산 행렬.

$$
M_{\mathbf{y}'\mathbf{A}\mathbf{y}}(t) = k_1 k_2 \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} k_3 e^{-(\mathbf{y}-\boldsymbol{\theta})'\mathbf{V}^{-1}(\mathbf{y}-\boldsymbol{\theta})/2} d\mathbf{y}
$$

$k_1 = 1/(2\pi)^{p/2}|\Sigma|^{1/2}$

(Problem 5.4), where $k_2 = (\sqrt{(2\pi)})^p|\mathbf{V}|^{1/2}e^{-[\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}-\boldsymbol{\theta}'\mathbf{V}^{-1}\boldsymbol{\theta}]/2}$ and $k_3 = 1/[(\sqrt{(2\pi)})^p|\mathbf{V}|^{1/2}]$. The multiple integral is equal to 1 since the multivariate normal density integrates to 1. Thus $M_{\mathbf{y}'\mathbf{A}\mathbf{y}}(t) = k_1k_2$. Substituting and simplifying, we obtain (5.8) (see Problem 5.5). ∎

Theorem 5.2c. If $\mathbf{y}$ is $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then

$$\text{var}(\mathbf{y}'\mathbf{A}\mathbf{y}) = 2\text{tr}[(\mathbf{A}\boldsymbol{\Sigma})^2] + 4\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\mu}. \quad (5.9)$$

Proof. The variance of a random variable can be obtained by evaluating the second derivative of the natural logarithm of its moment generating function at $t = 0$ (see hint to Problem 5.14). Let $\mathbf{C} = \mathbf{I} - 2t\mathbf{A}\boldsymbol{\Sigma}$. Then, from (5.8)

자연로그

$\frac{d^2 \ln M_{y'Ay}(t=0)}{dt^2}$

$$k(t) = \ln[M_{\mathbf{y}'\mathbf{A}\mathbf{y}}(t)] = -\frac{1}{2}\ln|\mathbf{C}| - \frac{1}{2}\boldsymbol{\mu}'(\mathbf{I} - \mathbf{C}^{-1})\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}.$$

Using (2.117), we differentiate $k(t)$ twice to obtain

$$k''(t) = \frac{1}{2}\frac{1}{|\mathbf{C}|^2}\left[\frac{d|\mathbf{C}|}{dt}\right]^2 - \frac{1}{2}\frac{1}{|\mathbf{C}|}\frac{d^2|\mathbf{C}|}{dt^2} - \frac{1}{2}\boldsymbol{\mu}'\mathbf{C}^{-1}\frac{d^2\mathbf{C}}{dt^2}\mathbf{C}^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$$
$$+ \boldsymbol{\mu}\left[\mathbf{C}^{-1}\frac{d\mathbf{C}}{dt}\right]^2\mathbf{C}^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$$

(Problem 5.6). A useful expression for $|\mathbf{C}|$ can be found using (2.97) and (2.107). Thus, if the eigenvalues of $\mathbf{A}\boldsymbol{\Sigma}$ are λ_i, $i = 1, \ldots, p$, we obtain

$$\begin{aligned}|\mathbf{C}| &= \prod_{i=1}^{p}(1 - 2t\lambda_i) \\ &= 1 - 2t\sum_i \lambda_i + 4t^2\sum_{i\neq j}\lambda_i\lambda_j - \cdots + (-1)^p 2^p t^p \lambda_1\lambda_2\cdots\lambda_p.\end{aligned}$$

Then $(d|\mathbf{C}|/dt) = -2\Sigma_i\lambda_i + 8t\Sigma_{i\neq j}\lambda_i\lambda_j +$ higher-order terms in t, and $(d^2|\mathbf{C}|/dt^2) = 8\Sigma_{i\neq j}\lambda_i\lambda_j +$ higher-order terms in t. Evaluating these expressions at $t = 0$, we obtain $|\mathbf{C}| = 1$, $(d|\mathbf{C}|/dt)|_{t=0} = -2\Sigma_i\lambda_i = -2\,\text{tr}(\mathbf{A}\boldsymbol{\Sigma})$ and $(d^2|\mathbf{C}|/dt^2)|_{t=0} = 8\Sigma_{i\neq j}\lambda_i\lambda_j$. For $t = 0$ it is also true that $\mathbf{C} = \mathbf{I}$, $\mathbf{C}^{-1} = \mathbf{I}$, $(d\mathbf{C}/dt)|_{t=0} = 2\mathbf{A}\boldsymbol{\Sigma}$ and $(d^2\mathbf{C}/dt)|_{t=0} = \mathbf{O}$. Thus

$$\begin{aligned}k''(0) &= 2[\text{tr}(\mathbf{A}\boldsymbol{\Sigma})]^2 - 4\sum_{i\neq j}\lambda_i\lambda_j + 0 + 4\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\mu} \\ &= 2\left\{[\text{tr}(\mathbf{A}\boldsymbol{\Sigma})]^2 - 2\sum_{i\neq j}\lambda_i\lambda_j\right\} + 4\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\mu}.\end{aligned}$$

By Problem 2.81, this can be written as

$$2\,\mathrm{tr}[(\mathbf{A\Sigma})^2] + 4\boldsymbol{\mu}'\mathbf{A\Sigma A}\boldsymbol{\mu}.$$

□

We now consider cov(**y**, **y**′**Ay**). To clarify the meaning of the expression cov(**y**, **y**′**Ay**), we denote **y**′**Ay** by the scalar random variable v. Then cov(**y**, v) is a column vector containing the covariance of each y_i and v:

$$\mathrm{cov}(\mathbf{y}, v) = E\{[\mathbf{y} - E(\mathbf{y})][v - E(v)]\} = \begin{pmatrix} \sigma_{y_1 v} \\ \sigma_{y_2 v} \\ \vdots \\ \sigma_{y_p v} \end{pmatrix}. \tag{5.10}$$

[On the other hand, cov(v, **y**) would be a row vector.] An expression for cov(**y**, **y**′**Ay**) is given in the next theorem.

