This document contains the technical details for power analysis for linear regression models within a clustered data setting in which clustered data are formed by therapists.

1. Simple linear regression with a single predictor

In this setting, we assume that there are n therapists and each therapist sees m patients. We have a response y and a predictor x and assume a linear relationship between y and x. Power is calculated for detecting a non-zero correlation between y and x. Input parameters are: the non-zero correlation and the intra-class correlation of the therapists' effect.

Let n denote the number of therapists and m the number of patients seen by each therapist. By treating n as the number of clusters and m as the size of each cluster, patients' responses, y_{it}, can be modeled using the following LMM:

$$\begin{aligned} y_{it} &= \beta_0 + x_i\beta_1 + b_i + \epsilon_{it}, \quad b_i \sim \left(0, \sigma_b^2\right), \quad \epsilon_{it} \sim N\left(0, \sigma^2\right), \quad 1 \le i \le n, \quad 1 \le t \le m, \quad (1) \\ \mathbf{y}_i &= \beta_0 \mathbf{1}_m + x_i\beta_1\mathbf{1}_m + b_i\mathbf{1}_m + \boldsymbol{\epsilon}_i \quad 1 \le i \le n, \end{aligned}$$

where x_i is a predictor of interest. In (1), the therapists' effect is explicitly modeled by the random effect b_i.

Now, let $E(x_i) = E(y_{it}) = 0$ and $Var(x_i) = Var(y_{it}) = 1$. Then, it follows from (1) that:

$$\beta_0 = 0, \quad var(x_i\beta_1 + b_i + \epsilon_{it}) = \beta_1^2 + \sigma_b^2 + \sigma^2 = 1. \qquad (2)$$

$$Corr(x_i, y_{it}) = E(x_i y_{it}) = E\left(x_i^2\right)\beta_1 + E(x_i b_i) + E(x_i\epsilon_{it}) = \beta_1. \qquad (3)$$

Thus, β_1 is the correlation between x_i and y_{it} and $\sigma_b^2 + \sigma^2 = 1 - \beta_1^2$.

Now, consider the hypothesis:

$$H_0: \beta_1 = 0, \quad \text{vs.} \quad H_a: \beta_1 = a, \qquad (4)$$

where a is the correlation to be detected.

It follows that the asymptotic variance of the MLE of β_1 is given by:

$$\Sigma_\beta = E^{-1}\left[X_1^\top V_1^{-1} X_1\right], \quad V_1 = \sigma_b^2 J_m + \sigma^2 I_m = \left(\sigma_b^2 + \sigma^2\right) C(\rho) = \left(1 - a^2\right) C(\rho), \qquad (5)$$

where J_m is a $m \times m$ matrix of 1's and $C(\rho)$ the uniform compound symmetry correlation matrix with $\rho = \frac{\sigma_b^2}{1-\beta_1^2}$. The within-cluster correlation ρ is widely known as the intra-class correlation. Thus, the asymptotic variance is a function of β_1 and the intra-class correlation ρ.

The above is similar to Example 5 in Tu et al. (2004) and so power can be computed by adding the following to the program:

a — correlation to be detected

ρ — within subject correlation for subjects seen by the same therapist

We can compute:

$$\sigma_b^2 = \rho\left(1 - a^2\right), \quad \sigma_b^2 + \sigma^2 = 1 - a^2.$$

2. Linear regression with multiple predictors

For convenience, we consider a linear regression with two predictors:

$$\begin{aligned} y_{it} &= \beta_0 + u_i\beta_1 + v_i\beta_2 + b_i + \epsilon_{it}, \quad b_i \sim \left(0, \sigma_b^2\right), \quad \epsilon_{it} \sim N\left(0, \sigma^2\right), \quad 1 \le i \le n, \quad 1 \le t \le m, \\ \mathbf{y}_i &= \beta_0\mathbf{1}_m + u_i\beta_1\mathbf{1}_m + v_i\beta_2\mathbf{1}_m + b_i\mathbf{1}_m + \boldsymbol{\epsilon}_i \quad 1 \le i \le n, \end{aligned}$$

We are interested in detecting a smaller effect size among β_1 and β_2 so we assume a working model as follows:

$$\begin{aligned} y_{it} &= \beta_0 + (u_i + v_i)\beta_1 + b_i + \epsilon_{it}, \quad b_i \sim \left(0, \sigma_b^2\right), \quad \epsilon_{it} \sim N\left(0, \sigma^2\right), \quad 1 \le i \le n, \quad 1 \le t \le m, \\ \mathbf{y}_i &= \beta_0\mathbf{1}_m + (u_i + v_i)\beta_1\mathbf{1}_m + b_i\mathbf{1}_m + \boldsymbol{\epsilon}_i \quad 1 \le i \le n, \end{aligned}$$

By viewing $u_i + v_i$ as a single x_i, we can apply the development in 1 above. Let $Var\left(u_i\right) = Var\left(v_i\right) = 1$ and $Corr\left(u_i, v_i\right) = \rho_{pred}$. In this case, $Var\left(x_i\right) = 2\left(1 + \rho_{pred}\right) \neq 1$. It follows from (1) that:

$$\beta_0 = 0, \quad var\left(x_i\beta_1 + b_i + \epsilon_{it}\right) = \beta_1^2\sigma_x^2 + \sigma_b^2 + \sigma^2 = 1.$$

$$Corr\left(x_i, y_{it}\right) = \frac{E\left(x_i y_{it}\right)}{\sigma_x} = \frac{1}{\sigma_x}\left[E\left(x_i^2\right)\beta_1 + E\left(x_i b_i\right) + E\left(x_i\epsilon_{it}\right)\right] = \sigma_x\beta_1.$$

Thus, $\sigma_x\beta_1$ is the correlation between x_i and y_{it} and $\sigma_b^2 + \sigma^2 = 1 - \sigma_x^2\beta_1^2$.

Now, consider the hypothesis:

$$H_0 : Corr\left(x_i, y_{it}\right) = 0, \quad \text{vs.} \quad H_a : Corr\left(x_i, y_{it}\right) = \sigma_x\beta_1 = a,$$

or equivalently,

$$H_0 : \beta_1 = 0, \quad \text{vs.} \quad H_a : \beta_1 = \frac{a}{\sigma_x},$$

where a is the correlation to be detected.

It follows that the asymptotic variance of the MLE of β_1 is given by:

$$\Sigma_\beta = E^{-1}\left[X_1^\top V_1^{-1} X_1\right], \quad V_1 = \sigma_b^2 J_m + \sigma^2 I_m = \left(\sigma_b^2 + \sigma^2\right)C\left(\rho\right) = \left(1 - \sigma_x^2\beta_1^2\right)C\left(\rho\right),$$

where J_m is a $m \times m$ matrix of 1's and $C\left(\rho\right)$ the uniform compound symmetry correlation matrix with $\rho = \frac{\sigma_b^2}{1-\sigma_x^2\beta_1^2} = \frac{\sigma_b^2}{1-a^2}$. The within-cluster correlation ρ is widely known as the intra-class correlation. Thus, the asymptotic variance is a function of β_1 and the intra-class correlation ρ.

The above is similar to Example 5 in Tu et al. (2004) and so power can be computed by adding the following to the program:

a — correlation to be detected

ρ — within subject correlation for subjects seen by the same therapist

ρ_{pred} — between-predictor correlation