Sample Sizes When the Variance is Estimated (Rule 2.1)

Discussion

The calculation of sample sizes when the variance is estimated in not discussed extensively in the literature. Perhaps one reason is that the formulation is not as straightforward since it will involve the *t* distribution which depends on the sample size. One suggestion that occurs is to substitute *t* values for the normal *z* values and iterate to a sample size. A letter to the editor of *Biometrics* by Steven Julious (*Biometrics*, **60:** 284-285) gives an better and elegant solution to the problem by defining an inflation factor to the sample size.
Julious gives as basic sample size formula

$$n_t = \frac{[TINV(1-\beta, m, z_{1-\alpha/2})]^2 s^2}{(\mu_1 - \mu_2)^2}$$

where TINV is the non-central *t* statistic with *m* degrees of freedom and non-centrality parameter $z_{1-\alpha/2}$ and area $1-\beta$. The subscript *t* on *n* is used to indicate that the sample size is based on the *t* distribution.
This equation is used instead of the usual formula,

$$n_z = \frac{[z_{1-\alpha/2} + z_{1-\beta}]^2 s^2}{(\mu_1 - \mu_2)^2}$$

Julious then calculates the Inflation Factor, *IF*,

$$IF = \frac{n_t}{n_z} = \frac{[TINV(1-\beta, m, z_{1-\alpha/2})]^2}{[z_{1-\alpha/2} + z_{1-\beta}]^2}$$

which does not depend on the sample sizes or the differences to be detected. The inflation factor depends on three quantities: the degrees of freedom on which the estimate of the variability is based, the power, and the Type I error (which is the non-centrality parameter for the non-central *t*). Table 1 presents the inflation factor for power of 0.80, 0.90 and 0.95. The column for power of 0.90 contains the values given by Julious (2004). The table indicates that as the power requirements increase the inflation factor increases also. For example, if the sample variance is based on 30 degrees of freedom, the inflation factor for the sample size calculation is 12% for a power of 0.95 but 6% for a power of 0.80.
While these adjustments can be substantial it's not clear how important they are. In many sample size calculations the investigator will assume a somewhat larger standard deviation than expected, or will calculate a series of sample sizes associated with a range of standard deviations (see for example ROM for April 2004).

Table 1. Inflation factor for two-sample two-tailed situation where the standard deviation is estimated from a sample with specified degrees of freedom and power of 0.80, 0.90, or 0.95. Type I error is assumed to be 0.05.

Degrees of Freedom	Power		
	0.80	0.90	0.95
10	1.19	1.30	1.43
15	1.12	1.19	1.26
20	1.09	1.14	1.19
25	1.07	1.11	1.15
30	1.06	1.09	1.12
40	1.04	1.07	1.09
50	1.03	1.05	1.07
100	1.02	1.03	1.04

Figure 1 provides another look at sample sizes associated with estimated standard deviations or variances. This figure is somewhat limited in that it assume an effect size of 0.50. The figure illustrates that the adjustment is highest with highest power (as in Table 1). The figure also illustrates that substituting a *t* statistic for the *z* statistics can be too conservative, that is, producing a larger sample size than the Julious (2004) approach. Another conclusion from the figure is that power drives everything. Greater power results in sample size adjustments greater than the adjustments for the estimation of the variability.

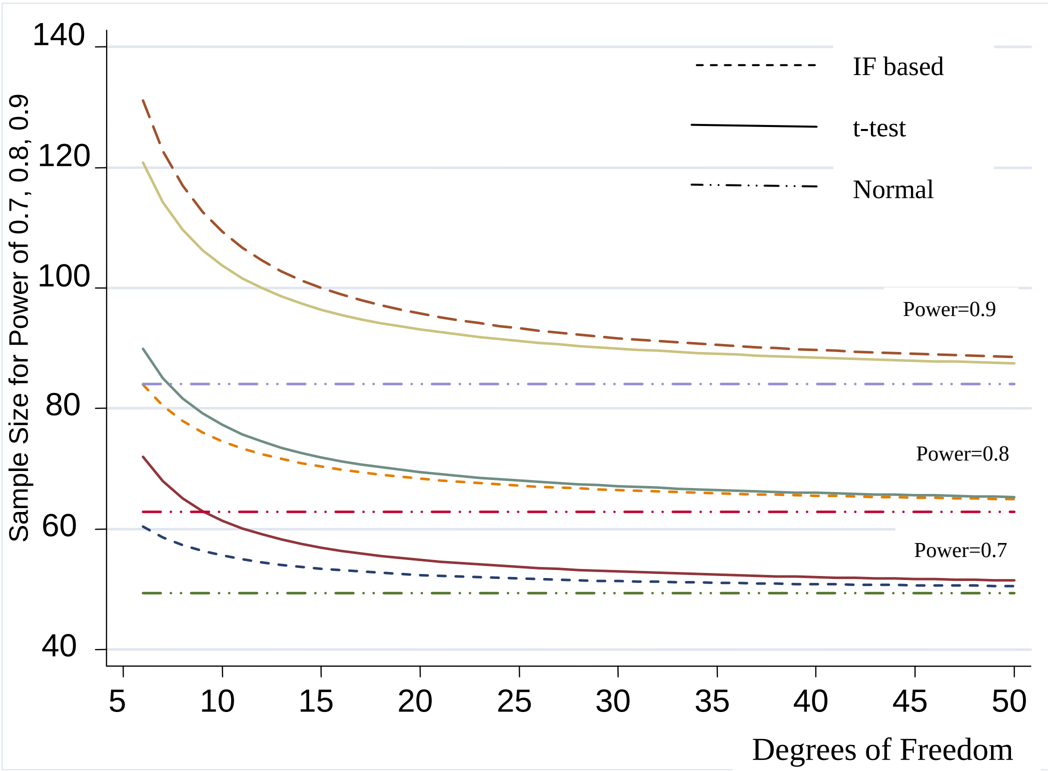

Figure 1. Sample sizes for two sample test based on normal distribution formula, substituting *t* statistics for the *z* statistics, and based on the Julious (2004) formulation. Effect size is 0.5