Multiple comparisons in ANCOVA

Stat 526
4-29-25

(1)

Suppose that $H_0: \tau_i = 0$ is rejected.

Need to know the properties of $\hat{\tau}_i - \hat{\tau}_j$

$$y_{ij} = \mu + \tau_i + \beta(x_{ij} - \bar{x}_{..}) + \varepsilon_{ij}$$

$$\bar{y}_{i.} = \mu + \tau_i + \beta(\bar{x}_{i.} - \bar{x}_{..}) + \bar{\varepsilon}_{i.}$$

$$\bar{y}_{..} = \mu + \bar{\varepsilon}_{..}$$

$$\hat{\beta} = \frac{E_{xy}}{E_{xx}} = \frac{1}{E_{xx}}\left[\sum_i\sum_j (x_{ij} - \bar{x}_{i.})(y_{ij} - \bar{y}_{i.})\right]$$

(2)

$$\hat{\beta} = \frac{1}{E_{xx}}\left[\sum_i\sum_j (x_{ij} - \bar{x}_{i.})\left(\beta(x_{ij} - \bar{x}_{i.}) + \varepsilon_{ij} - \bar{\varepsilon}_{i.}\right)\right]$$

$$= \frac{1}{E_{xx}}\left[\beta E_{xx} + \underbrace{\sum_i\sum_j (x_{ij} - \bar{x}_{i.})(\varepsilon_{ij} - \bar{\varepsilon}_{i.})}_{\sum_i\sum_j (x_{ij} - \bar{x}_{i.})\varepsilon_{ij} - \sum_i\left[\bar{\varepsilon}_{i.}\underbrace{\sum_j (x_{ij} - \bar{x}_{i.})}_{0}\right]}\right]$$

$$E(\hat{\beta}) = \beta + 0 \quad \therefore \hat{\beta} \text{ is unbiased}$$

$$V(\hat{\beta}) = \frac{1}{E_{xx}^2}\sum_i\sum_j (x_{ij} - \bar{x}_{i.})^2\sigma^2 = \frac{\sigma^2}{E_{xx}}$$

3

Let $\mu_i = \mu + \tau_i$

So $\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i = \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..})$

$$= \mu + \tau_i + \beta(\bar{x}_{i.} - \bar{x}_{..}) + \bar{\varepsilon}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..})$$

$$= \mu + \tau_i + (\beta - \hat{\beta})(\bar{x}_{i.} - \bar{x}_{..}) + \bar{\varepsilon}_{i.}$$

$E[\hat{\mu}_i] = \mu + \tau_i = \mu_i \qquad \therefore \hat{\mu}_i$ is unbiased

$$V[\hat{\mu}_i] = (\bar{x}_{i.} - \bar{x}_{..})^2 V(\hat{\beta}) + V(\bar{\varepsilon}_{i.})$$
$$+ 2\,\mathrm{Cov}[(\beta - \hat{\beta})(\bar{x}_{i.} - \bar{x}_{..}), \bar{\varepsilon}_{i.}]$$

4

$$V[\hat{\mu}_i] = (\bar{x}_{i.} - \bar{x}_{..})^2 \frac{\sigma^2}{E_{xx}} + \frac{\sigma^2}{n}$$
$$-2(\bar{x}_{i.} - \bar{x}_{..})\underbrace{\mathrm{Cov}(\hat{\beta}, \bar{\varepsilon}_{i.})}_{*}$$

$$* \quad \mathrm{Cov}(\hat{\beta}, \bar{\varepsilon}_{i.}) = \mathrm{Cov}\left(\beta + \frac{1}{E_{xx}}\sum_i\sum_j (x_{ij} - \bar{x}_{i.})\varepsilon_{ij}, \bar{\varepsilon}_{i.}\right)$$

$$= \frac{1}{E_{xx}}\sum_i\sum_j (x_{ij} - \bar{x}_{i.})\underbrace{\mathrm{Cov}(\varepsilon_{ij}, \bar{\varepsilon}_{i.})}_{\underbrace{\mathrm{Cov}\left(\varepsilon_{ij}, \frac{1}{n}\sum_j \varepsilon_{ij}\right)}_{\sigma^2/n}}$$

(5)

$$= \frac{\sigma^2}{n} \frac{1}{E_{xx}} \underbrace{\sum_i \sum_j (x_{ij} - \bar{x}_{i\cdot})}_{0} = 0$$

$$\therefore V[\hat{\mu}_i] = \sigma^2 \left[\frac{1}{n} + \frac{(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})^2}{E_{xx}} \right]$$

Finally, consider $\hat{\mu}_i - \hat{\mu}_j$ (same as $\hat{\tau}_i - \hat{\tau}_j$)

$$E[\hat{\mu}_i - \hat{\mu}_j] = \mu_i - \mu_j = (\mu + \tau_i) - (\mu + \tau_j)$$
$$= \tau_i - \tau_j$$

