

Notes for 751-752

Section 13

Martin Lindquist*

October 10, 2017

*Thanks to Brian Caffo and Ingo Ruczinski for supplying their notes, upon which much of this document is based.

13 Quadratic forms

We are often interested in working with random vectors that are combined into quadratic forms. The result is a function of random variables which is a scalar, and itself a random variable. Formally, a quadratic form is a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ that can be written:

$$f(\mathbf{y}) = \mathbf{y}' \mathbf{A} \mathbf{y} = \sum_{i,j} A_{ij} y_i y_j.$$

We have previously discussed how to compute the expected value of quadratic forms. Here we discuss their distributional properties and how they apply to linear models. But first we review some distributional results.

13.1 Distributional results

13.1.1 χ^2 distribution

A random variable U has a (central) χ^2 distribution with $n > 0$ degrees of freedom if it has a pdf given by

$$f_U(u|n) = \frac{1}{\Gamma(\frac{n}{2}) 2^{n/2}} u^{\frac{n}{2}-1} e^{-u/2} I(u > 0)$$

We write $U \sim \chi_n^2$. The moment generating function of U is given by

$$M_U(t) = (1 - 2t)^{-n/2}$$

for $t < 1/2$.

Below follow some important results related to the χ^2 -distribution.

Theorem: 13.1 If $Z \sim N(0, 1)$, then $U = Z^2 \sim \chi^2(1)$.

Theorem: 13.2 If $\mathbf{y} \sim N_p(\mathbf{0}, \mathbf{I})$, then $\mathbf{y}' \mathbf{y} \sim \chi^2(p)$.

Theorem: 13.3 If $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \Sigma)$, then $(\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \sim \chi^2(p)$.

13.1.2 Noncentral χ^2 distribution

A random variable V has a noncentral χ^2 distribution with $n > 0$ degrees of freedom and noncentrality parameter $\lambda > 0$ if it has a pdf given by

$$f_V(v|n, \lambda) = \sum_{j=0}^{\infty} \left(\frac{e^{-\lambda} \lambda^j}{j!} \right) f_U(v|n + 2j)$$

We write $V \sim \chi_n^2(\lambda)$. The moment generating function of V is given by

$$M_V(t) = (1 - 2t)^{-n/2} \exp \left(\frac{2t\lambda}{1 - 2t} \right)$$

for $t < 1/2$.

Below follow some important results related to the non-central χ^2 -distribution.

Theorem: 13.4 If $Y \sim N(\mu, 1)$, then $U = Y^2 \sim \chi_1^2(\lambda)$ where $\lambda = \mu^2/2$.

Theorem: 13.5 If U_1, U_2, \dots, U_m are independent random variables, where $U_i \sim \chi^2(n_i, \lambda_i)$ for $i = 1, 2, \dots, m$, then $U = \sum_i U_i \sim \chi^2(n, \lambda)$, where $n = \sum_i n_i$ and $\lambda = \sum_i \lambda_i$.

13.1.3 F distribution

A random variable W has a (central) F distribution with $m_1 > 0$ and $m_2 > 0$ degrees of freedom if it has a pdf given by

$$f_W(w|m_1, m_2) = \frac{\Gamma \left(\frac{m_1+m_2}{2} \right) \left(\frac{m_1}{m_2} \right)^{m_1/2} w^{(m_1-2)/2}}{\Gamma \left(\frac{m_1}{2} \right) \Gamma \left(\frac{m_2}{2} \right) \left(1 + \frac{m_1 w}{m_2} \right)^{(m_1+m_2)/2}}$$

for $0 < w < \infty$. We write $W \sim F_{m_1, m_2}$.

