Power Calculations for Multiple Treatments Design with an Isolated Treatment Effect

Assume T treatment conditions are being studied in either a completely randomized or randomized block design. Under the alternative hypothesis H_1, one treatment is distinguished from the other $T-1$ treatments by exhibiting differential expression for the gene. This computer routine calculates the individual power value $1-\beta_1$ for the design. This power value is the expected fraction of truly differentially expressed genes that will be correctly declared as differentially expressed by the tests.

The difference in expression between the distinguished treatment and the other treatments is μ_1 on the log-intensity scale. The number of replicates or blocks, as the case may be, is n. Hence, there are nT readings on each gene. If the error variance (expected mean square error) of the associated ANOVA model is denoted by σ^2, the non-centrality parameter has the following form.

$$\psi_1 = \frac{n(T-1)}{T}\left(\frac{|\mu_1|}{\sigma}\right)^2 \tag{1}$$

The following list summarizes notation for items used in the computation.

$E(R_0)$: Mean number of false positives.

$\psi_1 = \frac{n(T-1)}{T}\left(\frac{|\mu_1|}{\sigma}\right)^2$: Non-centrality parameter for the design.

T: Total number of treatment conditions.

G_0: Anticipated number of genes in the experiment that are *not* differentially expressed.

Example. Consider a randomized block design involving $T = 6$ treatments and $G_0 = 10,000$ undifferentially expressed genes. The investigator wishes to control the mean number of false positives at $E(R_0) = 2$ and to detect an isolated effect that amounts to a 1.5-fold difference between the distinguished treatment and the other treatments. The 1.5-fold difference represents a value of $|\mu_1| = \log_2(1.5) = 0.5850$ on a log-2 scale. The experimental error standard deviation σ is anticipated to be 0.30 on a log-2 scale. Thus, the ratio $|\mu_1|/\sigma$ equals $0.5850/0.30 = 1.950$. Eight sample units ($n = 8$) are to be used for each treatment condition. For these specifications we have $\psi_1 = 8\frac{(6-1)}{6}(1.950)^2 = 25.35$. The computer routine gives an individual power level of $1-\beta_1 = .70$. Thus, about 70 percent of genes that exhibit an isolated 1.5-fold differential expression in one of the $T = 6$ treatments are expected to be discovered with this study design.

This power calculation can be used iteratively to explore the effect on power of specific design changes. For example, if $n = 10$ sample units are used in lieu of $n = 8$ in each treatment condition, then recalculation of the non-centrality parameter gives $\psi_1 = 31.69$, and this specification gives an individual power level of $1-\beta_1 = .86$, which is somewhat better than previously.