Theorem 5.2d. If **y** is $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then

$$\mathrm{cov}(\mathbf{y}, \mathbf{y}'\mathbf{Ay}) = 2\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\mu}. \tag{5.11}$$

PROOF. By the definition in (5.10), we have

$$\mathrm{cov}(\mathbf{y}, \mathbf{y}'\mathbf{Ay}) = E\{[\mathbf{y} - E(\mathbf{y})][\mathbf{y}'\mathbf{Ay} - E(\mathbf{y}'\mathbf{Ay})]\}.$$

By Theorem 5.2a, this becomes

$$\mathrm{cov}(\mathbf{y}, \mathbf{y}'\mathbf{Ay}) = E\{(\mathbf{y} - \boldsymbol{\mu})[\mathbf{y}'\mathbf{Ay} - \mathrm{tr}(\mathbf{A\Sigma}) - \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}]\}.$$

Rewriting * $\mathbf{y}'\mathbf{Ay} - \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$ in terms of $\mathbf{y} - \boldsymbol{\mu}$ (see Problem 5.7), we obtain

$= (y-\mu)'A(y-\mu) + 2(y-\mu)'A\mu$

$$\begin{aligned} \mathrm{cov}(\mathbf{y}, \mathbf{y}'\mathbf{Ay}) &= E\{(\mathbf{y} - \boldsymbol{\mu})[(\mathbf{y} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{y} - \boldsymbol{\mu}) + 2(\mathbf{y} - \boldsymbol{\mu})'\mathbf{A}\boldsymbol{\mu} - \mathrm{tr}(\mathbf{A\Sigma})]\} \\ &= E[(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{y} - \boldsymbol{\mu})] + 2E[(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})'\mathbf{A}\boldsymbol{\mu}] \\ &\quad - E[(\mathbf{y} - \boldsymbol{\mu})\mathrm{tr}(\mathbf{A\Sigma})] \\ &= \mathbf{0} + 2\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\mu} - \mathbf{0}. \end{aligned} \tag{5.12}$$

The first term on the right side is **0** because all third central moments of the multivariate normal are zero. The results for the other two terms do not depend on normality (see Problem 5.7). ■

$* \; y'Ay = (y-\mu+\mu)'A(y-\mu+\mu)$
$= (y-\mu)'A(y-\mu) + (y-\mu)'A\mu + \mu'A(y-\mu) + \mu'A\mu$
$= (y-\mu)'A(y-\mu) + 2(y-\mu)'A\mu + \mu'A\mu$

Corollary 1. Let $\mathbf{B}$ be a $k \times p$ matrix of constants. Then

$$\mathrm{cov}(\mathbf{By}, \mathbf{y}'\mathbf{Ay}) = 2\mathbf{B}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\mu}. \tag{5.13}$$

□

For the partitioned random vector $\mathbf{v} = \begin{pmatrix} \mathbf{y} \\ \mathbf{x} \end{pmatrix}$, the bilinear form $\mathbf{x}'\mathbf{Ay}$ was introduced in (2.34). The expected value of $\mathbf{x}'\mathbf{Ay}$ is given in the following theorem.

Theorem 5.2e. Let $\mathbf{v} = \begin{pmatrix} \mathbf{y} \\ \mathbf{x} \end{pmatrix}$ be a partitioned random vector with mean vector and covariance matrix given by (3.32) and (3.33)

$$E\begin{pmatrix} \mathbf{y} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \end{pmatrix} \quad \text{and} \quad \mathrm{cov}\begin{pmatrix} \mathbf{y} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{pmatrix},$$

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

$$E(\mathbf{x}'\mathbf{Ay}) = \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}_{yx}) + \boldsymbol{\mu}_x'\mathbf{A}\boldsymbol{\mu}_y. \tag{5.14}$$

PROOF. The proof is similar to that of Theorem 5.2a; see Problem 5.10. □

Example 5.2b. To estimate the population covariance $\sigma_{xy} = E[(x - \mu_x)(y - \mu_y)]$ in (3.10), we use the sample covariance

$$s_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n - 1}, \tag{5.15}$$

where $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ is a bivariate random sample from a population with means μ_x and μ_y, variances σ_x^2 and σ_y^2, and covariance σ_{xy}. We can write (5.15) in the form

$$s_{xy} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{n - 1} = \frac{\mathbf{x}'[\mathbf{I} - (1/n)\mathbf{J}]\mathbf{y}}{n - 1}, \tag{5.16}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$. Since (x_i, y_i) is independent of (x_j, y_j) for $i \neq j$, the random vector $\mathbf{v} = \begin{pmatrix} \mathbf{y} \\ \mathbf{x} \end{pmatrix}$ has mean vector and covariance matrix

$$E\begin{pmatrix} \mathbf{y} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \end{pmatrix} = \begin{pmatrix} \mu_y \mathbf{j} \\ \mu_x \mathbf{j} \end{pmatrix},$$

$$\mathrm{cov}\begin{pmatrix} \mathbf{y} \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{pmatrix} = \begin{pmatrix} \sigma_y^2 \mathbf{I} & \sigma_{xy}\mathbf{I} \\ \sigma_{xy}\mathbf{I} & \sigma_x^2 \mathbf{I} \end{pmatrix},$$

where each $\mathbf{I}$ is $n \times n$. Thus for use in (5.14), we have $\mathbf{A} = \mathbf{I} - (1/n)\mathbf{J}$, $\mathbf{\Sigma}_{yx} = \sigma_{xy}\mathbf{I}$, $\boldsymbol{\mu}_x = \mu_x\mathbf{j}$, and $\boldsymbol{\mu}_y = \mu_y\mathbf{j}$. Hence

$E(x'Ay) = tr(A\Sigma_{yx}) + \mu_x' A \mu_y$

$$
\begin{aligned}
E\left[\mathbf{x}'\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{y}\right] &= \mathrm{tr}\left[\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\sigma_{xy}\mathbf{I}\right] + \mu_x\mathbf{j}'\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mu_y\mathbf{j} \\
&= \sigma_{xy}\mathrm{tr}\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right) + \mu_x\mu_y\underbrace{\left(\mathbf{j}'\mathbf{j} - \frac{1}{n}\mathbf{j}'\mathbf{j}\mathbf{j}'\mathbf{j}\right)}_{n-n} \\
&= \sigma_{xy}(n-1) + 0.
\end{aligned}
$$

Therefore

$$
E(s_{xy}) = \frac{E\left[\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})\right]}{n-1} = \frac{(n-1)\sigma_{xy}}{n-1} = \sigma_{xy}. \tag{5.17}
$$

∎

카이제곱 분포.

