Econometrics I TA Session

Wang Xin

May 19, 2022

1 Idempotent and Symmetric Matrix

The matrix $\mathbf{A}$ is idempotent and symmetric if and only if $\mathbf{A}^2 = \mathbf{A} = \mathbf{A}'$.

1.1 $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$

From the previous note, we know that

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

Therefore, the predicted value of $\mathbf{Y}$ can be written as

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \tag{1}$$

Let $\mathbf{N} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, then the below equations can be found:

$$\mathbf{N}' = (\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{N} \tag{2}$$

$$\mathbf{N}\mathbf{N} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{N} \tag{3}$$

1.2 $\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$

The residuals are defined as $\mathbf{e} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$, so using equation (1), we have

$$\mathbf{e} = \mathbf{Y} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Y}$$

Let $\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, similar to N, we can find

$$\mathbf{M}' = (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')' = \mathbf{I}' - (\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')' = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{M} \tag{4}$$

$$\begin{aligned}\mathbf{M}\mathbf{M} &= (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}') \\ &= \mathbf{I} + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{2}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \\ &= \mathbf{I} + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{2}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \\ &= \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{M}\end{aligned} \tag{5}$$

1.3 Other Relationships

Both **N** an **M** are projection matrices.

A projection matrix **P** describes the influence each response value has on each fitted value, whenever **P** is applied twice to any vector, it gives the same result as if it were applied once.

According to the properties of projection matrix, the below relationships can be found:

$$\mathbf{NX} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{X} \tag{6}$$

$$\mathbf{MX} = (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{X} = \mathbf{X} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{X} - \mathbf{X} = \mathbf{0} \tag{7}$$

$$\mathbf{MN} = \mathbf{NM} = \mathbf{0} \tag{8}$$

2 $\mathcal{X}^2$ Distribution

2.1 Definition

At TA Session, gamma distribution was introduced last week.

$$g(x) = \begin{cases} \dfrac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, & x > 0 \\ 0, & elsewhere \end{cases}$$

Let us now consider a special case of the gamma distribution in which $\alpha = r/2$, where r is a positive integer, and $\beta = 2$. A random variable X of the continuous type that has the pdf

$$f(x) = \begin{cases} \dfrac{x^{r/2-1}e^{-x/2}}{2^{r/2}\Gamma(r/2)}, & x > 0 \\ 0, & elsewhere \end{cases} \tag{9}$$

and the mgf

$$M(\theta) = (1 - 2\theta)^{-r/2}, \qquad \theta < \frac{1}{2}$$

is said to have a chi-square distribution ($\mathcal{X}^2$-distribution), and any $f(x)$ of this form is called a chi-square pdf.

2.2 Operation on Normal Variable

Theorem 1. *If the random variable* $X \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, $\sigma^2 < 0$, *then the random variable* $V = (X - \mu)^2/\sigma^2$ *is* $\mathcal{X}^2(1)$.

Proof. Because $V = W^2$, where $W = (X - \mu)/\sigma$ is $N(0,1)$, the cdf $G(v)$ for V is, for $v \geq 0$,

$$G(v) = P(W^2 \leq v) = P(-\sqrt{v} \leq w \leq \sqrt{v})$$
$$= 2\int_0^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw, \quad v \geq 0$$

and

$$G(v) = 0, \quad v < 0$$

If we change the variable of integration by writing $w = \sqrt{y}$, then

$$G(v) = 2\int_0^{v} \frac{1}{\sqrt{2\pi}\sqrt{y}} e^{-y/2} dy, \quad v \geq 0$$

Hence the pdf $g(v) = G'(v)$ of the continuous-type random variable V is

$$g(v) = \begin{cases} \frac{1}{\sqrt{2\pi}} v^{1/2-1} e^{-v/2} & 0 < v < \infty \\ 0, \quad elsewhere \end{cases}$$

Since $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, thus V is $\mathcal{X}^2(1)$. □