(6)

$$V(\hat{\mu}_i - \hat{\mu}_j) = V(\hat{\mu}_i) + V(\hat{\mu}_j) - 2\underbrace{\mathrm{Cov}(\hat{\mu}_i, \hat{\mu}_j)}_{**}$$

$$** = \mathrm{Cov}\big(\mu + \tau_i + (\beta - \hat{\beta})(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot}) + \bar{\varepsilon}_{i\cdot},$$
$$\mu + \tau_j + (\beta - \hat{\beta})(\bar{x}_{j\cdot} - \bar{x}_{\cdot\cdot}) + \bar{\varepsilon}_{j\cdot}\big)$$
$$= (\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})(\bar{x}_{j\cdot} - \bar{x}_{\cdot\cdot}) \mathrm{Cov}(\beta - \hat{\beta}, \beta - \hat{\beta})$$
$$+ (\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot}) \cancel{\mathrm{Cov}(\beta - \hat{\beta}, \bar{\varepsilon}_{j\cdot})}^{0} + (\bar{x}_{j\cdot} - \bar{x}_{\cdot\cdot}) \cancel{\mathrm{Cov}(\bar{\varepsilon}_{i\cdot}, \beta - \hat{\beta})}^{0}$$
$$+ \cancel{\mathrm{Cov}(\bar{\varepsilon}_{i\cdot}, \bar{\varepsilon}_{j\cdot})}^{0}$$

$$= (\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})(\bar{x}_{j\cdot} - \bar{x}_{\cdot\cdot}) V(\hat{\beta})$$

$$= (\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})(\bar{x}_{j\cdot} - \bar{x}_{\cdot\cdot}) \frac{\sigma^2}{E_{xx}}$$

So $V[\hat{\mu}_i - \hat{\mu}_j] =$

$$\sigma^2\left[\frac{1}{n} + \frac{(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})^2}{E_{xx}}\right] + \sigma^2\left[\frac{1}{n} + \frac{(\bar{x}_{j\cdot} - \bar{x}_{\cdot\cdot})^2}{E_{xx}}\right]$$

$$- 2(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})(\bar{x}_{j\cdot} - \bar{x}_{\cdot\cdot}) \frac{\sigma^2}{E_{xx}}$$

8

$$= \sigma^2\left[\frac{2}{n} + \frac{(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})^2 + (\bar{x}_{j\cdot} - \bar{x}_{\cdot\cdot})^2 - 2(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})(\bar{x}_{j\cdot} - \bar{x}_{\cdot\cdot})}{E_{xx}}\right]$$

$$= \sigma^2\left[\frac{2}{n} + \frac{(\bar{x}_{i\cdot} - \bar{x}_{j\cdot})^2}{E_{xx}}\right]$$

$$\therefore \quad \frac{\hat{\mu}_i - \hat{\mu}_j - (\tau_i - \tau_j)}{\sqrt{\sigma^2\left[\frac{2}{n} + \frac{(\bar{x}_{i\cdot} - \bar{x}_{j\cdot})^2}{E_{xx}}\right]}} \sim N(0,1)$$

⑨

$$\frac{\hat{\mu}_i - \hat{\mu}_j - (\tau_i - \tau_j)}{\sqrt{MSE\left[\frac{2}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{j.})^2}{E_{xx}}\right]}} \approx t_{df_E}$$

$\therefore$ a confidence interval for $\tau_i - \tau_j$ is

$$\hat{\mu}_i - \hat{\mu}_j \pm t_{df_E}\sqrt{MSE\left[\frac{2}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{j.})^2}{E_{xx}}\right]}$$

⑩

Repeated measures

factor (fixed)

person (random)

$$y_{ijk} = \mu + \tau_i + \beta_j + \tau\beta_{ij} + \varepsilon_{(ij)k}$$

$i = 1, \ldots, a \quad j = 1, \ldots, n \quad k = 1$

This is just a special case of the RCBD

It is also a generalization of matched pairs

We already know how to handle this if the factor is fixed.

What if the factor is random?

df		R i	R j	R k	EMS	denom
$a-1$	τ_i	1	n	1	$n\sigma^2_\tau + \sigma^2_{\tau\beta} + \sigma^2$	MS_{AB}
$n-1$	β_j	a	1	1	$a\sigma^2_\beta + \sigma^2_{\tau\beta} + \sigma^2$	MS_{AB}
~~$(a-1)(n-1)$~~	~~$\tau\beta_{ij}$~~	~~1~~	~~1~~	~~1~~	$\cancel{\sigma^2_{\tau\beta} + \sigma^2}$	~~MS_E~~
0	$\varepsilon_{(ij)k}$	1	1	1	σ^2	