Scheffé's Method

Scheffé's Method:

Suppose

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where

$$\boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I}).$$

Let

$$\boldsymbol{c}_1'\boldsymbol{\beta}, \ldots, \boldsymbol{c}_q'\boldsymbol{\beta}$$

be q estimable functions, where

$$\boldsymbol{c}_1, \ldots, \boldsymbol{c}_q$$

are linearly independent.

Let

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{c}_1' \\ \vdots \\ \boldsymbol{c}_q' \end{bmatrix}$$

and define

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_q \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1'\boldsymbol{\beta} \\ \vdots \\ \boldsymbol{c}_q'\boldsymbol{\beta} \end{bmatrix} = \boldsymbol{C\beta}.$$

Let $\boldsymbol{W} = \boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{C}'$ with diagonal elements denoted

$$w_{11}, \ldots, w_{qq}.$$

Then

$$\mathrm{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}}) = \sigma^2 \boldsymbol{W}$$

and

$$\mathrm{Var}(\hat{\theta}_j) = \mathrm{Var}(\boldsymbol{c}_j'\hat{\boldsymbol{\beta}}) = \sigma^2 w_{jj} \quad j = 1, \ldots, q.$$

For any $q \times 1$ vector $\boldsymbol{u}$ and any $k \in \mathbb{R}$, let $L(\boldsymbol{u}, k)$ denote the interval

$$[\boldsymbol{u}'\hat{\boldsymbol{\theta}} - k\sqrt{\hat{\sigma}^2\boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}}, \boldsymbol{u}'\hat{\boldsymbol{\theta}} + k\sqrt{\hat{\sigma}^2\boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}}].$$

We want to find $k \ni$

$$\mathbb{P}(\boldsymbol{u}'\boldsymbol{\theta} \in L(\boldsymbol{u}, k) \quad \forall\, \boldsymbol{u} \in \mathbb{R}^q) = 1 - \alpha.$$

Thus, we seek simultaneously coverage probability $1 - \alpha$ for an infinite set of intervals.

Show that

$$\boldsymbol{u}'\boldsymbol{\theta} \in L(\boldsymbol{u}, k) \quad \forall\, \boldsymbol{u} \in \mathbb{R}^q$$

$$\iff$$

$$\frac{(\boldsymbol{C}\hat{\boldsymbol{\beta}} - \boldsymbol{C}\boldsymbol{\beta})'(\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{C}')^{-1}(\boldsymbol{C}\hat{\boldsymbol{\beta}} - \boldsymbol{C}\boldsymbol{\beta})}{q\hat{\sigma}^2} \leq \frac{k^2}{q}.$$

First, note that

$$
\begin{aligned}
&\boldsymbol{u}'\boldsymbol{\theta} \in L(\boldsymbol{u}, k) \quad \forall\, \boldsymbol{u} \in \mathbb{R}^q \\
\iff &\boldsymbol{u}'\boldsymbol{\theta} \in L(\boldsymbol{u}, k) \quad \forall\, \boldsymbol{u} \in \mathbb{R}^q \setminus \{\mathbf{0}\} \\
\because\, &\mathbf{0}'\boldsymbol{\theta} = 0 \in [0, 0] = L(\mathbf{0}, k).
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{u}'\boldsymbol{\theta} \in L(\boldsymbol{u}, k) \quad &\iff \quad \boldsymbol{u}'\hat{\boldsymbol{\theta}} - k\sqrt{\hat{\sigma}^2\boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}} \leq \boldsymbol{u}'\boldsymbol{\theta} \leq \boldsymbol{u}'\hat{\boldsymbol{\theta}} + k\sqrt{\hat{\sigma}^2\boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}} \\
&\iff \quad -k\sqrt{\hat{\sigma}^2\boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}} \leq \boldsymbol{u}'\boldsymbol{\theta} - \boldsymbol{u}'\hat{\boldsymbol{\theta}} \leq k\sqrt{\hat{\sigma}^2\boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}} \\
&\iff \quad -k \leq \frac{\boldsymbol{u}'\boldsymbol{\theta} - \boldsymbol{u}'\hat{\boldsymbol{\theta}}}{\sqrt{\hat{\sigma}^2\boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}}} \leq k \\
&\iff \quad \left| \frac{\boldsymbol{u}'\boldsymbol{\theta} - \boldsymbol{u}'\hat{\boldsymbol{\theta}}}{\sqrt{\hat{\sigma}^2\boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}}} \right| \leq k \\
&\iff \quad \frac{[\boldsymbol{u}'(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})]^2}{\hat{\sigma}^2\boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}} \leq k^2.
\end{aligned}
$$