5.3 NONCENTRAL CHI-SQUARE DISTRIBUTION

Before discussing the noncentral chi-square distribution, we first review the central chi-square distribution. Let $z_1, z_2, \ldots, z_n$ be a random sample from the standard normal distribution $N(0, 1)$. Since the z's are independent (by definition of random sample) and each z_i is $N(0, 1)$, the random vector $\mathbf{z} = (z_1, z_2, \ldots, z_n)'$ is distributed as $N_n(\mathbf{0}, \mathbf{I})$. By definition

$z \sim N_n(0, I)$, $z'z \sim \chi^2(n)$

$$
\sum_{i=1}^{n} z_i^2 = \mathbf{z}'\mathbf{z} \text{ is } \chi^2(n); \tag{5.18}
$$

that is, the sum of squares of n independent standard normal random variables is distributed as a (central) chi-square random variable with n degrees of freedom.

The mean, variance, and moment generating function of a chi-square random variable are given in the following theorem.

Theorem 5.3a. If u is distributed as $\chi^2(n)$, then

카이제곱 분포의 평균, 분산, 적률 생성 함수

$$
E(u) = n, \tag{5.19}
$$

$$
\mathrm{var}(u) = 2n, \tag{5.20}
$$

$$
M_u(t) = \frac{1}{(1-2t)^{n/2}}. \tag{5.21}
$$

$u = z'Iz \sim \chi^2(n)$

Proof. Since u is the quadratic form $\mathbf{z}'\mathbf{I}\mathbf{z}$, $E(u)$, $\mathrm{var}(u)$, and $M_u(t)$ can be obtained by applying Theorems 5.2a, 5.2c, and 5.2b, respectively. ∎

Now suppose that $y_1, y_2, \ldots, y_n$ are independently distributed as $N(\mu_i, 1)$, so that $\mathbf{y}$ is $N_n(\boldsymbol{\mu}, \mathbf{I})$, where $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)'$. In this case, $\Sigma_{i=1}^n y_i^2 = \mathbf{y}'\mathbf{y}$ does not have a chi-square distribution, but $\Sigma(y_i - \mu_i)^2 = (\mathbf{y} - \boldsymbol{\mu})'(\mathbf{y} - \boldsymbol{\mu})$ is $\chi^2(n)$ since $y_i - \mu_i$ is distributed as $N(0,1)$.

The density of $v = \Sigma_{i=1}^n y_i^2 = \mathbf{y}'\mathbf{y}$, where the y's are independently distributed as $N(\mu_i, 1)$, is called the *noncentral chi-square distribution* and is denoted by $\chi^2(n, \lambda)$. The *noncentrality parameter* λ is defined as

$$\lambda = \frac{1}{2}\sum_{i=1}^n \mu_i^2 = \frac{1}{2}\boldsymbol{\mu}'\boldsymbol{\mu}. \tag{5.22}$$

Note that λ is not an eigenvalue here and that the mean of $v = \Sigma_{i=1}^n y_i^2$ is greater than the mean of $u = \Sigma_{i=1}^n (y_i - \mu_i)^2$: v의 평균 > u의 평균.

$$E\left[\sum_{i=1}^n (y_i - \mu_i)^2\right] = \sum_{i=1}^n E(y_i - \mu_i)^2 = \sum_{i=1}^n \mathrm{var}(y_i) = \sum_{i=1}^n 1 = n,$$

$$\begin{aligned} E\left(\sum_{i=1}^n y_i^2\right) &= \sum_{i=1}^n E(y_i^2) = \sum_{i=1}^n (\sigma_i^2 + \mu_i^2) = \sum_{i=1}^n (1 + \mu_i^2) \\ &= n + \sum_{i=1}^n \mu_i^2 = n + 2\lambda, \end{aligned}$$

2λ만큼 더 큼.

where λ is as defined in (5.22). The densities of u and v are illustrated in Figure 5.1.

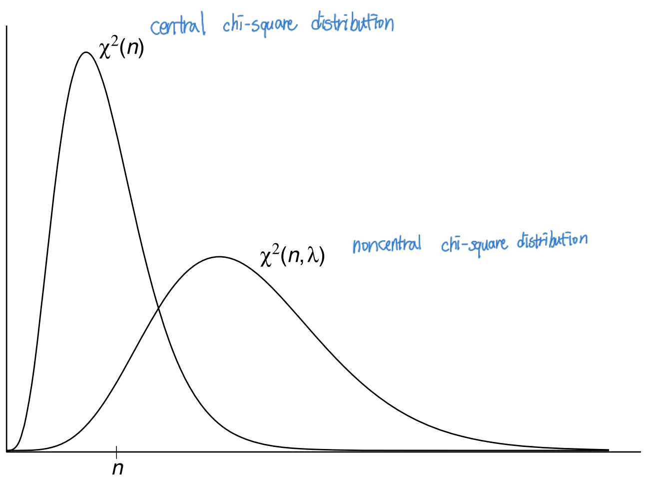

Figure 5.1 Central and noncentral chi-square densities.

The mean, variance, and moment generating function of a noncentral chi-square random variable are given in the following theorem.

