

Section 11: Comparisons & Contrasts

William Christensen



Consider the one-way ANOVA (BF[1]) model:

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad i = 1, \dots, I; \quad j = 1, \dots, n$$

The purpose of the F-test for the hypothesis

$$H_o : \alpha_1 = \alpha_2 = \dots = \alpha_I = 0$$

is screening – that is, to discover if some α_i is different from others.

The F-test does NOT help you see which means are different or best.

- Comparisons and contrasts are more focused on specific question groups



One-factor design:

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad i = 1, \dots, 5; \quad j = 1, \dots, 25$$

	"group"				
	no comp	pregnant (1)	pregnant (8)	virgin (1)	virgin (8)
Group Means	63.56	64.8	63.36	56.76	38.72

$$H_o : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0 \quad \begin{array}{l} \swarrow \\ \searrow \end{array} \text{general}$$

H_q : At least one α_i is different from the others

But more specific hypothesis/questions exist for group comparisons, where a comparison is a measure of distance between means for two groups of observed values.

	"group"				
	no comp	pregnant (1)	pregnant (8)	virgin (1)	virgin (8)
Group Means	63.56	64.8	63.36	56.76	38.72

1 Pairwise comparisons of μ_i vs. $\mu_{i'}$; (where $\mu_i = \mu + \alpha_i$ and $i \neq i'$)

$$\begin{array}{cccc} \mu_1 = \mu_2? & \mu_1 = \mu_3? & \mu_1 = \mu_4? & \mu_1 = \mu_5? \\ \mu_2 = \mu_3? & & \mu_2 = \mu_4? & \mu_2 = \mu_5? \\ & & \mu_3 = \mu_4? & \mu_3 = \mu_5? \\ & & & \mu_4 = \mu_5? \end{array}$$

In general, how many pairwise comparisons will there be for g groups?

- Is the true difference equal to zero? ←use hypothesis test
- What is the range of plausible values for true difference? ←CI

In Section 3, we discussed a general approach for evaluating a single difference in means (in an ANOVA setting):

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2, \text{df}_{\text{Error}}} \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \text{MS}_{\text{Error}}}$$

and

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \text{MS}_{\text{Error}}}} \sim t_{\text{df}_{\text{Error}}}$$

Here we extend this general approach to handle:

- multiple comparisons
- comparisons more complicated than just $\mu_1 - \mu_2$

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ≡ ≡ ≡ ↺ ↻ ↻

6/17

Multiple Comparisons

When we only have two groups, only one comparison needs to be made:

$$\mu_{1.} - \mu_{2.}$$

For 3 groups, we have

$$\begin{array}{ll} \mu_{1.} - \mu_{2.} & \mu_{1.} - \mu_{3.} \\ & \mu_{2.} - \mu_{3.} \end{array} \quad \binom{3}{2} = 3 \text{ comparisons}$$

For 5 groups, we have

$$\begin{array}{llll} \mu_{1.} - \mu_{2.} & \mu_{1.} - \mu_{3.} & \mu_{1.} - \mu_{4.} & \mu_{1.} - \mu_{5.} \\ & \mu_{2.} - \mu_{3.} & \mu_{2.} - \mu_{4.} & \mu_{2.} - \mu_{5.} \\ & & \mu_{3.} - \mu_{4.} & \mu_{3.} - \mu_{5.} \\ & & & \mu_{4.} - \mu_{5.} \end{array} \quad \binom{5}{2} = 10 \text{ comparisons}$$

- If there are no differences in the $\mu_{i.}$, what is the chance of at least one Type I error?

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ≡ ≡ ≡ ↺ ↻ ↻

7/17

Family-wise (Type I) error rate

- Def’n: The chance of at least one Type I error among a “family” of tests, assuming there are no differences in the group means.

Basic idea of multiple comparison procedures:

- Instead of the individual tests each having a Type I error probability of α (eg., 0.05), adjust your tests or C.I.'s so that the family-wise error rate is α .

Adjustments for multiple comparisons:

- instead of: difference $\pm(t_{\alpha/2, \text{df}_{\text{error}}})(\text{S.E.})$
- use: difference $\pm(\text{adjusted distance})(\text{S.E.})$

Let g be the number of groups.