If U_1 and U_2 are independent central χ^2 random variables with degrees of freedom m_1 and m_2 , respectively then

$$W = \frac{U_1/m_1}{U_2/m_2} \sim F_{m_1, m_2}.$$

13.1.4 Noncentral F distribution

A random variable W has a noncentral F distribution with $m_1 > 0$ and $m_2 > 0$ degrees of freedom and noncentrality parameter $\lambda > 0$ if it has a pdf given by

$$f_W(w|m_1, m_2, \lambda) = \sum_{j=0}^{\infty} \left(\frac{e^{-\lambda} \lambda^j}{j!} \right) \frac{\Gamma\left(\frac{m_1+2j+m_2}{2}\right) \left(\frac{m_1+2j}{m_2}\right)^{(m_1+2j)/2} w^{(m_1+2j+2)/2}}{\Gamma\left(\frac{m_1+2j}{2}\right) \Gamma\left(\frac{m_2}{2}\right) \left(1 + \frac{m_1 w}{m_2}\right)^{(m_1+2j+m_2)/2}}$$

for $0 < w < \infty$. We write $W \sim F_{m_1, m_2}(\lambda)$.

If $U_1 \sim \chi_{m_1}^2(\lambda)$ and $U_2 \sim \chi_{m_2}^2$ be independent, then

$$W = \frac{U_1/m_1}{U_2/m_2} \sim F_{m_1, m_2}(\lambda).$$

13.2 Distribution of quadratic forms

In linear model theory, test statistics arise from sums of squares (which are special cases of quadratic forms) with χ^2 distributions. Below follows some useful results regarding quadratic forms that will be extremely useful when we start performing hypothesis testing.

Theorem: 13.6 If $\mathbf{y} \sim N_p(\mathbf{0}, \mathbf{I})$ and \mathbf{A} is symmetric idempotent of rank r then

$$\mathbf{y}'\mathbf{A}\mathbf{y} \sim \chi^2(r).$$

Theorem: 13.7 If $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \mathbf{I})$ then

$$\mathbf{y}'\mathbf{A}\mathbf{y} \sim \chi^2(r, \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}/2)$$

if and only if \mathbf{A} is symmetric idempotent of rank r

Theorem: 13.8 If $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \Sigma)$ then

$$\mathbf{y}'\mathbf{A}\mathbf{y} \sim \chi^2(r, \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}/2)$$

if and only if $\mathbf{A}\Sigma$ is idempotent of rank r

Example: 13.9 To illustrate consider the linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$. Let $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ and consider the following partition of the sum of squares:

$$\mathbf{y}'\mathbf{y} = \mathbf{y}'\mathbf{H}\mathbf{y} + \mathbf{y}'(\mathbf{I} - \mathbf{H})\mathbf{y}.$$

Let us begin by studying the term $\mathbf{y}'(\mathbf{I} - \mathbf{H})\mathbf{y}$. Dividing this term by σ^2 we obtain

$$\mathbf{y}'\mathbf{A}\mathbf{y}$$

where

$$\mathbf{A} = \sigma^{-2}(\mathbf{I} - \mathbf{H}).$$

Note that $\mathbf{A}\mathbf{V} = (\mathbf{I} - \mathbf{H})$ is idempotent with rank $n - p$. Also note that

$$\lambda = \frac{1}{2}\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = \frac{1}{2}(\mathbf{X}\boldsymbol{\beta})'(\mathbf{I} - \mathbf{H})\mathbf{X}\boldsymbol{\beta} = 0.$$

Hence, by Theorem 13.8 it holds that

$$\mathbf{y}'(\mathbf{I} - \mathbf{H})\mathbf{y}/\sigma^2 \sim \chi^2(n - p).$$

Now consider the term $\mathbf{y}'\mathbf{H}\mathbf{y}$. Dividing this term by σ^2 we obtain

$$\mathbf{y}'\mathbf{A}\mathbf{y}$$

where

$$\mathbf{A} = \sigma^{-2}\mathbf{H}.$$

Note that $\mathbf{A}\mathbf{V} = \mathbf{H}$ is idempotent with rank p . Also note that

$$\lambda = \frac{1}{2}\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} = \frac{1}{2\sigma^2}(\mathbf{X}\boldsymbol{\beta})'\mathbf{H}\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\beta}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}/(2\sigma^2).$$

Hence,

$$\mathbf{y}'\mathbf{H}\mathbf{y} \sim \chi^2(p, \boldsymbol{\beta}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}/(2\sigma^2)).$$

13.3 Independence

Let's continue by discussing some independence results related to quadratic forms. In particular, we seek conditions for when $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{B}\mathbf{y}$ are independent and when $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ are independent.