Theorem 5.3b. If v is distributed as $\chi^2(n, \lambda)$, then

$$E(v) = n + 2\lambda, \tag{5.23}$$

$$\mathrm{var}(v) = 2n + 8\lambda, \tag{5.24}$$

$$M_v(t) = \frac{1}{(1-2t)^{n/2}} e^{-\lambda[1-1/(1-2t)]}. \tag{5.25}$$

Proof. For $E(v)$ and $\mathrm{var}(v)$, see Problems 5.13 and 5.14. For $M_v(t)$, use Theorem 5.2b. ∎

↳ $var(y'Ay) = 2tr(A\Sigma)^2 + 4\mu'A\Sigma A\mu$ 이용.

$var(y'Iy) = 2tr(I)^2 + 4\mu'\mu = 2n + 8\lambda$

Corollary 1. If $\lambda = 0$ (which corresponds to $\mu_i = 0$ for all i), then $E(v)$, $\mathrm{var}(v)$, and $M_v(t)$ in Theorem 5.3b reduce to $E(u)$, $\mathrm{var}(u)$, $M_u(t)$ for the central chi-square distribution in Theorem 5.3a. Thus

$$\chi^2(n, 0) = \chi^2(n). \tag{5.26}$$

∎

The chi-square distribution has an additive property, as shown in the following theorem.

〈 카이제곱분포의 성질들. 〉

Theorem 5.3c. If $v_1, v_2, \ldots, v_k$ are independently distributed as $\chi^2(n_i, \lambda_i)$, then

합의 분포.

$$\sum_{i=1}^{k} v_i \text{ is distributed as } \chi^2\left(\sum_{i=1}^{k} n_i, \sum_{i=1}^{k} \lambda_i\right). \tag{5.27}$$

∎

Corollary 1. If $u_1, u_2, \ldots, u_k$ are independently distributed as $\chi^2(n_i)$, then

$$\sum_{i=1}^{k} u_i \text{ is distributed as } \chi^2\left(\sum_{i=1}^{k} n_i\right).$$

∎

5.4 NONCENTRAL F AND t DISTRIBUTIONS

5.4.1 Noncentral F Distribution

Before defining the noncentral F distribution, we first review the central F. If u is $\chi^2(p)$, v is $\chi^2(q)$, and u and v are independent, then by definition

$u \sim \chi^2(p)$, $v \sim \chi^2(q)$

$$w = \frac{u/p}{v/q} \text{ is distributed as } F(p, q), \tag{5.28}$$

$w = \frac{u/p}{v/q} \sim F(p,q)$

the (central) F distribution with p and q degress of freedom. The mean and variance of w are given by

$$E(w) = \frac{q}{q-2}, \qquad \mathrm{var}(w) = \frac{2q^2(p+q-2)}{p(q-1)^2(q-4)}. \tag{5.29}$$

Now suppose that u is distributed as a noncentral chi-square random variable, $\chi^2(p, \lambda)$, while v remains central chi-square random variable, $\chi^2(q)$, with u and v independent. Then

$u \sim \chi^2(p, \lambda)$, $v \sim \chi^2(q)$

$$z = \frac{u/p}{v/q} \text{ is distributed as } F(p, q, \lambda), \tag{5.30}$$

the *noncentral F distribution* with noncentrality parameter λ, where λ is the same noncentrality parameter as in the distribution of u (noncentral chi-square distribution). The mean of z is

$$E(z) = \frac{q}{q-2}\left(1 + \frac{2\lambda}{p}\right), \quad > E(w) = \frac{q}{q-2} \tag{5.31}$$

which is, course, greater than $E(w)$ in (5.29).

When an F statistic is used to test a hypothesis H_0, the distribution will typically be central if the (null) hypothesis is true and noncentral if the hypothesis is false. Thus the noncentral F distribution can often be used to evaluate the power of an F test. The *power* of a test is the probability of rejecting H_0 for a given value of λ. If F_α is the upper α percentage point of the central F distribution, then the power, $P(p, q, \alpha, \lambda)$, can be defined as

H_0를 기각할 확률

$$P(p, q, \alpha, \lambda) = \mathrm{Prob}\ (z \geq F_\alpha), \tag{5.32}$$

q, α, λ가 커지면 이 확률값도 커짐, p가 커지면, 이 확률값은 작아짐!

$z \sim F(p, q, \lambda)$

where z is the noncentral F random variable defined in (5.30). Ghosh (1973) showed that $P(p, q, \alpha, \lambda)$ increases if q or α or λ increases, and $P(p, q, \alpha, \lambda)$ decreases if p increases. The power is illustrated in Figure 5.2.

The power as defined in (5.32) can be evaluated from tables (Tiku 1967) or directly from distribution functions available in many software packages. For example, in SAS, the noncentral F-distribution function PROBF can be used to find the power in (5.32) as follows:

$$P(p, q, \alpha, \lambda) = 1 - \mathrm{PROBF}(F_\alpha, p, q, \lambda).$$

A probability calculator for the F and other distributions is available free of charge from NCSS (download at www.ncss.com).

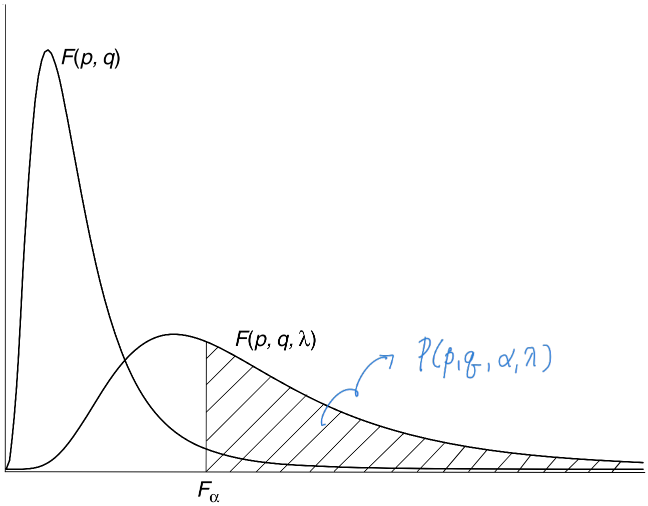

Figure 5.2 Central F, noncentral F, and power of the F test (shaded area).

