It's time to simulate!

Based on the parameters you entered or the parameters we obtained from your data, we will now simulate similar data under H0 (i.e., there is only one population/cluster) to detect Type-1 risk, or the risk to find clusters when they do not exist, and under H1 (i.e., there are two cluster in the population) if you also want to know what is your power to detect two clusters when your data have the specified parameters and the clusters really exist.

If you select 'Type1', we will only simulate data under H0 and Cohen's ds will be set to 0.

If you select 'Power', we will also simulate data under H1 and Cohen's ds will be sampled from the preferred range. In this case we will count how many times the cluster analys exactly detects 2 clusters

Select an Option:

Mclust

Select type of analysis:

Power

N = your sample size

20 200 2,000

20 220 420 620 820 1,020 1,220 1,420 1,620 1,820 2,000

Cohen's d (meaningful only when computing Power):

0 1 1.5

0 0.15 0.3 0.45 0.6 0.75 0.9 1.05 1.2 1.35 1.5

Cohen's d will be automatically set to 0 for Type-1 error simulations

Maximum computation time (in seconds):

10 60 300

10 40 70 100 130 160 190 220 250 280 300

We will simulate data until time expires

Start simulation

Results of the simulation

ATTENTION: if you have many indicators and a large sample size, the simulation might be slow and the system crash for computation limits. Please download the R code and run your own simulation.

Method	Iterations	Indicators	SampleSize	Correlations	Skewness	Kurtosi	power	T1error
mclust	197/180	3	200	[0.08; 0.2]	[0.02; 0.19]	[-0.04; 0.17]	0.69	0.26

Results with fewer than 1,000 iterations may not be stable.

POWER ANALYSIS — During the desired time we managed to simulate 197 random datasets with 3 indicators and 200 observations. Correlations between variables ranged between 0.08 and 0.20. Skewness ranged between 0.02 and 0.19. Kurtosis ranged between -0.04 and 0.17. We applied the cluster analysis with the mclust method to each dataset and calculated the probability of incurring in an error. Effect size/cluster separation (Cohen's d) range between 1.00 and 1.50 across indicators. Your estimated power to detect two clusters is 0.69 with an Adjusted Rand index of 0.35

TYPE I ERROR — During the desired time we managed to simulate 180 random datasets with 3 indicators and 200 observations. Correlations between variables ranged between 0.08 and 0.20. Skewness ranged between 0.02 and 0.19. Kurtosis ranged between -0.04 and 0.17. We applied the cluster analysis with the mclust method to each dataset and calculated the probability of incurring in an error. Your estimated probability of type-1 errors is 0.26