Economics 361
Joint Hypotheses Tests

Jun Ishii *
Department of Economics
Amherst College

Fall 2023

For simplicity, consider hypothesis tests involving the OLS estimator, with all of the Classical Normal Regression Model assumptions (Linearity, Spherical Errors, Full Rank, Normality). Thus,

$$b \mid X \quad \sim \quad N(\beta, \sigma^2 (X'X)^{-1})$$

This also means that

$$\begin{aligned} Rb - r \mid X \quad &\sim \quad N(R\beta - r, \sigma^2 R(X'X)^{-1}R') \\ \frac{Rb - r}{\sqrt{\sigma^2 R(X'X)^{-1}R'}} \mid X \quad &\sim \quad N(\frac{R\beta - r}{\sqrt{\sigma^2 R(X'X)^{-1}R'}}\ ,\ 1) \end{aligned}$$

where R is (1 x K) matrix of constants and r a single (1 x 1) constant.

Any single (simple) linear hypothesis involving β can be written as $R\beta - r = 0$. The test statistic $\frac{Rb-r}{\sqrt{\sigma^2 R(X'X)^{-1}R'}}$ can be used as we know its sampling distribution under the null hypothesis

$$\frac{Rb - r}{\sqrt{\sigma^2 R(X'X)^{-1}R'}} \mid X \quad \sim \quad N(0, 1) \quad \text{under } H_0 : R\beta - r = 0$$

Suppose we want to test the following *joint* hypotheses

$$\begin{aligned} H_0 : &\qquad \beta_1 + \beta_2 = 5 \text{ and } \beta_1 - \beta_2 = 3 \\ H_a : &\qquad \beta_1 + \beta_2 \neq 5 \text{ or } \beta_1 - \beta_2 \neq 3 \end{aligned}$$

for an OLS model where $E[Y|X] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 = X\beta$

The hypotheses can be re-written as

$$\begin{aligned} \beta_1 + \beta_2 - 5 = 0 \quad &\Rightarrow \quad \underbrace{\begin{pmatrix} 0 & 1 & 1 \end{pmatrix}}_{=R_1} \beta - \underbrace{5}_{=r_1} = 0 \\ \beta_1 - \beta_2 - 3 = 0 \quad &\Rightarrow \quad \underbrace{\begin{pmatrix} 0 & 1 & -1 \end{pmatrix}}_{=R_2} \beta - \underbrace{3}_{=r_2} = 0 \end{aligned}$$

*Office: Converse Hall 315 Phone: (413) 542-2901 E-mail: jishii@amherst.edu

If we were testing each hypothesis individually, we would be using the following test statistics

$$\text{Testing } H_0 : \beta_1 + \beta_2 = 5 \qquad \text{Test Statistic} = \frac{R_1 b - r_1}{\sqrt{\sigma^2 R_1 (X'X)^{-1} R_1'}}$$
$$\text{Testing } H_0 : \beta_1 - \beta_2 = 3 \qquad \text{Test Statistic} = \frac{R_2 b - r_2}{\sqrt{\sigma^2 R_2 (X'X)^{-1} R_2'}}$$

One might be tempted to conduct the joint hypotheses using the above two test statistic as follows:

- Test each hypothesis individually
- If any of the hypotheses is individually rejected, then reject the joint hypotheses
- If none of the hypotheses is individually rejected, then fail to reject the joint hypotheses

But for the above procedure, how does one pick the appropriate critical values for each of the individual hypothesis tests? Suppose we want to test the joint hypotheses with a significance level of 5% (probability of type I error fixed at 5%). We might be tempted to test each individual hypothesis using a critical value corresponding to a 5% significance level. But this leads to a joint hypotheses test with a significance level **greater** than 5%.

While each hypothesis may be rejected (when true) with 5% probability, the joint hypotheses (according to the procedure above) is rejected when any of the hypotheses is individually rejected (when true); the significance level for the joint hypotheses will end up being between 5% and 10% (5% only if the two hypothesis tests always lead to the same conclusion and 10% when they never lead to the same conclusion)

So, theoretically, we would need to pick critical values for the two test statistic (C_1, C_2) such that

$$\Pr\left(\frac{R_1 b - r_1}{\sqrt{\sigma^2 R_1 (X'X)^{-1} R_1'}} \in C_1 \text{ or } \frac{R_2 b - r_2}{\sqrt{\sigma^2 R_2 (X'X)^{-1} R_2'}} \in C_2 \mid \text{when } R_1\beta = r_1 \text{ and } R_2\beta = r_2\right) = 0.05$$

This gets complicated. Keep in mind that not only do we want to pick critical values C_1 and C_2 that satisfy the above, we also want to pick C_1 and C_2 that minimize the probability of Type II error for the chosen significance level.

So let us consider another test statistic, related to the ones above. Note that we can write the two hypotheses jointly using some matrix notation

$$\begin{pmatrix} \beta_1 + \beta_2 - 5 = 0 \\ \beta_1 - \beta_2 - 3 = 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} - \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \beta - \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

Replace β with b

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} - \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \underbrace{\begin{pmatrix} R_1 \\ R_2 \end{pmatrix}}_{=R} b - \underbrace{\begin{pmatrix} r_1 \\ r_2 \end{pmatrix}}_{=r}$$

Note that the joint hypotheses can also be expressed as $R\beta - r = 0$, the differences being that R is now (M x k) and r now (M x 1), where M is the number of hypotheses jointly being tested; for the above example $M = 2$.