5.4.2 Noncentral t Distribution

We first review the central t distribution. If z is $N(0,1)$, u is $\chi^2(p)$, and z and u are independent, then by definition

$z \sim N(0,1)$, $u \sim \chi^2(p)$

$$t = \frac{z}{\sqrt{u/p}} \text{ is distributed as } t(p), \tag{5.33}$$

the (central) t distribution with p degrees of freedom.

Now suppose that y is $N(\mu, 1)$, u is $\chi^2(p)$, and y and u are independent. Then

$y \sim N(\mu, 1)$, $u \sim \chi^2(p)$

$$t = \frac{y}{\sqrt{u/p}} \text{ is distributed as } t(p, \mu), \tag{5.34}$$

the noncentral t distribution with p degrees of freedom and noncentrality parameter μ. If y is $N(\mu, \sigma^2)$, then

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

since by (3.4), (3.9), and Theorem 4.4a(i), y/σ is distributed as $N(\mu/\sigma, 1)$.

5.5 DISTRIBUTION OF QUADRATIC FORMS

It was noted following Theorem 5.3a that if $\mathbf{y}$ is $N_n(\boldsymbol{\mu}, \mathbf{I})$, then $(\mathbf{y}-\boldsymbol{\mu})'(\mathbf{y}-\boldsymbol{\mu})$ is $\chi^2(n)$. If $\mathbf{y}$ is $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we can extend this to

$$(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu}) \text{ is } \chi^2(n). \tag{5.35}$$

To show this, we write $(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})$ in the form

$$\begin{aligned}(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu}) &= (\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{-1/2}(\mathbf{y}-\boldsymbol{\mu}) \\ &= \left[\boldsymbol{\Sigma}^{-1/2}(\mathbf{y}-\boldsymbol{\mu})\right]'\left[\boldsymbol{\Sigma}^{-1/2}(\mathbf{y}-\boldsymbol{\mu})\right] \\ &= \mathbf{z}'\mathbf{z},\end{aligned}$$

($A=CDC$ 일때.)

$A^{1/2}=CD^{1/2}C'$

where $\mathbf{z}=\boldsymbol{\Sigma}^{-1/2}(\mathbf{y}-\boldsymbol{\mu})$ and $\boldsymbol{\Sigma}^{-1/2}=(\boldsymbol{\Sigma}^{1/2})^{-1}$, with $\boldsymbol{\Sigma}^{1/2}$ given by (2.109). The vector $\mathbf{z}$ is distributed as $N_n(\mathbf{0}, \mathbf{I})$ (see Problem 5.17); therefore, $\mathbf{z}'\mathbf{z}$ is $\chi^2(n)$ by definition [see (5.18)]. Note the analogy of $(\mathbf{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})$ to the univariate random variable $(y-\mu)^2/\sigma^2$, which is distributed as $\chi^2(1)$ if y is $N(\mu, \sigma^2)$.

In the following theorem, we consider the distribution of quadratic forms in general. In the proof we follow Searle (1971, p. 57). For alternative proofs, see Graybill (1976, pp. 134–136) and Hocking (1996, p. 51).

Theorem 5.5. Let $\mathbf{y}$ be distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, let $\mathbf{A}$ be a symmetric matrix of constants of rank r, and let $\lambda=\frac{1}{2}\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$. Then $\mathbf{y}'\mathbf{A}\mathbf{y}$ is $\chi^2(r, \lambda)$, if and only if $\mathbf{A}\boldsymbol{\Sigma}$ is idempotent.

$y \sim N_p(\mu, \Sigma)$. A : rank$(A)=r$ 인 대칭행렬. ($\lambda=\frac{1}{2}\mu' A\mu$)

$y'Ay \sim \chi^2(r,\lambda)$ $\Longleftrightarrow$ $A\Sigma$가 멱등행렬.

Proof. By Theorem 5.2b the moment generating function of $\mathbf{y}'\mathbf{A}\mathbf{y}$ is

$$M_{\mathbf{y}'\mathbf{A}\mathbf{y}}(t) = |\mathbf{I}-2t\mathbf{A}\boldsymbol{\Sigma}|^{-1/2}e^{-(1/2)\boldsymbol{\mu}'[\mathbf{I}-(\mathbf{I}-2t\mathbf{A}\boldsymbol{\Sigma})^{-1}]\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}.$$

$A\Sigma$의 고유값

By (2.98), the eigenvalues of $\mathbf{I}-2t\mathbf{A}\boldsymbol{\Sigma}$ are $1-2t\lambda_i$, $i=1,2,\ldots,p$, where λ_i is an eigenvalue of $\mathbf{A}\boldsymbol{\Sigma}$. By (2.107), $|\mathbf{I}-2t\mathbf{A}\boldsymbol{\Sigma}|=\prod_{i=1}^{p}(1-2t\lambda_i)$. By (2.102), $(\mathbf{I}-2t\mathbf{A}\boldsymbol{\Sigma})^{-1}=\mathbf{I}+\sum_{k=1}^{\infty}(2t)^k(\mathbf{A}\boldsymbol{\Sigma})^k$, provided $-1<2t\lambda_i<1$ for all i. Thus $M_{\mathbf{y}'\mathbf{A}\mathbf{y}}(t)$ can be written as

$(I-A)^{-1}=I+A+A^2+A^3+\cdots$

$$M_{\mathbf{y}'\mathbf{A}\mathbf{y}}(t) = \left(\prod_{i=1}^{p}(1-2t\lambda_i)^{-1/2}\right)e^{-(1/2)\boldsymbol{\mu}'\left[-\sum_{k=1}^{\infty}(2t)^k(\mathbf{A}\boldsymbol{\Sigma})^k\right]\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}.$$

λ_i 중 r개는 1, 나머지 고유값은 0.