$$
\begin{aligned}
&\therefore \boldsymbol{u}'\boldsymbol{\theta} \in L(\boldsymbol{u}, k) \quad \forall\, \boldsymbol{u} \in \mathbb{R}^q \setminus \{\boldsymbol{0}\} \\
&\iff \frac{[\boldsymbol{u}'(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})]^2}{\hat{\sigma}^2 \boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}} \leq k^2 \quad \forall\, \boldsymbol{u} \in \mathbb{R}^q \setminus \{\boldsymbol{0}\} \\
&\iff \max_{\boldsymbol{u} \neq \boldsymbol{0}} \frac{[\boldsymbol{u}'(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})]^2}{\hat{\sigma}^2 \boldsymbol{u}'\boldsymbol{W}\boldsymbol{u}} \leq k^2 \\
&\iff \frac{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})'\boldsymbol{W}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})}{\hat{\sigma}^2} \leq k^2
\end{aligned}
$$

(by C-S generalization in previous notes)

$$\iff \frac{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})' \boldsymbol{W}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})}{q\hat{\sigma}^2} \leq \frac{k^2}{q}$$

$$\iff \frac{(\boldsymbol{C}\hat{\boldsymbol{\beta}} - \boldsymbol{C}\boldsymbol{\beta})' (\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{C}')^{-1} (\boldsymbol{C}\hat{\boldsymbol{\beta}} - \boldsymbol{C}\boldsymbol{\beta})}{q\hat{\sigma}^2} \leq \frac{k^2}{q}.$$

Thus, we have

$$\begin{aligned}
&\mathbb{P}(\boldsymbol{u}'\boldsymbol{\theta} \in L(\boldsymbol{u}, k) \quad \forall\, \boldsymbol{u} \in \mathbb{R}^q) \\
&= \mathbb{P}\left[\frac{(\boldsymbol{C}\hat{\boldsymbol{\beta}} - \boldsymbol{C}\boldsymbol{\beta})'(\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{C}')^{-1}(\boldsymbol{C}\hat{\boldsymbol{\beta}} - \boldsymbol{C}\boldsymbol{\beta})}{q\hat{\sigma}^2} \leq \frac{k^2}{q}\right].
\end{aligned}$$

What shall we choose for k to make this probability equal to $1 - \alpha$?

$$\frac{(\boldsymbol{C}\hat{\boldsymbol{\beta}} - \boldsymbol{C}\boldsymbol{\beta})'(\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{C}')^{-1}(\boldsymbol{C}\hat{\boldsymbol{\beta}} - \boldsymbol{C}\boldsymbol{\beta})}{q\hat{\sigma}^2} \sim F_{q,n-r}.$$

Thus, if

$$k = \sqrt{qF_{q,n-r,\alpha}},$$

then

$$\frac{k^2}{q} = F_{q,n-r,\alpha}$$

so that the simultaneous coverage probability is $1 - \alpha$.

Example:

Suppose an experiment was conducted using a completely randomized design with 10 subjects in each of 4 treatment groups.

The treatment groups were defined by the combinations of levels from 2 factors: diet (1 or 2) and exercise program (1 or 2).

The model

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$$

was fit the response data

$$\begin{aligned} y_{ijk} \quad = \quad & \text{measure of overall health} \\ & \text{for diet } i\text{, exercise program } j\text{,} \\ & \text{subject } k \; (i = 1, 2;\; j = 1, 2;\; k = 1, \ldots, 10). \end{aligned}$$

The parameter μ_{ij} represents the mean response for diet i, exercise program j $(i = 1, 2;\; j = 1, 2)$.

