

mcstatsim: A Functional Programming-Based Monte Carlo Simulation Tool with Parallel Computing Support

Imad EL BADISY¹

¹ Mohammed VI Center for Research and Innovation, Rabat, Morocco

DOI:

Software

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

Submitted:

Published:

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

Summary

Monte Carlo simulation studies, which generate pseudo-random observations from a known truth, are essential for conducting computationally intensive experiments. These simulations provide researchers with powerful tools to evaluate, compare, and assess the performance of various statistical methods, ensuring their reliable application (Boulesteix et al., 2020). Despite their utility, managing the code and maintaining reproducibility becomes increasingly challenging as the complexity of the simulation design grows.

Organizing Monte Carlo simulations is crucial, especially when the design involves numerous varying parameters. Efficiently handling large simulations often requires parallel execution, which depends on the operating system's capabilities to support multicore or multisession parallelism. Additionally, an important aspect of reporting a simulation study is the visualization of results. This requires data to be in a ready format for manipulation, and while data frames are suitable for this purpose, each iteration produces its own set of results. Therefore, a structured and organized approach is necessary to handle and analyze all the data effectively.

Statement of Need

Monte Carlo simulations are a critical component in many statistical analyses and research studies. However, the complexity of setting up and managing simulations can be a barrier, particularly when handling large parameter grids and ensuring reproducibility. `mcstatsim` addresses these challenges by providing a functional programming approach to streamline simulation setups, manage parameters efficiently, and facilitate parallel execution to improve performance.

Key Features

The `mcstatsim` package offers several key features that enhance the efficiency and usability of Monte Carlo simulations:

- It employs a functional programming approach, centralizing the simulation process into a single higher-order function. This enhances manageability and usability by using pure functions that ensure deterministic and stateless computation, leading to reproducible results.
- Parallel execution is leveraged to speed up simulations, supporting both Unix-based and Windows operating systems.

- It delivers a structured output by presenting results in a single dataframe format, simplifying analysis and visualization.
- Additionally, it includes various simulation targets, providing a range of performance measures and their Monte Carlo estimations (Morris, White, & Crowther, 2019).
- The package also features a progression bar, allowing users to track the simulation's progress in real time, enhancing user experience and workflow monitoring.

The functional programming paradigm is particularly suited for Monte Carlo simulations due to its emphasis on immutability and modularity. By breaking down the simulation design into modular, reusable functions, `mcstatsim` allows users to construct complex simulations concisely and systematically.

The core function `runsim()` handles the execution of simulations in `mcstatsim`. It processes simulation parameters via an expanded grid, mapping them to the user-defined simulation function through `mcpmap()`. The results are combined internally and returned deliberately in a structured dataframe, addressing the limitations of list outputs and ensuring scalability and adaptability to various experimental designs.

Conclusion

The `mcstatsim` package offers a simple, yet efficient framework for conducting Monte Carlo simulations using a functional programming approach. Its simplicity and effectiveness make it a valuable tool for researchers in statistical and biostatistical fields, enabling them to perform complex simulations with ease.

References

- Boulesteix, A.-L., Groenwold, R. H., Abrahamowicz, M., Binder, H., Briel, M., Hornung, R., Morris, T. P., et al. (2020). Introduction to statistical simulations in health research. *BMJ open*, 10(12), e039921.
- Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in medicine*, 38(11), 2074–2102.