Bonferoni t-test

http://marekrychlik.com/node/61

{The problem can be formulated as follows: Given contrasts (between means) C_1, C_2, ..., C_n, we would like to test the null hypothesis that all of them are 0. We have two significance levels:

- {For an individual contrast. We will call it α_C.
- {Overall, for all contrasts simultaneously. We will call it α_E

{We must assume that the estimators of the contrasts are independent random variables. Perhaps we should recall that a contrast is a property of the joint distribution of the treatment groups:

{

$$C_j = \sum_{i=1}^{t} c_{ij}\mu_i.$$

{An estimator of the contrast is a function of the sample, or a random variable. The mean of the j-th treatment group is estimated as usual:

$$\hat{\mu}_i = \frac{1}{r_i}\sum_{j=1}^{r_i} y_{ij} = \bar{y}_{i\cdot}$$

{The estimators of contrasts are:

$$\hat{C}_j = \sum_{i=1}^{t} c_{ij}\hat{\mu}_i$$

{We also recall that the condition of independence, provided that the usual assumptions of normality and homoskedasticity (homogeneity of varances) is that the covariance between contrasts is 0. We already calculated the covariance to be:

$$Cov(\hat{C}_j, \hat{C}_s) = \sigma^2 \sum_{j=1} \frac{c_{ij}c_{is}}{r_i}$$

where σ^2 is the overall population variance.

{For equal treatment groups, the condition $Cov(\hat{C}_j, \hat{C}_s) = 0$ for $j \neq s$ is equivalent to saying that the matrix (c_{ij}) has orthogonal columns. We recall that matrices of this sort are used to describe contrasts in the package software R. For uneven sizes of treatment groups, we modify the notion of orthogonality, so that the covariance condition still holds.

{The Bonferoni t-test is based on a very simple observation that, if the contrasts are independent random variables, the probability of all of them simultanously being smaller than thenbsp; critical value is $(1-\alpha_C)^n$. Thus, the probability that at least one of them exceeds the critical value is $1-(1-\alpha_C)^n$. Thus the equation

$$\alpha_E = 1-(1-\alpha_C)^n$$

. Hence, in order to test the simultaneous null hypothesis, we simply perform n individual tests with the significance level which is determined from this formula, based on a given value of α_E. Hence:

{

$$\alpha_C = 1-(1-\alpha_E)^{\frac{1}{n}}$$

{Another observation, which together with the prior observations, constitutes the Bonferoni t-test, is the following inequality:

{

$$(1-\alpha_C)^n \geq 1 - n\alpha_C$$

{which holds for all values $\alpha_C \in [0,1]$. Thus, if $\alpha_E = 1-(1-\alpha_C)^n$ then $\alpha_E \leq 1-(1-n\alpha_C) = n\alpha_C$. Hence, if we set:

{

$$\alpha_C = \frac{\alpha_E}{n}$$

{then we get a slightly weaker test, but asymptotically, as $\alpha_C \to 0$, we get essentially no difference.

{Finally, we know that the null hypothesis for an individual contrast is based on the Student t-distribution of the statistic:

{

$$t_j = \frac{\hat{C}_j}{\hat{s}\sqrt{\sum_{i=1}^{t} c_{ij}^2/r_i}} \qquad j = 1, 2 \ldots, n$$

{where

$$\hat{s}^2 = \frac{1}{N-t}\sum_{i,j}(y_{ij} - \hat{\mu}_i)^2$$

is the estimator of the variance.