Theorem: 13.10 Let $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \Sigma)$, \mathbf{A} be idempotent of rank m , and \mathbf{B} be idempotent of rank s , and suppose $\mathbf{B}\Sigma\mathbf{A} = \mathbf{0}$. Then $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{B}\mathbf{y}$ are independent.

Theorem: 13.11 Let $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \Sigma)$, \mathbf{A} be idempotent of rank m , and \mathbf{B} be idempotent of rank s , and suppose $\mathbf{B}\Sigma\mathbf{A} = \mathbf{0}$. Then $\mathbf{y}'\mathbf{A}\mathbf{y}$ and $\mathbf{y}'\mathbf{B}\mathbf{y}$ are independent.

Example: 13.12 Consider the linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$. In a previous example we showed that

$$\mathbf{y}'(\mathbf{I} - \mathbf{H})\mathbf{y}/\sigma^2 \sim \chi_{n-p}^2.$$

and

$$\mathbf{y}'\mathbf{H}\mathbf{y}/\sigma^2 \sim \chi_p^2(\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}/(2\sigma^2)).$$

Now let

$$\mathbf{y}'(\mathbf{I} - \mathbf{H})\mathbf{y}/\sigma^2 = \mathbf{y}'\mathbf{A}\mathbf{y}$$

where $\mathbf{A} = \sigma^{-2}(\mathbf{I} - \mathbf{H})$. In addition, let

$$\mathbf{y}'\mathbf{H}\mathbf{y}/\sigma^2 = \mathbf{y}'\mathbf{B}\mathbf{y}$$

where $\mathbf{B} = \sigma^{-2}\mathbf{H}$.

Hence, it holds that

$$\begin{aligned} \mathbf{B}\Sigma\mathbf{A} &= \sigma^{-2}\mathbf{H}\sigma^2\mathbf{I}\sigma^{-2}(\mathbf{I} - \mathbf{H}) \\ &= \mathbf{0} \end{aligned}$$

Thus, $\mathbf{y}'(\mathbf{I} - \mathbf{H})\mathbf{y}/\sigma^2$ and $\mathbf{y}'\mathbf{H}\mathbf{y}/\sigma^2$ are independent quadratic forms and it holds that

$$\begin{aligned} F &= \frac{\mathbf{y}'\mathbf{H}\mathbf{y}/p}{\mathbf{y}'(\mathbf{I} - \mathbf{H})\mathbf{y}/(n-p)} \\ &= \frac{\sigma^{-2}\mathbf{y}'\mathbf{H}\mathbf{y}/p}{\sigma^{-2}\mathbf{y}'(\mathbf{I} - \mathbf{H})\mathbf{y}/(n-p)} \\ &\sim F_{p, n-p}(\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}/2\sigma^2) \end{aligned}$$

13.4 Cochran's Theorem

Cochran's Theorem can be used to determine the distributions of partitioned sums of squares of normally distributed random variables. It allows us to split the sum of the squares of observations into a number of quadratic forms where each corresponds to some cause of variation. This result can be used to construct the ANOVA tables one typically encounters when fitting linear models.

Theorem: 13.13 Let $\mathbf{y} \sim N_p(\mu, \sigma^2 \mathbf{I})$ and suppose that $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k$ are $n \times n$ symmetric idempotent matrices with $\text{rank}(\mathbf{A}_i) = s_i$. If

$$\mathbf{A}_1 + \mathbf{A}_2 + \dots + \mathbf{A}_k = \mathbf{I},$$

then

$$\mathbf{y}'\mathbf{A}_1\mathbf{y}/\sigma^2, \mathbf{y}'\mathbf{A}_2\mathbf{y}/\sigma^2, \dots, \mathbf{y}'\mathbf{A}_k\mathbf{y}/\sigma^2$$

follow independent $\chi^2_{s_i}(\lambda_i)$ distributions, where $\lambda_i = \mu'\mathbf{A}_i\mu/2\sigma^2$, for $i = 1, 2, \dots, k$ and $\sum_{i=1}^k s_i = n$.