Worked Example for Group or Cluster Regression Discontinuity Designs: Assignment Score is a Non-Outcome Variable at Pre-Test

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 an assignment score at pre-test is above some cut-off value (results are the same if the assignment rule is reversed). The remaining g groups are assigned to the control condition – thus, the treatment group sampling rate p is 0.50. Each group has m participants. Let σ_y^2 be the total variance of the outcome and let ρ be the intraclass correlation coefficient (ICC). Using the RDD effect described in equation 11 of Schochet (2009) and the unmatched, unstratified group-randomized trial analysis described by Murray (1998) the minimum detectable different (Δ) that is available at the level of power specified can be written as

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

Here R_m^2 is the proportion of member-level variation explained by covariates, R_g^2 is the proportion of group-level variation explained by covariates, ρ_{ts} is the correlation between the assignment score and intervention indicator, $t_{\alpha/2}$ 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. The quantity $\frac{1}{1-\rho_{ts}^2}$ is interpreted as the RDD effect.

Assume the assignment score follows a normal distribution, with the mean of the distribution at the threshold value. Thus, the percentage of the score distribution lying above the cut-off value q equals p – i.e., $q = p = 0.50$. Then from table 2 in Schochet (2009), the correlation between intervention and assignment score, ρ_{ts}, is

$$\begin{aligned}\rho_{ts} &= \frac{\phi\left(\Phi^{-1}(p)\right)}{\sqrt{p(1-p)}} \\ &= \frac{\phi\left(\Phi^{-1}(0.50)\right)}{\sqrt{0.50(1-0.50)}} \\ &= 0.7979,\end{aligned}$$

where ϕ is the standard normal probability density function and Φ is the inverse distribution function for the standard normal distribution. This yields an RDD effect of $\frac{1}{1-\rho_{ts}^2} = \frac{1}{1-0.7979^2} = 2.7519$.

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

Let $m = 55$.

Let $g = 34$ with one group-level covariate. Then:

df = $2g - 3 - df_g = 2(34) - 3 - 1 = 64$.

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

$t_{\alpha/2} = 1.9977$ and $t_{\beta} = 0.8473$.

Let $\sigma_y^2 = 1$ and $\rho = 0.15$.

Let $R_g^2 = 0.50$ and $R_m^2 = 0.50$.

Then

$$\begin{aligned}\Delta &= \sqrt{2\left(\frac{\sigma_y^2(1-\rho)(1-R_m^2) + m\sigma_y^2\rho\left(1-R_g^2\right)}{mg}\left(\frac{1}{1-\rho_{ts}^2}\right)\right)\left(t_{\alpha/2}+t_{\beta}\right)^2} \\ &= \sqrt{2\left(\frac{1^2(1-0.15)(1-0.50) + 55\cdot 1^2\cdot 0.15\cdot(1-0.50)}{55\cdot 34}(2.7519)\right)(1.9977+0.8473)^2} \\ &= 0.33\end{aligned}$$

The minimum detectable difference given these parameter estimates is 0.33 standard deviation units.

References

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

Schochet PZ. **Statistical Power for Regression Discontinuity Designs in Education Evaluations.** J Educ Behav Stat. 2009;34(2):238-66. https://doi.org/10.3102/1076998609332748.