2.5: Contrast Analysis

The paired comparisons discussed in sections 2.2 and 2.3 have the limitation that the comparisons are made only between treatment mean pairs. The contrast analysis procedure can be used to carry out comparisons of a much wider context such as comparisons of treatment level groups or even testing of trends prompting regression modeling to express the response vs. treatment relationship with treatment as a numerical predictor. In the context of a single factor ANOVA model, a linear contrast can be defined as a linear combination of the treatment means such that their numerical coefficients add to zero. Mathematically, a contrast can be represented by...

$$A=\sum_{i=1}^{T} a_i \bar{y}_i \tag{2.5.1}$$

where $\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_T$ represent the sample treatment means and $\sum_{i=1}^{T} a_i=0$. The quantity A is a sample statistic and serves as an estimate for the population contrast $\sum_{i=1}^{T} a_i \mu_i$. By choosing the numerical coefficients appropriately, linear contrasts can be used to make different comparisons among groups of treatment means but not limited to only mean pairs. The table below gives 4 linear contrasts defined in terms of the 3 fertilizer levels F1, F2, F3, and the Control in the greenhouse example.

Table: Greenhouse example contrasts

Ex	a_1	a_2	a_3	a_4	**Contrast**
1	1	-1	0	0	F1-F2
2	1	1	1	-3	F1+F2+F3-3C
3	1	1	-2	0	F1+F2-2F3
4	0	1	-1	0	F2-F3

Notice that values of each list of a_i $(i=1,2,3,4)$ add to zero. The first contrast compares the first two fertilizer types in terms of their means, and the second compares the means of the 3 fertilizer types with the Control mean. The third is a comparison between the combined effect of fertilizer types 1 and 2 with fertilizer type 3, while the last contrast compares the second and third fertilizer types.

A pair of contrasts $A=\sum_{i=1}^{T} a_i \bar{y}_i$ and $B=\sum_{i=1}^{T} b_i \bar{y}_i$ is orthogonal if the products of their numerical coefficients add to zero. This can be expressed mathematically as

$$\sum_{i=1}^{T} a_i b_i=0 \tag{2.5.2}$$

A set of contrasts is said to be orthogonal if every pair of contrasts in the set is orthogonal. Two orthogonal contrasts are not correlated which means that if A and B are orthogonal, then $\text{Covariance}(A, B)=0$. Furthermore, the sum of squares of the treatment usually displayed in the ANOVA table can be partitioned into a set of $(T-1)$ orthogonal contrasts each with 1 degree of freedom. Note that the maximal number of orthogonal contrasts associated with a treatment of T levels is $(T-1)$ and each of them would be associated with one specific comparison independent of each other. In the table above, contrasts 1, 2, and 3 form an orthogonal set of contrasts and contrast 4 cannot be admitted into this set.

The statistical significance of a linear contrast, which can be equated to testing for the zero contrast value can be formulated using the null and alternative hypotheses:

$$H_0: \sum_{i=1}^{T} a_i \mu_i=0 \text{ vs. } H_A: \sum_{i=1}^{T} a_i \mu_i \neq 0 \tag{2.5.3}$$

and can be tested using either,

$$t=\frac{\sum_{i=1}^{T} a_i \bar{y}_i}{\sqrt{\text{MSE} \sum_{i=1}^{T} \frac{a_i^2}{n_i}}} \text{ with } (N-T) \text{ degrees of freedom} \tag{2.5.4}$$

or

$$F=\frac{\left(\sum_{i=1}^{T}a_i\bar{y}_i\right)^2}{\text{MSE}\sum_{i=1}^{T}\frac{a_i^2}{n_i}} \tag{2.5.5}$$

with numerator and denominator degrees of freedom equal to 1 and $(N-T)$ respectively.

Note that MSE can be obtained from the ANOVA table. Applying the above formula, the t-statistic for testing contrast 2 above is...

$$t=\frac{\sum_{i=1}^{T}a_i\bar{y}_i}{\sqrt{\text{MSE}\sum_{i=1}^{T}\frac{a_i^2}{n_i}}}=\frac{28.6+25.867+29.2+(3*21)}{3.052\times\sqrt{\frac{(1+1+1+9)}{6}}}=8.365 \tag{2.5.6}$$

with $df=20$ and has a p-value of .0028, indicating that the average plant height due to the combined treatment of the 3 fertilizer types differs significantly from the average plant height yielded by the control.

The above testing procedure is applicable to non-orthogonal contrasts as well. But, as non-orthogonal contrasts are not guaranteed to be uncorrelated, the conclusions arrived at may be "overlapping" and lead to redundancies. In Chapter 3, examples are provided to illustrate how software can be used to conduct contrast testing. The hypothesis testing for trends using contrasts will be discussed in Chapter 10: ANCOVA II.