Validation guide

December 27, 2021

Each document lays out results for estimating statistical power, minimum detectable effect size (MDES), and sample size for a specific design. The document may also contain multiple models, such as constant, fixed, or random treatment effects.

We compare results of Monte Carlo simulations, PUMP, and the PowerUp R package.

Adjustment procedures:

- Bonferroni
- Benjamini Hochberg (BH)
- Holm
- Westfall-Young Single Step (WY-SS)
- Westfall-Young Step Down (WY-SD)

Power Validation

In this section, we validate power results for different definitions of power.

Power definitions:

- Individual power for each outcome (M = 3): D1indiv, D2indiv, D3indiv
- Mean individual power
- Minimum power: min1, min2
- Complete power

Parameters that vary:

- school size: $\bar{n}$ is 100, 75, 50
- R2: $R^2 = 0$ by default, we also try 0.6
- rho: $\rho = 0.5$ by default, we also try $\rho = 0.2$ and $\rho = 0.8$
- true positives: the ATE in effect size units is $(0.125, 0.125, 0.125)$ by default, we also try $(0.125, 0, 0)$
- ICC: ICC = 0.2 by default. If relevant, we try ICC = 0.7
- omega: $\omega = 0.1$ by default. If relevant, we try $\omega = 0.8$

MDES and sample size validation

We choose one default scenario for each design and model, then input the already-calculated D1 individual power and see if the output MDES is the same as the original input MDES.

Similarly, for sample size validation, we input the already-calculated D1 individual power and see if the output sample size (either $\bar{n}$, J, or K depending on design) is the same as the original sample size.

Remark In some cases, there will be a discrepancy between the input sample size and the return sample size when the power curve is relatively flat, and a variety of sample sizes result in similar power.