Worked Example for Group or Cluster Regression Discontinuity Designs: Assignment Score is the Outcome at Pre-Test, Cohort Design

This calculator will use the parameter estimates you have provided to calculate the required sample size of groups or clusters per condition for the intervention effect and the level of power that you specified.

In addition, the calculator will use the other parameter estimates to calculate the detectable difference Δ that you specified as a function of the number of groups or clusters g per condition and the number of members m in each group or cluster as described below.

Assume a group or cluster regression discontinuity design (RDD) assigns g groups to the intervention condition if the group-level summary of the outcome at pre-test is above the mean of group-level summaries of the outcome at pre-test (results are the same if the assignment rule is reversed). The remaining g groups are assigned to the control condition. Each group has m participants. Let σ_y^2 be the total variance of the outcome, ρ be the intraclass correlation coefficient (ICC), τ_g be the cluster autocorrelation or over-time correlation at the group-level, and τ_m be the individual autocorrelation or overtime correlation at the member-level. Further, let $r = \frac{m\rho\tau_g}{1+(m-1)\rho} + \frac{(1-\rho)\tau_m}{1+(m-1)\rho}$. From equation 21 in Pennell et al. (2011), the variance of the cohort intervention effect estimator $\hat{\Delta}$ can be written as:

$$\mathrm{Var}(\hat{\Delta}) = 2\frac{m\rho\sigma_y^2(1-R_g^2) + (1-\rho)\sigma_y^2(1-R_m^2)}{gm}\left(\frac{1-r^2}{1-2\pi^{-1}}\right)$$

The quantity π is the familiar mathematical constant, approximately 3.14. Based on Murray (1998), the values R_g^2 and R_m^2 represent the proportion group-level variation and member-level variation, respectively, that are accounted for by covariates.

We can then write the following expression for the minimum detectable difference (Δ) that is available with the level of power that you specified as a function of the number of groups, the members per group, and the ICC according to the following equation.

$$\begin{aligned}\Delta &= \sqrt{\mathrm{Var}(\hat{\Delta})(t_{\alpha/2}+t_\beta)^2}\\ &= \sqrt{\left(2\frac{m\rho\sigma_y^2(1-R_g^2) + (1-\rho)\sigma_y^2(1-R_m^2)}{gm}\left(\frac{1-r^2}{1-2\pi^{-1}}\right)\right)(t_{\alpha/2}+t_\beta)^2}\end{aligned}$$

Here $t_{\alpha/2}$ is the t-value selected based on the two-tailed alpha level and available degrees of freedom, and t_β is the t-value selected based on the desired power and available degrees of freedom.

As an example, consider the following set of parameter estimates.

Let $m = 100$ and $g = 14$ with 1 degree of freedom (df) used for group-level covariates. Then:

df = $2g - 3 - df_g = 2(14) - 3 - 1 = 24$.

For a two-tailed type 1 error rate of 5% and power of 80%,

$t_{\alpha/2} = 2.0595$ and $t_\beta = 0.8526$.

Let $\sigma_y^2 = 1$ and $\rho = 0.01$.

Let $R_g^2 = 0.10$ and $R_m^2 = 0.70$.

Let $\tau_g = 0.23$ and $\tau_m = 0.26$. Then $r = \frac{100 \cdot 0.01 \cdot 0.23}{1+(100-1)\cdot 0.01} + \frac{(1-0.01)\cdot 0.26}{1+(100-1)\cdot 0.01} = 0.2449$.

Then

$$\Delta = \sqrt{\left(2\frac{m\rho\sigma_y^2\left(1-R_g^2\right)+(1-\rho)\sigma_y^2(1-R_m^2)}{gm}\left(\frac{1-r^2}{1-2\pi^{-1}}\right)\right)\left(t_{\alpha/2}+t_\beta\right)^2}$$

$$= \sqrt{\left(2\frac{100\cdot 0.01\cdot 1\cdot(1-0.10)+(1-0.01)\cdot 1\cdot(1-0.70)}{14\cdot 100}\left(\frac{1-0.2449^2}{1-2\pi^{-1}}\right)\right)(2.0595+0.8526)^2}$$

$$= 0.194$$

The minimum detectable difference given these parameter estimates is 0.194 standard deviation units.

References

Murray DM. **Design and analysis of group-randomized trials.** 1998 New York: Oxford University Press.

Pennell ML, Hade EM, Murray DM, Rhoda DA. **Cutoff designs for community-based intervention studies.** Stat Med. 2011;30(15):1865-82. Epub 2011/04/17. PMID: 21500240.