2.3 F-distribution

Consider two independent chi-square random variables U and V having r_1 and r_2 degrees of freedom, respectively. The joint pdf $h(u,v)$ of U and V is then

$$h(u,v) = \begin{cases} \frac{u^{r_1/2-1} v^{r_2/2-1} e^{-(u+v)/2}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}}, & 0 < u, v < \infty \\ 0, \quad elsewhere \end{cases}$$

We define the new random variable

$$W = \frac{U/r_1}{V/r_2}$$

and we propose finding the pdf $g_1(w)$ of W. The equations

$$w = \frac{u/r_1}{v/r_2}, \quad z = v,$$

define a one-to-one transformation that maps the set $\mathcal{S} = (u,v) : 0 < u, v < \infty$ onto the set $T = (w,z) : 0 < w, z < \infty$. Since $u = (r_1/r_2)zw, v = z$, the absolute value of the Jacobian of the transformation is $|J| = (r_1/r_2)z$. The joint pdf $g(w,z)$ of the random variables W and $Z = V$ is then

$$g(w,z) = \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} (\frac{r_1 zw}{r_2})^{\frac{r_1-2}{2}} z^{\frac{r_2-2}{2}} exp\left[-\frac{z}{2}\left(\frac{r_1 w}{r_2} + 1\right)\right] \frac{r_1 z}{r_2} \quad (10)$$

provided that $(w, z) \in T$, and zero elsewhere. The marginal pdf $g_1(w)$ of W is the

$$g_1(w) = \int_{-\infty}^{+\infty} g(w, z)dz$$
$$= \int_0^{\infty} \frac{(r_1/r_2)^{r_1/2} w^{\frac{r_1}{2}-1}(r_2/r_1)}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} z^{\frac{r_1-2}{2}} z^{\frac{r_2-2}{2}} \frac{r_1 z}{r_2} exp\left[-\frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right)\right] dz$$
$$= \int_0^{\infty} \frac{(r_1/r_2)^{r_1/2} w^{\frac{r_1}{2}-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} z^{\frac{r_1+r_2}{2}-1} exp\left[-\frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right)\right] dz$$

If we change the variable of integration by writing $y = \frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right)$, then

$$\frac{dy}{dz} = \frac{1}{2}\left(\frac{r_1 w}{r_2}+1\right)$$

$$g_1(w) = \frac{(r_1/r_2)^{r_1/2} w^{\frac{r_1}{2}-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \int_0^{\infty} (\frac{2y}{r_1 w/r_2+1})^{\frac{r_1+r_2}{2}-1} e^{-y} \frac{2}{r_1 w/r_2+1} dy$$
$$= \frac{(r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\Gamma(r_2/2)} \frac{w^{\frac{r_1}{2}-1}}{(r_1 w/r_2+1)^{(r_1+r_2)/2}} \int_0^{\infty} y^{\frac{r_1+r_2}{2}-1} e^{-y} dy$$

In the above equation, a gamma function is contained

$$\Gamma(\frac{r_1+r_2}{2}) = \int_0^{\infty} y^{\frac{r_1+r_2}{2}-1} e^{-y} dy$$

Therefore, the pdf $g_1(w)$ of W can be simplified as

$$g_1(w) = \begin{cases} \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\Gamma(r_2/2)} \frac{w^{r_1/2-1}}{(1+r_1 w/r_2)^{(r_1+r_2)/2}}, & 0 < w < \infty \\ 0, & elsewhere \end{cases} \tag{11}$$

Accordingly, if U and V are independent chi-square variables with r_1 and r_2 degrees of freedom, respectively, then $W = (U/r_1)/(V/r_2)$ has the pdf $g_1(w)$. The distribution of this random variable is usually called an F-distribution; and we often call the ratio, which we have denoted by W, F. That is

$$F = \frac{U/r_1}{V/r_2} \tag{12}$$

It should be observed that an F-distribution is completely determined by the two parameters r_1 and r_2.