METHOD	ADJUSTMENT	WHEN TO USE	SIZE
Fisher's LSD	$t_{\alpha/2, dfError}$ (no adjustment)	Screening study (you're willing to live with high familywise error rate)	Most narrow
Tukey's HSD	$q/\sqrt{2}$ (q is based on expected difference between highest group mean minus lowest group mean when in fact all g of the group means are equal—depends on g and dfError)	Looking at all pairwise (group mean) comparisons	Between Scheffe' and LSD
Scheffe'	$\sqrt{(F_{\alpha, g-1, dfError})(g-1)}$ (uses the F critical value you would use for testing group)	Post hoc comparisons (data snooping)	Widest (usually)
Bonferroni	$t_{\alpha/2k, dfError}$ (where k is the number of comparisons you have planned)	Limited number of <i>a priori</i> comparisons	Narrower than Tukey if k is small

EX SAS: fruit flies (pairwise comparisons)

Contrasts

Consider the comparisons implied by the structure of our groups.
We'll focus on weighted estimates of the form

$$c_1\mu_1. + c_2\mu_2. + \dots + c_I\mu_I.$$

■ EX BF[1] - fruit flies

■ ex. 1

$$\mu_{1.} - \mu_{2.} \Rightarrow$$

$\underline{c_1}$	$\underline{c_2}$	$\underline{c_3}$	$\underline{c_4}$	$\underline{c_5}$
1	-1	0	0	0

■ ex. 2

$$\underbrace{\frac{\mu_{2.} + \mu_{3.}}{2} - \frac{\mu_{4.} + \mu_{5.}}{2}}_{\text{pregnant vs. virgin}} \Rightarrow$$

0	.5	.5	-.5	-.5
---	----	----	-----	-----

■ ex. 3

$$\underbrace{\mu_{1.} - \frac{\mu_{2.} + \mu_{3.} + \mu_{4.} + \mu_{5.}}{4}}_{\text{no comp vs. comp}} \Rightarrow$$

1	-.25	-.25	-.25	-.25
---	------	------	------	------

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ≡ ≡ ↺ ↻ ↻

10/17

Def'n: For linear estimates of $c_1\mu_1. + c_2\mu_2. + \dots + c_I\mu_I.$, the linear estimator is called a contrast if $\sum_{i=1}^I c_i = 0$

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ≡ ≡ ↺ ↻ ↻

11/17

- **EX** BF[1] - fruit flies

	$\underline{c_1}$	$\underline{c_2}$	$\underline{c_3}$	$\underline{c_4}$	$\underline{c_5}$
■ ex. 1 $\mu_{1.} - \mu_{2.} \Rightarrow$	1	-1	0	0	0
■ ex. 2 $\underbrace{\frac{\mu_{2.} + \mu_{3.}}{2} - \frac{\mu_{4.} + \mu_{5.}}{2}}_{\text{pregnant vs. virgin}} \Rightarrow$	0	.5	.5	-.5	-.5
■ ex. 3 $\underbrace{\mu_{1.} - \frac{\mu_{2.} + \mu_{3.} + \mu_{4.} + \mu_{5.}}{4}}_{\text{no comp vs. comp}} \Rightarrow$	1	-.25	-.25	-.25	-.25

- Which of the examples above (ex.1, ex.2, ex.3) are contrasts?

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ≡ ≡ ≡ ↺ ↻

12/17

Philosophical Interlude

Do we always use a multiple comparison procedure to adjust α when looking at contrasts/comparisons?