The ε_{ijk} terms are assumed to be iid $N(0, \sigma^2)$.

A summary of the data is as follows:

$$\bar{y}_{11\cdot} = 9 \quad \bar{y}_{12\cdot} = 7$$
$$\bar{y}_{21\cdot} = 8 \quad \bar{y}_{22\cdot} = 3$$
$$\hat{\sigma}^2 = 5.$$

Suppose we want to construct a set of confidence intervals using a method that gives simultaneous coverage probability at least 95%.

Suppose the confidence intervals will be used to address the following questions:

1. Diet main effect?

2. Exercise program main effect?

3. Diet-by-exercise program interaction?

4. Difference between diet 1 and diet 2 under exercise program 1?

5. Difference between exercise program 1 and 2 under diet 1?

6. Diet 1, exercise program 1 vs. mean of other treatments?

What estimable function of

$$\boldsymbol{\beta} = \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{21} \\ \mu_{22} \end{bmatrix}$$

is of interest in each of the questions 1 through 6, respectively?

1. $\frac{\mu_{11}+\mu_{12}}{2} - \frac{\mu_{21}+\mu_{22}}{2} = \frac{\mu_{11}+\mu_{12}-\mu_{21}-\mu_{22}}{2}$

2. $\frac{\mu_{11}+\mu_{21}}{2} - \frac{\mu_{12}+\mu_{22}}{2} = \frac{\mu_{11}-\mu_{12}+\mu_{21}-\mu_{22}}{2}$

3. $(\mu_{11} - \mu_{12}) - (\mu_{21} - \mu_{22}) = \mu_{11} - \mu_{12} - \mu_{21} + \mu_{22}$

4. $\mu_{11} - \mu_{21}$

5. $\mu_{11} - \mu_{12}$

6. $\mu_{11} - \frac{(\mu_{12}+\mu_{21}+\mu_{22})}{3} = \frac{3\mu_{11}-\mu_{12}-\mu_{21}-\mu_{22}}{3}$.

Compute an estimate and standard error for each estimable function of interest.

$$\boldsymbol{X} = \underset{4\times 4}{\boldsymbol{I}} \otimes \underset{10\times 1}{\boldsymbol{1}}, \qquad \boldsymbol{X}'\boldsymbol{X} = 10 \underset{4\times 4}{\boldsymbol{I}}$$

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y} = \begin{bmatrix} \bar{y}_{11\cdot} \\ \bar{y}_{12\cdot} \\ \bar{y}_{21\cdot} \\ \bar{y}_{22\cdot} \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ 8 \\ 3 \end{bmatrix}$$

$$\text{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) = \sigma^2 \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{c} = \frac{\sigma^2}{10}\boldsymbol{c}'\boldsymbol{c}$$

$$se(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) = \sqrt{\frac{\hat{\sigma}^2}{10}\boldsymbol{c}'\boldsymbol{c}} = \sqrt{\frac{5}{10}\boldsymbol{c}'\boldsymbol{c}} = \sqrt{\boldsymbol{c}'\boldsymbol{c}/2}.$$

1. $\boldsymbol{c}_1'\hat{\boldsymbol{\beta}} = 2.5 \qquad se_1 = \sqrt{1/2}$
2. $\boldsymbol{c}_2'\hat{\boldsymbol{\beta}} = 3.5 \qquad se_2 = \sqrt{1/2}$
3. $\boldsymbol{c}_3'\hat{\boldsymbol{\beta}} = -3 \qquad se_3 = \sqrt{2}$
4. $\boldsymbol{c}_4'\hat{\boldsymbol{\beta}} = 2 \qquad se_4 = 1$
5. $\boldsymbol{c}_5'\hat{\boldsymbol{\beta}} = 1 \qquad se_5 = 1$
6. $\boldsymbol{c}_6'\hat{\boldsymbol{\beta}} = 3 \qquad se_6 = \sqrt{2/3}$.

