EPI809

Testing in Multiple Regression Models

- Long & Short regressions, and various ways of formulating H_0 & Ha

- Type-I error

- Single-test p-values & type-I error rate

- Multiple testing

- Sequential testing:
    - F-test, if significant,
    - Examine t-test for individual predictors

Various ways of describing H0 and Ha

$$\begin{cases} Model\text{:}\ \ y_i = \mu + x_{1i}\beta_1 + x_{2i}\beta_2 + x_{3i}\beta_3 + x_{4i}\beta_4 + \varepsilon_i \\ H_0\text{:}\,\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \\ H_\text{a}\text{:}\ \ \text{At least one of the}\ \ \beta's \neq 0 \end{cases}$$

$$\begin{cases} H_0\text{:}\ y_i = \mu + \varepsilon_i \\ H_\text{a}\text{:}\ y_i = \mu + x_{1i}\beta_1 + x_{2i}\beta_2 + x_{3i}\beta_3 + x_{4i}\beta_4 + \varepsilon_i \end{cases}$$

$$\begin{cases} H_0\text{:}\,\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \\ H_\text{a}\text{:}\ y_i = \mu + x_{1i}\beta_1 + x_{2i}\beta_2 + x_{3i}\beta_3 + x_{4i}\beta_4 + \varepsilon_i \end{cases}$$

F-test Vs t-test

- The preceding example is a 3 DF test.

- For 1-DF tests, the F-test and the t-test are equivalent:

    - The p-values are identical
    - the t-stat=sqrt(F-stat)
    (verify this with an example)

- However, for test involving more than 1 DF the two tests behave different.

- How do we proceed?

    - For more than 1DF tests we use the F-test
    - If the F–test leads to rejection, then, at least one of the individual effects is non-zero, and only then we evaluate individual effects using the t-test.

- Why don't we just focus on the t-test?

- The problem is that when we carry multiple test, rejection based on individual p-values do not control type-I error rate accurately. This problem emerges from multiple-testing/

Hypothesis Testing & Type-I Error

State of nature	Do not reject H_0	Reject H_0
H_0 holds	✓	Type-I error
H_a holds	Type-II error	✓

- We use p-values to control type-I error rate (i.e., the probability of rejecting H0 when H0 holds).
- P-values are estimates of type-I error rate.
- We reject when p-values are small to keep the probability of type-I error low.
- (If the p-value is correct) and we conduct 1 test p(type-I error)=pvalue
- What if we to two tests? What is the probability of making at least one mistake?

Hypothesis Testing & Type-I Error

$$\begin{cases} H_0\colon\ y_i = \mu + \varepsilon_i \\ H_a\colon\ y_i = \mu + x_{1i}\beta_1 + x_{2i}\beta_2 + \varepsilon_i \end{cases}$$

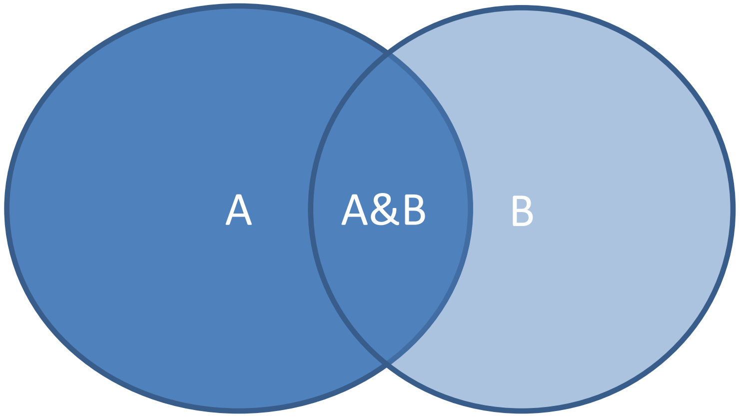
$pvalue_1 = P(rejecting\ H_{01}\colon \beta_1$=0| β_1 =0, μ and β_2 free)

$pvalue_2 = P(rejecting\ H_{02}\colon \beta_2$=0| β_2 =0, μ and 1 free)

What is the probability of making at least one mistake (FWER=Family-wise error rate)?

$$P[(reject\ H_{01}|H_{01},H_{02})\ \ or\ \ (reject\ H_{02}|H_{01},H_{02})\ \ or\ \ (reject\ H_{01}\&H_{02}|H_{01},H_{02})]$$

$p(A\ or\ B)$= $p(A) + p(B)$- $p(A\&B)$



Thus, if we reject t-tests at an **α**-level (significance, say 0.05), and we conduct multiple tests, the probability of making at leas 1 type-I error can be larger than **α** (up to **α** times the number of tests)

To address this problem, if our test involves more than 1 DF, we use the F-test

Post-hoc analysis. What if the F-test is rejected?

- If the F-test is rejected, we can then proceed and evaluate individual t-tests, but we still need to account for multiple testing.

- For factors, when comparing the effects of various levels, two common approaches are Duncan's test and Tukey's Honest Significant Difference (HSD)

- Both tests are implemented in the package Agricolae (See examples)