- Opinions/philosophies vary
- For this class, we will take the conservative approach and use multiple comparison adjustments when testing multiple contrasts

EX Testing 3 contrasts? Use Bonferroni adjustment and compare each contrast's p-value with $\alpha/3$ instead of α

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ≡ ≡ ≡ ↺ ↻

13/17

Contrasts in SAS GLM for a BF[k]

- (Step 1) Write down contrast in terms of treatment means (or cell means)
- (Step 2) Write contrast in terms of model parameters (eg., $\alpha_i, \beta_j, (\alpha\beta)_{ij}$)
- (Step 3) Write contrast or estimate statement in PROC GLM
 - **EX** BF[1] - fruitflies: $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$
mean for pregnant vs. mean for virgin:

■ Step 1: $\frac{\mu_{2.} + \mu_{3.}}{2} - \frac{\mu_{4.} + \mu_{5.}}{2}$

■ Step 2: $\frac{\mu + \alpha_2 + \mu + \alpha_3 - (\mu + \alpha_4 + \mu + \alpha_5)}{2}$
 $= \frac{1}{2}(\alpha_2 + \alpha_3 - \alpha_4 - \alpha_5)$

■ Step 3: contrast 'preg vs virgin' group 0 .5 .5 -.5 -.5;

14/17

Ex. BF[2] — wear data: $y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$

Means for Factor Levels and Cells

Proportion of filler

	25% (1)	50% (2)	75% (3)	
Filler type 1	164	187.5	232	194.5 ← $\bar{y}_{1.}$
Filler type 2	148.5	113.5	143.5	135.167 ← $\bar{y}_{2.}$
	156.25 ← $\bar{y}_{.1.}$	150.5 ← $\bar{y}_{.2.}$	187.75 ← $\bar{y}_{.3.}$	$\bar{y}_{...} = 164.833$

"Cell (2,1)"

$\bar{y}_{21.}$

$\bar{y}_{13.}$

(Step 1) $\mu_{1..} - \mu_{2..}$ (average of filler 1 - avg of filler 2)

- Note: $\mu_{1..}$ and $\mu_{2..}$ are not treatment means — need to write these group means in terms of the treatment means $\mu_{11.}, \mu_{12.}, \dots, \mu_{23.}$

$$\mu_{1..} - \mu_{2..} = \frac{\mu_{11.} + \mu_{12.} + \mu_{13.}}{3} - \frac{\mu_{21.} + \mu_{22.} + \mu_{23.}}{3}$$

15/17

(Step 2)

$$\begin{aligned} & \frac{1}{3} \begin{pmatrix} \mu + \alpha_1 + \beta_1 + (\alpha\beta)_{11} \\ + \mu + \alpha_1 + \beta_2 + (\alpha\beta)_{12} \\ + \mu + \alpha_1 + \beta_3 + (\alpha\beta)_{13} \end{pmatrix} - \frac{1}{3} \begin{pmatrix} \mu + \alpha_2 + \beta_1 + (\alpha\beta)_{21} \\ + \mu + \alpha_2 + \beta_2 + (\alpha\beta)_{22} \\ + \mu + \alpha_2 + \beta_3 + (\alpha\beta)_{23} \end{pmatrix} \\ &= \alpha_1 + \frac{1}{3}(\alpha\beta)_{11} + \frac{1}{3}(\alpha\beta)_{12} + \frac{1}{3}(\alpha\beta)_{13} \\ &\quad - \left(\alpha_2 + \frac{1}{3}(\alpha\beta)_{21} + \frac{1}{3}(\alpha\beta)_{22} + \frac{1}{3}(\alpha\beta)_{23} \right) \end{aligned}$$

(Step 3)

contrast 'f1-f2' filler $\begin{matrix} 1 & -1 \\ \alpha_1 & \alpha_2 \end{matrix}$

filler*prop $\begin{matrix} .333 & .333 & .333 & -.333 & -.333 & -.333; \\ (\alpha\beta)_{11} & (\alpha\beta)_{12} & (\alpha\beta)_{13} & (\alpha\beta)_{21} & (\alpha\beta)_{22} & (\alpha\beta)_{23} \end{matrix}$

Actually, SAS is smart enough to know what you mean if you say:

contrast 'f1-f2' filler 1-1;

since this is a simple comparison of main effect levels

Navigation icons: back, forward, search, etc.

16/17

EX BF[2] - wear data

What about average of (filler 1 with proportion 3)
minus average of (filler 2 with proportion 2)

(Step 1)

(Step 2)

(Step 3)

Navigation icons: back, forward, search, etc.

17/17