Sample Size for a Comparative Study

When Comparing Two Independent Group Means

To calculate the optimal sample size the following 4 quantities must be pre-specified:

1. Assuming there is a true underlying difference, how certain do you want to be of detecting this? i.e. Power, generally we want power = 90% (should be at least 80%)
2. What significance level is difference criterion? The cut off below which we will reject the null hypothesis, generally p=0.05 (5%).
3. Clinically important effect size, d, you wish to detect in the test i.e. the smallest difference in means that it would be clinically meaningful to detect.
4. Variability i.e. Standard Deviation (SD) of the outcome if interest

The above information may be obtained from published studies with similar outcomes and populations or you may need to carry out a pilot study. Then;

\[
patients \text{ per group} = f(\alpha, \beta) \times \frac{2 \times SD^2}{(d)^2}
\]

Where \(f(\alpha, \beta) = 7.85\) or 10.5 for 80% or 90% power respectively with 5% significance. Significance (risk of type I error) is almost always set at 5%.

Example: We wish to conduct a study to compare the mean time taken for patients to recover from two different knee arthroplasty procedures with:
- 90% power
- Significance = 0.05
- The clinically meaningful difference we wish to detect is \(d = 2.5\) days
- \(SD = 5\) days

With 90% power and 0.05 significance \(f(\alpha, \beta) =10.5\)

\[
patients \text{ per group} = f(\alpha, \beta) \times \frac{2 \times SD^2}{(d)^2}
\]

\[
= 10.5 \times \frac{(2 \times 5^2)}{(2.5)^2}
\]

\[
= 84
\]

\(\rightarrow\) 168 patients overall need to be recruited for the trial
When Comparing Two Independent Proportions

To calculate the optimal sample size the following 4 quantities must be pre-specified:
1. Assuming there is a true underlying difference, how certain do you want to be of detecting this? i.e. **Power**, generally we want power = 90% (should be at least 80%)
2. What **significance level** is difference criterion? The cut off below which we will reject the null hypothesis, generally p=0.05 (5%).
3. The assumed proportion that you wish to detect in group 1, \( p_1 \)
4. The assumed proportion that you wish to detect in group 2, \( p_2 \) \([p_1 - p_2 \text{ is the smallest difference in proportions that is clinically important} \]

Then,

\[
\text{patients per group} = \frac{p_1(1-p_1) + p_2(1-p_2)}{(p_2 - p_1)^2} \times f(\alpha, \beta)
\]

**Example:** The UK PACE study compared fitting a single chamber vs. a dual chamber pacemaker on mortality for patients with atrioventricular block. Assumptions:
\( p_1 = 0.24 \) [24% mortality in single chamber group]
\( p_2 = 0.18 \) [18% mortality in dual chamber group]
90% power, 5% significance \([f(\alpha, \beta) = 10.5]\)

\[
\text{patients per group} = \frac{0.35(1-0.35) + 0.15(1-0.15)}{(0.35 - 0.15)^2} \times 10.5 = 70
\]