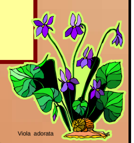
Introduction to Linear Contrasts

Επιστημονική Επιμέλεια

Δρ. Γεώργιος Μενεξές

Τομέας Φυτών Μεγάλης Καλλιέργειας και Οικολογίας,
Εργαστήριο Γεωργίας



Linear contrasts

- Pairwise comparisons of means are a special case of linear contrasts
- Comparison of one group with another group with general **weights**

$$\mu_1 = \mu_2$$
$$\mu_3 = \mu_4$$
$$\mu_1 = (\mu_3 + \mu_4)/2$$

- Linear contrasts take the form of a linear combination of the means

$$L = a_1\overline{Y}_1 + a_2\overline{Y}_2 + \ldots + a_k\overline{Y}_k = \sum a_j\overline{Y}_j$$

Linear contrasts

- Restriction

$$\sum a_j = 0$$

- Suppose we had three means and wanted to compare the first and the second only. This could be done by having $a_1 = 1$, $a_2 = -1$, and $a_3 = 0$

$$L = (1)\bar{Y}_1 + (-1)\bar{Y}_2 + 0\bar{Y}_3 = \bar{Y}_1 - \bar{Y}_2$$

Linear contrasts

- Linear contrasts allow us to express the sum of squared differences between the ***means*** of sets of treatments (formula assumes equal sample sizes)

$$SS_{\text{contrast}} = \frac{nL^2}{\sum a_j^2} = \frac{n\left(\sum a_j \bar{Y}_j\right)^2}{\sum a_j^2}$$

Example

- Suppose we had $\overline{Y}_1 = 1.5, \overline{Y}_2 = 2.0, \overline{Y}_3 = 3.0, n = 10$
- We can easily compute

$$SS_{\text{treat}} = n\sum(\overline{Y}_j - \overline{Y})^2$$
$$= 10\left((1.5-2.17)^2 + (2-2.17)^2 + (3-2.17)^2\right)$$
$$= 10(0.44 + 0.028 + 0.694) = 11.667$$

- Let's say we wanted to compare the average of treatments 1 and 2 to treatment 3. Thus

$$L = \sum a_j \overline{Y}_j = (1)1.5 + (1)2.0 + (-2)3.0 = -2.5$$

$$SS_{\text{contrast}_1} = \frac{nL^2}{\sum a_j^2} = \frac{10(-2.5)^2}{6} = 10.417$$

Example

- Now suppose we wanted to compare groups 1 and 2:

$$L = \sum a_j \overline{Y}_j = (1)1.5 + (-1)2.0 + (0)3.0 = -0.5$$

$$SS_{\text{contrast}_2} = \frac{nL^2}{\sum a_j^2} = \frac{10(-0.5)^2}{6} = 1.25$$

- Note that the sum of the two SS_{contrast} is

$$SS_{\text{treat}} = SS_{\text{contrast}_1} + SS_{\text{contrast}_2}$$
$$11.667 = 10.417 + 1.25$$

- In this case, we can say that the contrasts ***completely partition*** SS_{treat}

F test for contrast

- Note that the ***absolute*** value of the contrast weights does not matter

Means:	Y_1	Y_2	Y_3	Y_4	Y_5
a_j:	2	2	2	-3	-3
	1	1	1	-1.5	-1.5

- The ***significance*** of a contrast can be tested with an F test

$$F = \frac{MS_{\text{contrast}}}{MS_{\text{error}}} = \frac{nL^2 / \sum a_j^2}{MS_{\text{error}}}$$

Contrasts

- The square root of the F for a simple contrast ($a_1 = 1, a_2 = -1$) is the same value obtained in a t test
  - ***t*** **tests are special cases of linear contrasts**
- Note however that if you run several contrasts, the familywise error will be much larger than α
  - Bonferroni correction can be used
  - Run fewer contrasts! Only as many as needed
  - If they were really a priori, probably ok without correction (some people would agree with that)

Orthogonal contrasts

- Some contrasts are independent of one another, while others “share” information
- Independent contrasts are called ***orthogonal***

$$\sum a_j = 0$$

$$\sum b_j = 0$$

$$\sum a_j b_j = 0$$

- Contrasts given by a_j and b_j are orthogonal

Orthogonal contrasts

- It is useful to have orthogonal contrasts because they **exactly partition** SS_{treat}
- However, it is not strictly necessary that all contrasts be orthogonal
  - But remember that in this case the contrasts do **not** convey independent information

Degrees of freedom

- For a contrast, we have

$$F = \frac{MS_{\text{contrast}}}{MS_{\text{error}}}$$

- What are the ***numerator*** degrees of freedom?
- Consider *df* for SS_{contrast}
- A contrast always compares two quantities

$$\mu_1 = \mu_2$$
$$\mu_3 = \mu_4$$
$$\mu_1 = (\mu_3 + \mu_4)/2$$

Degrees of freedom

- Thus SS_{contrast} has $df = 1$
- Another way to think of the *df* is that the *F* for a contrast can in fact be written in the usual way

$$F = \frac{(SSE(R) - SSE(F))/(df_R - df_F)}{SSE(F)/df_F}$$

- To determine df_R, we must determine the number of ***independent*** parameters in the **restricted model** which are associated with the contrast
- Consider the following null hypothesis:

$$H_0 : \tfrac{1}{3}\mu_1 + \tfrac{1}{3}\mu_2 + \tfrac{1}{3}\mu_3 - \mu_4 = 0$$

Degrees of freedom

- The corresponding ***restricted*** model is

$$Y_{ij} = \mu_j + \varepsilon_{ij}$$

where $1/3\mu_1 + 1/3\mu_2 + 1/3\mu_3 - \mu_4 = 0$

- Model has 4 parameters but only 3 are ***independent***
- In the general case for *a* groups, we would have $a - 1$ independent parameters
- Thus

$$df_R - df_F = [N - (a - 1)] - (N - a) \\ = 1$$

Bibliography

- **Φασούλας, Α. Κ. (2006). *Στοιχεία Πειραματικής Στατιστικής*. Θεσσαλονίκη.**
- **Steel, R. & Torrie, J. (1986). *Principles and Procedures of Statistics: A Biometrical Approach*. Singapore: McGraw-Hill Book Company.**
- **Gomez, K. & Gomez, A. (1984). *Statistical Procedures for Agricultural Research*. Singapore: John Willey & Sons, Inc.**
- **Kuehl, R. (2000). *Designs of Experiments: Statistical Principles of Research Design and Analysis*. Pacific Grove: Duxbury Thomson Learning.**