This document contains the technical details for power analysis for comparing groups nested within clusters such as therapists and schools.

Theory

In this setting, we assume that n clusters formed by therapy groups or schools, nested within each of which are subjects assigned to different treatment conditions.

Consider only two treatment conditions with complete data case and equal number of clusters. Let n denote the number of clusters and m_k the number of subjects in the kth treatment condition. Let y_{kgi} denote the response of ith subject from kth treatment condition and gth cluster. the following LMM:

$$E\left(y_{gki}\right)=\mu_k, \quad Var\left(y_{gki}\right)=\sigma^2, \quad 1\leq g\leq n, \quad 1\leq i\leq m_k, \quad k=1,2.$$

In a matrix form, let

$$\begin{aligned}
\mathbf{y}_g &= \begin{pmatrix} y_{g11} & \cdots & y_{g1m_1} & y_{g1(m_1+1)} & \cdots & y_{g1(m_1+m_2)} \end{pmatrix}^\top, \\
\boldsymbol{\mu} &= \begin{pmatrix} \mu_1 & \cdots & \mu_1 & \mu_2 & \cdots & \mu_2 \end{pmatrix}^\top, \quad 1\leq g\leq n, \\
V &= \begin{pmatrix} \sigma^2 & \sigma^2\rho & \cdots & \sigma^2\rho \\ & \sigma^2 & \cdots & \sigma^2\rho \\ & & \ddots & \vdots \\ & & & \sigma^2 \end{pmatrix} = \sigma^2 C\left(\rho\right).
\end{aligned}$$

Then, we have:

$$E\left(\mathbf{y}_g\right)=\boldsymbol{\mu}, \quad Var\left(\mathbf{y}_g\right)=V=\sigma^2 C\left(\rho\right), \quad 1\leq g\leq n.$$

Now, let $\overline{\mathbf{y}}=\frac{1}{n}\sum_{g=1}^n \mathbf{y}_g$. Then,

$$\sqrt{n}\left(\overline{\mathbf{y}}-\boldsymbol{\mu}\right)\rightarrow_d N\left(\mathbf{0},\Sigma=\sigma^2 C\left(\rho\right)\right).$$

Consider the hypothesis:

$$H_0:\mu_1-\mu_2=K\boldsymbol{\mu}=\begin{pmatrix} \frac{1}{m_1}\mathbf{1}_{m_1} \\ \frac{-1}{m_2}\mathbf{1}_{m_2} \end{pmatrix}^\top \boldsymbol{\mu}=0, \quad \text{vs.} \quad H_0:\mu_1-\mu_2=d.$$

The quadratic $Q_{n0}^2=n\left[K\left(\overline{\mathbf{y}}-\boldsymbol{\mu}\right)\right]^\top\left(K\Sigma K^\top\right)^{-1}\left[K\left(\overline{\mathbf{y}}-\boldsymbol{\mu}\right)\right]$, has an asymptotic central χ_s^2 distribution. Thus, under H_0 and H_a, $Q_n^2=n\left(K\overline{\mathbf{y}}\right)^\top\left(K\Sigma K^\top\right)^{-1}K\overline{\mathbf{y}}$ has a central and non-central χ_s^2 distributions:

$$H_0: Q_n^2\sim\chi_s^2\left(0\right), \qquad H_a: Q_n^2\sim\chi_s^2\left(c\right),$$

where $c=nd^2\left(K\Sigma K^\top\right)^{-1}$. Let $F_{\chi_s^2(c)}$ denote the cdf of $\chi_s^2\left(c\right)$. For a given level of type I error α, let p_α denote the αth percentile of $\chi_s^2\left(0\right)$. The power function, φ, for the linear contrasts (**??**) is given by:

$$\varphi\left(n,c,\alpha\right)=1-F_{\chi_s^2(c)}\left(p_{1-\alpha}\right), \quad \alpha=1-F_{\chi_s^2(0)}\left(p_{1-\alpha}\right).$$