Determine appropriate intervals to address questions 1 through 6.

Each interval is of the form

$$\boldsymbol{c}_j'\hat{\boldsymbol{\beta}} \pm kse_j.$$

How shall we choose k?

If we use the Bonferroni approach, then

$$k = t_{40-4, \frac{0.05}{(2)(6)}} \approx 2.79.$$

This approach would be legitimate if we were interested in these 6 intervals, and only these 6 intervals, prior to observing the data.

Alternatively, we can consider Scheffé intervals, $k = \sqrt{qF_{q,40-4,0.05}}$.

What is the value of q in our situation?

Recall that Scheffé intervals have the property

$$\mathbb{P}(\boldsymbol{u}'\boldsymbol{\theta} \in L(\boldsymbol{u}, k) \quad \forall\, \boldsymbol{u} \in \mathbb{R}^q) = 1 - \alpha,$$

where

$$\boldsymbol{\theta} = \boldsymbol{C}\boldsymbol{\beta}$$

for some $\underset{q \times p}{\boldsymbol{C}}$ of rank q.

If we choose $\underset{q\times p}{\boldsymbol{C}}$ so that

$$\boldsymbol{c}_1'\boldsymbol{\beta}, \ldots, \boldsymbol{c}_q'\boldsymbol{\beta} \in \{\boldsymbol{u}'\boldsymbol{\theta} : \boldsymbol{u} \in \mathbb{R}^q\},$$

then we will have Scheffé intervals that give simultaneous coverage at least $1-\alpha$ for $\boldsymbol{c}_1'\boldsymbol{\beta}, \ldots, \boldsymbol{c}_6'\boldsymbol{\beta}$.

Because $qF_{q,n-r,\alpha}$ is an increasing function of q, we would like to pick q no larger than necessary.

The matrix

$$\begin{bmatrix} 2\boldsymbol{c}_1' \\ 2\boldsymbol{c}_2' \\ \boldsymbol{c}_3' \\ \boldsymbol{c}_4' \\ \boldsymbol{c}_5' \\ 3\boldsymbol{c}_6' \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & -1 & -1 & -1 \end{bmatrix} \text{ has rank 3.}$$

To see this, note that each row is a contrast vector (sums to zero) and is thus in $\mathcal{N}(\underset{4\times 1}{\mathbf{1}'})$, which has dim 3.

Thus, rank ≤ 3. The last 3 rows are LI, so the rank is exactly 3.

Thus,

$$dim(span\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_6\}) = 3.$$

We can take $q = 3$ to obtain

$$k = \sqrt{3F_{3,40-4,0.05}} \approx 2.93.$$

The resulting intervals

$$\boldsymbol{c}_j'\hat{\boldsymbol{\beta}}_j \pm 2.93se_j \quad j = 1, \ldots, 6$$

have simultaneous coverage at least 95%.

We could also include additional intervals for other $\boldsymbol{c}'\boldsymbol{\beta}$ without changing k or losing the guarantee of simultaneous coverage probability at least 95%, as long as $\boldsymbol{c} \in \mathcal{C}(\boldsymbol{C}')$.

We have not specified $\boldsymbol{C}$ explicitly, but we can choose $\boldsymbol{C} \ni$ the rows of $\boldsymbol{C}$ form a basis for the set of all contrasts; i.e., $\mathcal{N}(\underset{4\times 1}{\mathbf{1}'})$.

This Scheffé method for all possible contrasts allows us to construct as many intervals as we wish and still have simultaneous coverage probability at least $1 - \alpha$, provided each interval is for a $\boldsymbol{c}'\boldsymbol{\beta} \ni \mathbf{1}'\boldsymbol{c} = 0$.

We can even examine the data to decide which contrasts appear to be of most interest.

When using the Bonferroni method, intervals of interest must be preplanned before observing the data.

If the contrasts in this example were preplanned, Bonferroni would be preferred over Scheffé because the Bonferroni intervals are narrower.