Additionally, we know that as each test statistic is distributed Normal. For this particular example, the vector of the two test statistic will also be jointly distributed Normal[1]

$$Rb - r \quad \sim \quad N\left(\begin{pmatrix} R_1\beta - r_1 \\ R_2\beta - r_2 \end{pmatrix} \; , \; \underbrace{\sigma^2 R(X'X)^{-1}R'}_{(MxM)} \right)$$

Note that

$$\sigma^2 R(X'X)^{-1}R' \quad = \quad \begin{pmatrix} \mathrm{Var}(R_1 b - r_1 \mid X) & \mathrm{Cov}(R_1 b - r_1, R_2 b - r_2 \mid X) \\ \mathrm{Cov}(R_1 b - r_1, R_2 b - r_2 \mid X) & \mathrm{Var}(R_2 b - r_2 \mid X) \end{pmatrix}$$

To see the above explicitly,

$$\begin{aligned}
\sigma^2 R(X'X)^{-1}R' &= R\sigma^2(X'X)^{-1}R' \\
&= \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \sigma^2(X'X)^{-1} \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \mathrm{Var}(b_0|X) & \mathrm{Cov}(b_0, b_1|X) & \mathrm{Cov}(b_0, b_2|X) \\ \mathrm{Cov}(b_0, b_1|X) & \mathrm{Var}(b_1|X) & \mathrm{Cov}(b_1, b_2|X) \\ \mathrm{Cov}(b_0, b_2|X) & \mathrm{Cov}(b_1, b_2|X) & \mathrm{Var}(b_2|X) \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \\
&= \begin{pmatrix} \underbrace{\mathrm{Var}(b_1|X) + \mathrm{Var}(b_2|X) + 2\mathrm{Cov}(b_1, b_2|X)}_{=\mathrm{Var}(R_1 b - r_1|X)} & \underbrace{\mathrm{Var}(b_1|X) - \mathrm{Var}(b_2|X)}_{=\mathrm{Cov}(R_1 b - r_1, R_2 b - r_2|X)} \\ \mathrm{Var}(b_1|X) - \mathrm{Var}(b_2|X) & \underbrace{\mathrm{Var}(b_1|X) + \mathrm{Var}(b_2|X) - 2\mathrm{Cov}(b_1, b_2|X)}_{=\mathrm{Var}(R_2 b - r_2|X)} \end{pmatrix}
\end{aligned}$$

Under the null joint hypotheses, we know that

$$Rb - r \quad \sim \quad N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \; , \; \underbrace{\sigma^2 R(X'X)^{-1}R'}_{(MxM)} \right) \qquad \text{under } H_0$$

Using the above result, we can additionally show that

$$\underbrace{(Rb - r)' \, [R\sigma^2(X'X)^{-1}R']^{-1} \, (Rb - r)}_{\chi^2 \text{ stat}} \mid X \;\sim\; \chi^2_2 \qquad \text{under } H_0$$

Without going into too much detail, the above result comes from recognizing that the test statistic is simply the sum of squared standard Normal random variable. As shown earlier, the square of a standard Normal is distributed χ^2_1. The sum of M random variables independently distributed χ^2_1 can be shown to be distributed χ^2_M (change of variables again).

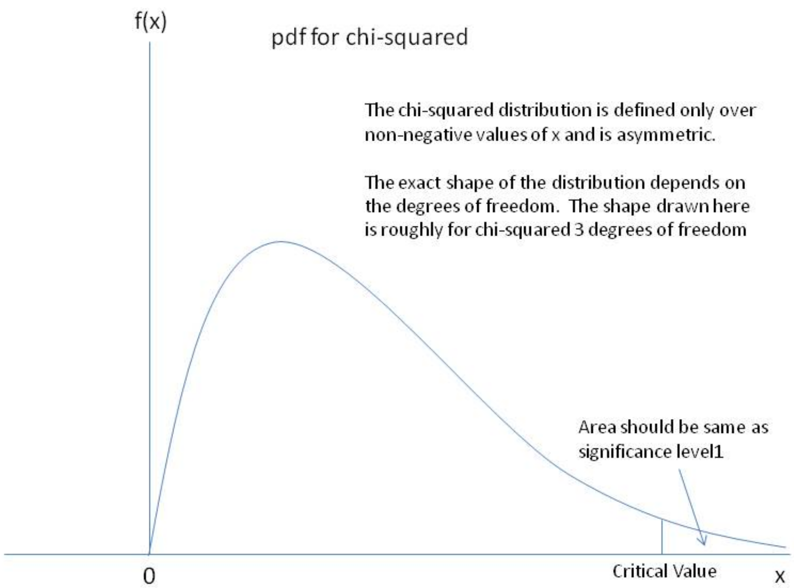
[1]This result stems from the fact that both test statistics are linear combinations of b

This χ^2 test statistic is much easier to use. The critical value can be calculated in a manner similar to the one-sided single null hypothesis tests. The χ^2 random variable only takes non-negative values and is not symmetric. The appropriate critical values are those values on the right tail (large positive values) that satisfy the chosen significance level.

So for a 5% significance level and $M = 2$, we reject if

$$(Rb - r)' \, [R\sigma^2(X'X)^{-1}R']^{-1} \, (Rb - r) > 5.99$$

The value 5.99 comes from Table A.2 in the Goldberger text (pp. 382-383)



The joint hypotheses test above generalizes to

- $M > 2$ joint linear hypotheses

  R is (M x k); r is (M x 1); test statistic is distributed χ^2_M

- When σ^2 is unknown

  use s^2 instead of σ^2 and divide the whole term by M (# of joint hypotheses)

  test statistic is now distributed $F_{M,N-k}$ rather than χ^2_M

- Hypotheses tests involving GLS instead of OLS

  use b^{gls} instead of b; use $(X'\Sigma^{-1}X)^{-1}$ instead of $\sigma^2(X'X)^{-1}$

See "Hypothesis Test Cheat Sheet" for more details