Suppose that $\mathbf{A\Sigma}$ is idempotent of rank r (the rank of $\mathbf{A}$); then r of the λ_i's are equal to 1, $p - r$ of the λ_i's are equal to 0, and $(\mathbf{A\Sigma})^k = \mathbf{A\Sigma}$. Therefore,

$$M_{\mathbf{y}'\mathbf{Ay}}(t) = \left(\prod_{i=1}^{r} (1-2t)^{-1/2}\right) e^{-(1/2)\boldsymbol{\mu}'\left[-\sum_{k=1}^{\infty}(2t)^k\right]\mathbf{A\Sigma\Sigma}^{-1}\boldsymbol{\mu}}$$

$$= (1-2t)^{-r/2} e^{-1/2\boldsymbol{\mu}'[1-(1-2t)-1]\mathbf{A}\boldsymbol{\mu}},$$

p $2t\lambda_i$ $1-(1-2t)^{-1}$ $(2t)^k(A\Sigma)^k \to (2t)^k A\Sigma$.

오타. $(1-2t)^{-1}$ 가 맞음.

* 무한 등비급수 사용.

$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$, for $|k|<1$

$\sum_{k=1}^{\infty}(2t)^k = \sum_{k=0}^{\infty}(2t)^k - 1 = \frac{1}{1-2t} - 1 = -1 + (1-2t)^{-1}$

provided $-1 < 2t < 1$ or $-\frac{1}{2} < t < \frac{1}{2}$, which is compatible with the requirement that the moment generating function exists for t in a neighborhood of 0. Thus

$$M_{\mathbf{y}'\mathbf{Ay}}(t) = \frac{1}{(1-2t)^{r/2}} e^{-(1/2)\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}[1-1/(1-2t)]},$$

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

$M_v(t) = \frac{1}{(1-2t)^{n/2}} e^{-\lambda[1-1/(1-2t)]}$

which by (5.25) is the moment generating function of a noncentral chi-square random variable with degrees of freedom $r = \text{rank}(\mathbf{A})$ and noncentrality parameter $\lambda = \frac{1}{2}\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$.

For a proof of the converse, namely, if $\mathbf{y}'\mathbf{Ay}$ is $\chi^2(r, \lambda)$, then $\mathbf{A\Sigma}$ is idempotent; see Driscoll (1999). ∎

Some corollaries of interest are the following (for additional corollaries, see Problem 5.20).

Corollary 1. If $\mathbf{y}$ is $N_p(\mathbf{0}, \mathbf{I})$, then $\mathbf{y}'\mathbf{Ay}$ is $\chi^2(r)$ if and only if $\mathbf{A}$ is idempotent of rank r. ∎

$y \sim N_p(0, I)$ 일 때, $y'Ay \sim \chi^2(r) \iff$ A는 계수가 r인 멱등행렬

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

$y \sim N_p(\mu, \sigma^2 I)$ 일 때, $\frac{y'Ay}{\sigma^2} \sim \chi^2(r, \frac{\mu'A\mu}{2\sigma^2}) \iff$ A는 계수가 r인 멱등행렬

Example 5. To illustrate Corollary 2 to Theorem 5.5, consider the distribution of $(n-1)s^2/\sigma^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2/\sigma^2$, where $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ is distributed as $N_n(\mu\mathbf{j}, \sigma^2\mathbf{I})$ as in Examples 5.1 and 5.2. In (5.2) we have $\sum_{i=1}^{n}(y_i - \bar{y})^2 = \mathbf{y}'[\mathbf{I} - (1/n)\mathbf{J}]\mathbf{y}$. The matrix $\mathbf{I} - (1/n)\mathbf{J}$ is shown to be idempotent in Problem 5.2. Then by Theorem 2.13d, rank $[\mathbf{I} - (1/n)\mathbf{J}] = \text{tr}[\mathbf{I} - (1/n)\mathbf{J}] = n - 1$. We next find λ, which is given by

$$\lambda = \frac{\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}}{2\sigma^2} = \frac{\mu\mathbf{j}'(\mathbf{I} - \frac{1}{n}\mathbf{J})\mu\mathbf{j}}{2\sigma^2} = \frac{\mu^2(\mathbf{j}'\mathbf{j} - \frac{1}{n}\mathbf{j}'\mathbf{Jj})}{2\sigma^2}$$

$$= \frac{\mu^2(n - \frac{1}{n}\mathbf{j}'\mathbf{jj}'\mathbf{j})}{2\sigma^2} = \frac{\mu^2[n - \frac{1}{n}(n)(n)]}{2\sigma^2} = 0.$$

$\sum_{i=1}^{n}(y_i - \bar{y})^2/\sigma^2$

Therefore, $\mathbf{y}'[\mathbf{I} - (1/n)\mathbf{J}]\mathbf{y}/\sigma^2$ is $\chi^2(n-1)$. ∎

$\chi^2(n-1, 0) = \chi^2(n-1)$

5.6 INDEPENDENCE OF LINEAR FORMS AND QUADRATIC FORMS

In this section, we discuss the independence of (1) a linear form and a quadratic form, (2) two quadratic forms, and (3) several quadratic forms.

For an example of (1), consider $\bar{y}$ and s^2 in a simple random sample or $\hat{\beta}$ and s^2 in a regression setting. To illustrate (2), consider the sum of squares due to regression and the sum of squares due to error. An example of (3) is given by the sums of squares due to main effects and interaction in a balanced two-way analysis of variance.

균형 잡힌 2원 배치법, ANOVA

We begin with the independence of a linear form and a quadratic form.

Theorem 5.6a. Suppose that $\mathbf{B}$ ($k \times p$) is a $k \times p$ matrix of constants, $\mathbf{A}$ ($p \times p$) is a $p \times p$ symmetric matrix of constants, and $\mathbf{y}$ is distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then $\mathbf{By}$ and $\mathbf{y}'\mathbf{Ay}$ are independent if and only if $\mathbf{B\Sigma A} = \mathbf{O}$.

