Multiple Testing

Familywise Error Rate (FWER)

The simultaneous interval estimation methods we learned about correspond to simultaneous testing procedures that control the familywise error rate (FWER).

The familywise error rate (FWER) is the probability of one or more Type I errors when conducting a family of tests.

For example, suppose

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}),$$

and we wish to test

$$H_{0j} : \boldsymbol{c}_j' \boldsymbol{\beta} = 0 \quad \text{for} \quad j = 1, \ldots, m,$$

where each H_{0j} is a testable null hypothesis.

If for each $j = 1, \ldots, m$, we reject $H_{0j} \Longleftrightarrow 0 \notin I_j$ where

$$I_j = \left[\boldsymbol{c}_j' \hat{\boldsymbol{\beta}} - t_{n-r, \frac{\alpha}{2m}} \sqrt{\hat{\sigma}^2 \boldsymbol{c}_j' (\boldsymbol{X}'\boldsymbol{X})^- \boldsymbol{c}_j}, \boldsymbol{c}_j' \hat{\boldsymbol{\beta}} + t_{n-r, \frac{\alpha}{2m}} \sqrt{\hat{\sigma}^2 \boldsymbol{c}_j' (\boldsymbol{X}'\boldsymbol{X})^- \boldsymbol{c}_j} \right],$$

then the FWER will be bounded above by α.

Prove that this is true.

As another example, suppose

$$y_{ij} = \mu_i + \varepsilon_{ij}, \quad i = 1, \ldots, t;\; j = 1, \ldots, n,$$

where $\varepsilon_{11}, \varepsilon_{12}, \ldots, \varepsilon_{tn} \overset{i.i.d.}{\sim} N(0, \sigma^2)$.

Consider testing the family of null hypotheses

$$H_0^{(i,i^*)} : \mu_i = \mu_{i^*}, \quad 1 \leq i < i^* \leq t.$$

Suppose for each $i < i^*$ pair, we reject $H_0^{(i,i^*)}$ iff

$$0 \notin \left[\bar{y}_{i\cdot} - \bar{y}_{i^*\cdot} - \frac{\hat{\sigma}}{\sqrt{n}} R_{t,t(n-1),\alpha}, \bar{y}_{i\cdot} - \bar{y}_{i^*\cdot} + \frac{\hat{\sigma}}{\sqrt{n}} R_{t,t(n-1),\alpha}\right].$$

Then FWER = α.

Strong Control of FWER

The methods we learned about (Bonferroni, Scheffé, Tukey) correspond to multiple testing procedures that provide strong control of the FWER.

A method for testing a family of null hypotheses $H_{01}, \ldots, H_{0m}$ provides strong control of FWER at level α iff FWER $\leq \alpha$ regardless of which or how many nulls in the family are true.

Weak Control of FWER

In contrast, a method provides weak control of FWER at level α if $\text{FWER} \leq \alpha$ whenever all null hypotheses in the family $(H_{01}, \ldots, H_{0m})$ are true.

Strong control of FWER $\Rightarrow$ weak control of FWER.

Show by example that weak control $\not\Rightarrow$ strong control.

- Many multiple testing procedures that provide only weak control of the FWER have been published in the statistical literature.
- Methods that provide weak control of the FWER tend to be more powerful than methods that provide strong control, but weak control is rarely sufficient.
- Thus, methods that provide strong control of FWER are preferred.