MDSC 643.01, Biostatistics I
Brant, Fall 2002

Comparing Cross-classified Samples

Two Factor Analysis of Variance

Randomized Block Designs

Consider experiments where multiple measurements are made on a single experimental "entity" - e.g. human subject, agricultural block (field), generalizing the paired-observation experimental design.

In general we have a setting where k treatments are to be compared where all k have been applied to each of b blocks, resulting in observations

y_{ij} = observation from j'th block for treatment i.

Initial comparisons: As usual with continuous outcomes we start with boxplots, in particular comparative boxplots by treatment (or by individual)

Also consider corresponding means, $\bar{y}_{i.}$ (treatment means) and $\bar{y}_{.j}$ (block means)

Using similar definition to that used in One Way Analysis of Variance, we can define sums of squares that reflect the degree of difference (variance) between treatment and block means.

Between treatment SS = $\sum k(\bar{y}_{i.} - \bar{y})^2$ and

Between block SS = $\sum b(\bar{y}_{.j} - \bar{y})^2$ (note k 's and b 's replace the n_i 's)

It's customary to incorporate these sums of squares in a

Two Way Analysis of variance table.

Source	d.f.	Sum of Squares	Mean Square	F-value
Treatments	k-1	Treatment SS	Treatment SS/(k-1)	Treatment MS/Residual MS
Blocks	b-1	Block SS	Block SS/(b-1)	*
Residual	(k-1)(b-1)	Residual SS	Residual MS =Residual SS/(k-1)(b-1)	
Total	bk-1	Total SS		

* in general, it is not meaningful to test equality of block effects.

In this table we use the term Residual instead of Within (group) for the SS and MS's that reflect random variation. This table represents a book-keeping device that permits calculating the Residual SS by subtracting the Treatment and Block SS's from the Total SS.

To test the hypothesis of equal treatment means, $H_0: \mu_1 = \mu_2 = \mu_3 = \ldots = \mu$ we use

$$F = \frac{Treatment\ SS}{(k-1)\, s^2}$$

, just as we did in the One Way classification, noting that in this case we define s^2 by the $Residual\ MS$. Because of this we use denominator degrees of freedom = $(b-1)(k-1)$, instead of $n-k$

The numerator degrees of freedom = $k-1$, just as before.

e.g.

Pairwise Comparisons

We can examine differences between μ_i and μ_j , using $t = \frac{\bar{y}_i - \bar{y}_j}{s.e.(\bar{y}_i - \bar{y}_j)}$ as always, noting that $s.e.(\bar{y}_i - \bar{y}_j) = s\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$ uses $s = \sqrt{Residual\ MS}$, based on Two-Way ANOVA table and degrees of freedom $(b-1)(k-1)$ for Student's t Confidence intervals can also be constructed in the usual way.