By와 $y'Ay$는 독립, $\Leftrightarrow$ $B\Sigma A = O$.

Proof. Suppose $\mathbf{B\Sigma A} = \mathbf{O}$. We prove that $\mathbf{By}$ and $\mathbf{y}'\mathbf{Ay}$ are independent for the special case in which $\mathbf{A}$ is symmetric and idempotent. For a general proof, see Searle (1971, p. 59).

Assuming that $\mathbf{A}$ is symmetric and idempotent, $\mathbf{y}'\mathbf{Ay}$ can be written as

$$\mathbf{y}'\mathbf{Ay} = \mathbf{y}'\mathbf{A}'\mathbf{Ay} = (\mathbf{Ay})'\mathbf{Ay}.$$

$\to cov(z,w) = cov(Ay, By) = A\Sigma B'$

If $\mathbf{B\Sigma A} = \mathbf{O}$, we have by (3.45)

$$\mathbf{B\Sigma A} = \text{cov}(\mathbf{By}, \mathbf{Ay}) = \mathbf{O}.$$

$y \sim N_p(\mu, \Sigma)$, $cov(Ay, By) = A\Sigma B' = O \Rightarrow$ Ay와 By는 독립.

Hence, by Corollary 2 to Theorem 4.4c, $\mathbf{By}$ and $\mathbf{Ay}$ are independent, and therefore $\mathbf{By}$ and the function $(\mathbf{Ay})'\mathbf{Ay}$ are also independent (Seber 1977, pp. 17, 33–34).

We now establish the converse, namely, if $\mathbf{By}$ and $\mathbf{y}'\mathbf{Ay}$ are independent, then $\mathbf{B\Sigma A} = \mathbf{O}$. By Corollary 1 to Theorem 5.2d, cov$(\mathbf{By}, \mathbf{y}'\mathbf{Ay}) = \mathbf{0}$ becomes

$cov(By, y'Ay) = 2B\Sigma A\mu$

$$2\mathbf{B\Sigma A}\boldsymbol{\mu} = \mathbf{0}.$$

Since this holds for all possible $\boldsymbol{\mu}$, we have $\mathbf{B\Sigma A} = \mathbf{O}$ [see (2.44)]. ■

Note that $\mathbf{B\Sigma A} = \mathbf{O}$ does not imply $\mathbf{A\Sigma B} = \mathbf{O}$. In fact, the product $\mathbf{A\Sigma B}$ will not be defined unless $\mathbf{B}$ has p rows.

Corollary 1. If $\mathbf{y}$ is $N_p(\boldsymbol{\mu}, \sigma^2\mathbf{I})$, then $\mathbf{By}$ and $\mathbf{y}'\mathbf{Ay}$ are independent if and only if $\mathbf{BA} = \mathbf{O}$. ■

$y \sim N_p(\mu, \sigma^2 I)$ 일 때, By와 $y'Ay$는 독립 $\Leftrightarrow$ $BA = O$.

$y \sim N_n(\mu \mathbf{j}, \sigma^2 I)$

Example 5.6a. To illustrate Corollary 1, consider $s^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2/(n-1)$ and $\bar{y} = \sum_{i=1}^{n} y_i/n$, where $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ is $N_n(\mu\mathbf{j}, \sigma^2\mathbf{I})$. As in Example 5.1, $\bar{y}$ and s^2 can be written as $\bar{y} = (1/n)\mathbf{j}'\mathbf{y}$ and $s^2 = \mathbf{y}'[\mathbf{I} - (1/n)\mathbf{J}]\mathbf{y}/(n-1)$. By Corollary 1, $\bar{y}$ is independent of s^2 since $\underbrace{(1/n)\mathbf{j}'}_{A}\underbrace{[\mathbf{I} - (1/n)\mathbf{J}]}_{B} = \mathbf{0}'$. ∎

$(\frac{1}{n} \cdots \frac{1}{n}) - \frac{1}{n^2}\underbrace{(1 \cdots 1)\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}}_{(n \cdots n)} = (\frac{1}{n} \cdots \frac{1}{n}) - (\frac{1}{n} \cdots \frac{1}{n})$

We now consider the independence of two quadratic forms.

Theorem 5.6b. Let $\mathbf{A}$ and $\mathbf{B}$ be symmetric matrices of constants. If $\mathbf{y}$ is $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ are independent if and only if $\mathbf{A}\boldsymbol{\Sigma}\mathbf{B} = \mathbf{O}$.

$y \sim N_p(\mu, \Sigma)$ 일 때, $y'Ay$ 와 $y'By$는 독립 $\Leftrightarrow A\Sigma B = O$.

Proof. Suppose $\mathbf{A}\boldsymbol{\Sigma}\mathbf{B} = \mathbf{O}$. We prove that $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ are independent for the special case in which $\mathbf{A}$ and $\mathbf{B}$ are symmetric and idempotent. For a general proof, see Searle (1971, pp. 59–60) or Hocking (1996, p. 52).

