5.4.1 Sample Size for Point Hypotheses

Let α be the nominal level of significance (i.e., the probability of committing a type I error that one is willing to tolerate); and $\phi = 1 - \beta$, the power one wishes to have to detect a difference of at least Δ magnitude, where β is the probability of making a type II error. In the interest of balance, we assume $n_1 = n_2 = n_e$. In other words, each sequence will be allocated the same number of subjects at random. The sample size per sequence n_e for the hypotheses of equality Equation 5.3.1 can be determined by the formulation given in the following:

$$n_e \geq 2[t(\alpha/2,\ 2n-2) + t(\beta,\ 2n-2)]^2[\widehat{\sigma}_d/\Delta]^2, \tag{5.4.1}$$

where $\widehat{\sigma}_d$ can usually be obtained from previous studies. According to the power approach used on the 80/20 rule, the sample size should be large enough to provide a power of 80% for detection of a difference of the magnitude at least 20% of the unknown reference mean. Thus, Equation 5.4.1 can be simplified as

$$n_e \geq [t(\alpha/2,\ 2n-2) + t(\beta,\ 2n-2)]^2[\mathrm{CV}/20]^2, \tag{5.4.2}$$

where

$$\mathrm{CV} = 100 \times \frac{\sqrt{2\widehat{\sigma}_d^2}}{\mu_R} = 100 \times \frac{\sqrt{\mathrm{MSE}}}{\mu_R}.$$

The total number of subjects required for the standard 2×2 crossover design is $N = 2n_e$. Since the degrees of freedom $(2n - 2)$ in both Equations 5.4.1 and 5.4.2 are unknown, a numerical iterative procedure is required to solve for n_e. To illustrate this, let us consider the following example.

Example 5.4.1

Suppose we would like to conduct a bioequivalence study to compare average bioavailability of a new formulation with a reference formulation as discussed in Example 3.6.1. The design was chosen to be the standard 2×2 crossover design and the 80/20 rule will be used to determine bioequivalence in average bioavailability between two formulations. The next question then is how many subjects are needed to have 80% power to detect a 20% difference. From the data from Example 3.6.1, we have

$$\mathrm{CV} = 100 \times \frac{\sqrt{167.246}}{82.559} = 15.66.$$

Let us first guess $n_e = 9$. This gives degrees of freedom $2n - 2 = 18 - 2 = 16$, $t(0.025,\ 16) = 2.12$ and $t(0.2,\ 16) = 0.865$. By Equation 5.4.2,

$$n_e = (2.12 + 0.865)^2(15.66/20)^2 = 5.5 \cong 6.$$

We then start with $n_e = 6$ and repeat the same calculation, which gives degrees of freedom $= 12 - 2 = 10$,

$$t(0.025,\ 10) = 2.228,$$
$$t(0.2,\ 10) = 0.879.$$

Again, by Equation 5.4.2, we have

$$n_e = (2.228 + 0.879)^2(15.66/20)^2 = 5.9 \cong 6,$$

which is very close to the previous solution.

Therefore, a total of $N = (2)(6) = 12$ subjects are needed based on the 80/20 rule.

As we have pointed out earlier, the power approach based on the 80/20 rule is an ad hoc method for assessment of average bioequivalence that may not be statistically valid. The sample size determined by Equation 5.4.1 or Equation 5.4.2 may not be large enough to provide sufficient power if other methods for interval hypotheses such as Schuirmann's two one-sided tests procedure are used.

5.4.2 Sample Size for Interval Hypotheses

As discussed in Chapter 4, the classic (or shortest) confidence interval, Schuirmann's two one-sided tests procedure, as well as Rodda and Davis's Bayesian method can lead to the same conclusion for determination of bioequivalence in average bioavailability. Therefore, in this section, we will focus on sample size determination based upon Schuirmann's two one-sided tests procedure for interval hypotheses.

As indicated in Section 5.3, the power function $\phi_s(\theta)$ for Schuirmann's two one-sided t tests procedure is symmetric about 0 when the ± 20 rule is used for the assessment of average bioequivalence. Furthermore, the intra-subject variability has an influence on the power function. To illustrate this, Phillips (1990) provided several graphs for the power of Schuirmann's two one-sided t tests procedure for various sample sizes and CVs. Some graphs are presented in Figure 5.4.1. Because calculation for the exact power for Schuirmann's two one-sided t tests procedure requires complicated numerical integration as discussed in Section 5.3, the sample size determination based on the power function is complicated and difficult to obtain. However, an approximate sample size based on the power function can be obtained using some familiar traditional methods (Liu and Chow, 1992a).

We first consider the case where $\theta = \mu_T - \mu_R = 0$ and $n_1 = n_2 = n$. Here,

$$\frac{Y}{\sqrt{\frac{2}{n}\widehat{\sigma}_d^2}}$$

has a central t distribution with $2n - 2$ degrees of freedom. The power at $\theta = 0$ is then given by