




POSSA: Power simulation for sequential analyses and multiple hypotheses

Gáspár Lukács ¹

1 Tilburg University, Department of Methodology and Statistics, The Netherlands

DOI: [10.21105/joss.04643](https://doi.org/10.21105/joss.04643)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Frederick Boehm](#)  

Reviewers:

- [@mingzhuang](#)
- [@mmrabe](#)

Submitted: 20 July 2022

Published: 28 August 2022

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

Summary

Conventional experiment designs have a fixed sample size: observations are collected until the planned sample size is reached, and then the experiment is concluded. Alternatively however, sequential designs can also be used (e.g., [Lakens, 2014](#)). In that case, interim analyses can be performed during data collection, and whenever the analysis suggests sufficient evidence for the presence or absence of the given effect, the experiment can be concluded without further data collection. This can drastically reduce the required sample sizes, saving time, expenses, and effort. However, sequential analyses require adjustments for Type 1 error rate (i.e., the ratio of false significant findings) and affect statistical power (which is crucial to determine the required sample size, e.g., [Cohen, 1988](#); [Perugini et al., 2018](#)). Furthermore, regardless of sequential analysis, in case of multiple hypotheses (multiple tests included in an analysis), for both Type 1 error rate and power, correction is necessary (e.g., [Khandis & Gangestad, 2020](#)).

The present POSSA R package serves to perform, via a simulation framework, the necessary adjustments and calculate power for sequential analyses as well as for multiple hypotheses for practically any significance test.

Statement of need

While sequential designs can be extremely useful, their application is limited by the lack of software that performs the necessary statistical adjustments. Recently, there have been several new related software solutions (e.g., the R packages by [Anderson, 2022](#); [Pahl, 2022](#); [Wassmer & Pahlke, 2022](#); for more, see [Weigl & Ponocny, 2020](#)), but all of these apply only to single and specific parametric tests. Solutions for simulation-based power analysis also exist (e.g., [Green & MacLeod, 2016](#); [Hughes, 2017](#); [Lakens & Caldwell, 2021](#)), but none specifically for sequential designs or multiple hypotheses. The POSSA R package allows power and Type 1 error calculations and corrections for analyses in fixed as well as sequential designs including any, and any number of, statistical null hypothesis tests.

Example

A simplistic example is presented below to demonstrate how POSSA works.

First, create a custom sampling function, simulating the data that would be obtained in an actual single run of the experiment in case of (a) existing true effect in the population, and (b) no effect. In this example, it is simply two normally distributed sets of numbers.

```
customSample = function(sampleSize) {  
  list(  
    # baseline  
    sample1 = rnorm(sampleSize, mean = 0, sd = 10),
```

```
# no effect
sample2_h0 = rnorm(sampleSize, mean = 0, sd = 10),
# true effect
sample2_h1 = rnorm(sampleSize, mean = 5, sd = 10)
)
}
```

Second, create a corresponding custom testing function, including all tests that would be conducted on the obtained samples. Here, it is simply a single *t*-test.

```
customTest = function(sample1, sample2_h0, sample2_h1) {
  c(
    # no effect
    p_h0 = t.test(sample1, sample2_h0, 'less', var.equal = TRUE)$p.value,
    # true effect
    p_h1 = t.test(sample1, sample2_h1, 'less', var.equal = TRUE)$p.value
  )
}
```

Third, pass the sampling and testing functions to the `POSSA::sim` function, with any desired observation numbers (at the interim stops and at the final stop). This will simulate and return a great number of *p* values.

```
dfPvalsSeq = sim(fun_obs = customSample,
                 n_obs = c(27, 54, 81),
                 fun_test = customTest)
```

Fourth and last, pass the obtained `dfPvalsSeq` data.frame to the `POSSA::pow` function. In this case, `alpha_locals = NA` specifies that all local alphas should be identical (Pocock, 1977).

```
pow(dfPvalsSeq, alpha_locals = NA)
```

The information returned by this function includes the values of the local alphas (to achieve, by default, a Type 1 error rate of .05), the actual Type 1 error rate, the statistical power, and the average expected sample sizes in case of a true effect and in case of no effect.

The same workflow can be used to accommodate any other analyses.

References

- Anderson, K. (2022). *gsDesign: Group sequential design*. <https://CRAN.R-project.org/package=gsDesign>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum. <https://doi.org/10.4324/9780203771587>
- Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Hughes, J. (2017). *paramtest: Run a function iteratively while varying parameters*. <https://CRAN.R-project.org/package=paramtest>
- Khandis, R. B., & Gangestad, S. (2020). On attenuated interactions, measurement error, and statistical power: Guidelines for social and personality psychologists. *Personality and Social Psychology Bulletin*, 46(12), 1702–1711. <https://doi.org/10.1177/0146167220913363>
- Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses: Sequential analyses. *European Journal of Social Psychology*, 44(7), 701–710. <https://doi.org/10.1002/ejsp.2023>

- Lakens, D., & Caldwell, A. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, 4(1), 251524592095150. <https://doi.org/10.1177/2515245920951503>
- Pahl, R. (2022). *GroupSeq: Group sequential design probabilities - with graphical user interface*. <https://CRAN.R-project.org/package=GroupSeq>
- Perugini, M., Gallucci, M., & Costantini, G. (2018). A practical primer to power analysis for simple experimental designs. *International Review of Social Psychology*, 31(1), 20. <https://doi.org/10.5334/irsp.181>
- Pocock, S. J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika*, 64(2), 191–199. <https://doi.org/10.1093/biomet/64.2.191>
- Wassmer, G., & Pahlke, F. (2022). *rpact: Confirmatory adaptive clinical trial design and analysis*. <https://CRAN.R-project.org/package=rpact>
- Weigl, K., & Ponocny, I. (2020). Group sequential designs applied in psychological research. *Methodology*, 16(1), 75–91. <https://doi.org/10.5964/meth.2811>