Assuming that $\mathbf{A}$ and $\mathbf{B}$ are symmetric and idempotent, $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ can be written as $\mathbf{y}'\mathbf{A}\mathbf{y} = \mathbf{y}'\mathbf{A}'\mathbf{A}\mathbf{y} = (\mathbf{A}\mathbf{y})'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y} = \mathbf{y}'\mathbf{B}'\mathbf{B}\mathbf{y} = (\mathbf{B}\mathbf{y})'\mathbf{B}\mathbf{y}$. If $\mathbf{A}\boldsymbol{\Sigma}\mathbf{B} = \mathbf{O}$, we have [see (3.45)]

$$\mathbf{A}\boldsymbol{\Sigma}\mathbf{B} = \text{cov}(\mathbf{A}\mathbf{y}, \mathbf{B}\mathbf{y}) = \mathbf{O}.$$

Hence, by Corollary 2 to Theorem 4.4c, $\mathbf{A}\mathbf{y}$ and $\mathbf{B}\mathbf{y}$ are independent. It follows that the functions $(\mathbf{A}\mathbf{y})'(\mathbf{A}\mathbf{y}) = \mathbf{y}'\mathbf{A}\mathbf{y}$ and $(\mathbf{B}\mathbf{y})'(\mathbf{B}\mathbf{y}) = \mathbf{y}'\mathbf{B}\mathbf{y}$ are independent (Seber 1977, pp. 17, 33–34). ∎

Note that $\mathbf{A}\boldsymbol{\Sigma}\mathbf{B} = \mathbf{O}$ is equivalent to $\mathbf{B}\boldsymbol{\Sigma}\mathbf{A} = \mathbf{O}$ since transposing both sides of $\mathbf{A}\boldsymbol{\Sigma}\mathbf{B} = \mathbf{O}$ gives $\mathbf{B}\boldsymbol{\Sigma}\mathbf{A} = \mathbf{O}$ ($\mathbf{A}$ and $\mathbf{B}$ are symmetric).

Corollary 1. If $\mathbf{y}$ is $N_p(\boldsymbol{\mu}, \sigma^2\mathbf{I})$, then $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ are independent if and only if $\mathbf{A}\mathbf{B} = \mathbf{O}$ (or, equivalently, $\mathbf{B}\mathbf{A} = \mathbf{O}$). ∎

Example 5.6b. To illustrate Corollary 1, consider the partitioning in (5.1), $\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2 + n\bar{y}^2$, which was expressed in (5.3) as

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

If $\mathbf{y}$ is $N_n(\mu\mathbf{j}, \sigma^2\mathbf{I})$, then by Corollary 1, $\mathbf{y}'[\mathbf{I} - (1/n)\mathbf{J}]\mathbf{y}$ and $\mathbf{y}'[(1/n)\mathbf{J}]\mathbf{y}$ are independent if and only if $[\mathbf{I} - (1/n)\mathbf{J}][(1/n)\mathbf{J}] = \mathbf{O}$, which is shown in Problem 5.2. ∎

The distribution and independence of several quadratic forms are considered in the following theorem.

Theorem 5.6c. Let $\mathbf{y}$ be $N_n(\boldsymbol{\mu}, \sigma^2\mathbf{I})$, let $\mathbf{A}_i$ be symmetric of rank r_i for $i = 1, 2, \ldots, k$, and let $\mathbf{y}'\mathbf{A}\mathbf{y} = \sum_{i=1}^k \mathbf{y}'\mathbf{A}_i\mathbf{y}$, where $\mathbf{A} = \sum_{i=1}^k \mathbf{A}_i$ is symmetric of rank r. Then

(i) $\mathbf{y}'\mathbf{A}_i\mathbf{y}/\sigma^2$ is $\chi^2(r_i, \boldsymbol{\mu}'\mathbf{A}_i\boldsymbol{\mu}/2\sigma^2)$, $i = 1, 2, \ldots, k$.
(ii) $\mathbf{y}'\mathbf{A}_i\mathbf{y}$ and $\mathbf{y}'\mathbf{A}_j\mathbf{y}$ are independent for all $i \neq j$.
(iii) $\mathbf{y}'\mathbf{A}\mathbf{y}/\sigma^2$ is $\chi^2(r, \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}/2\sigma^2)$.

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

이 중 2개 이상 만족시키는 것과 필요충분 조건,,

(a) Each $\mathbf{A}_i$ is idempotent.
(b) $\mathbf{A}_i\mathbf{A}_j = \mathbf{O}$ for all $i \neq j$.
(c) $\mathbf{A} = \sum_{i=1}^k \mathbf{A}_i$ is idempotent.

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

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

Proof. See Searle (1971, pp. 61–64). □

A : $n \times n$ 대칭행렬, $A = \sum_{i=1}^{k} A_i$ (A_i : $n \times n$ 대칭행렬). $\Rightarrow$ $\left\{ \begin{array}{l} A\text{는 멱등.} \\ A_1 \cdots A_k \text{ 도 멱등.} \\ A_iA_j = O. \text{ for } i \neq j \end{array} \right.$

Note that by Theorem 2.13g, any two of (a), (b), or (c) implies the third.

Theorem 5.6c pertains to partitioning a sum of squares into several component sums of squares. The following corollary treats the special case where $\mathbf{A} = \mathbf{I}$; that is, the case of partitioning the total sum of squares $\mathbf{y}'\mathbf{y}$ into several sums of squares.

Corollary 1. Let $\mathbf{y}$ be $N_n(\boldsymbol{\mu}, \sigma^2\mathbf{I})$, let $\mathbf{A}_i$ be symmetric of rank r_i for $i = 1, 2, \ldots, k$, and let $\mathbf{y}'\mathbf{y} = \sum_{i=1}^k \mathbf{y}'\mathbf{A}_i\mathbf{y}$. Then (i) each $\mathbf{y}'\mathbf{A}_i\mathbf{y}/\sigma^2$ is $\chi^2(r_i, \boldsymbol{\mu}'\mathbf{A}_i\boldsymbol{\mu}/2\sigma^2)$ and (ii) the $\mathbf{y}'\mathbf{A}_i\mathbf{y}$ terms are mutually independent if and only if any *one* of the following statements holds:

(a) Each $\mathbf{A}_i$ is idempotent.
(b) $\mathbf{A}_i\mathbf{A}_j = \mathbf{O}$ for all $i \neq j$.
(c) $n = \sum_{i=1}^k r_i$. □

Note that by Theorem 2.13h, condition (c) implies the other two conditions. Cochran (1934) first proved a version of Corollary 1 to Theorem 5.6c.

PROBLEMS

5.1 Show that $\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n\bar{y}^2$ as in (5.1).