INSTITUTE FOR DEFENSE ANALYSES

DATAWorks 2024: Monte Carlo Power Evaluation with skpr and skprJMP Live Demo Poster

Robert Atkins, Project Leader

Tyler Morgan-Wall

April 2024

IDA Product ID 3002094

INSTITUTE FOR DEFENSE ANALYSES
730 East Glebe Road
Alexandria, Virginia 22305

About This Publication
This work was conducted by the Institute for Defense Analyses (IDA) under contract HQ0034-19-D-0001, Task BD-9-2299(37), "Shipboard Air Defense Systems," for the Office of the Director, Operational Test and Evaluation. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments
The IDA Technical Review Committee was chaired by Dr. V. Bram Lillard and consisted of Dr. Breeana G. Anderson, Dr. George M. Khoury, and Dr. Jason P. Sheldon from the Operational Evaluation Division, and Dr. Heather M. Wojton from the Executive Division.

For more information:
Dr. Robert Atkins, Project Leader
ratkins@ida.org • (703) 933-6546

Dr. V. Bram Lillard, Director, Operational Evaluation Division
vlillard@ida.org • (703) 845-2230

INSTITUTE FOR DEFENSE ANALYSES

IDA Product ID 3002094

DATAWorks 2024: Monte Carlo Power Evaluation with skpr and skprJMP Live Demo Poster

Robert Atkins, Project Leader

Tyler Morgan-Wall

Executive Summary

Design of Experiments (DoE) provides a solid analytical basis for test planning because it includes tools to characterize the design space, maximizes the value of each run or mission, and links the planning of a test to the data analysis that follows the test. Importantly, DoE allows testers to characterize and minimize risk for a test by calculating statistical power for a particular design and analysis method, providing decision makers with a quantitative assessment of the likelihood of an experiment discovering an effect if one exists. Because of all these benefits, the Government favors test designs that use DoE.

However, the most common power calculation software packages used in the Department of Defense do not support logistic regression, an important and widely used data analysis method used to analyze designs with a binomial response. To address this capability gap, IDA issued guidance in 2014 that recommended the use of the signal-to-noise ratio (SNR) approximation method to calculate power for these types of designs. Since then, developments in statistical software and improvements in computational power have made more accurate and reliable methods available for this purpose.

This poster compares two software products developed at IDA that provide a more accurate and robust Monte Carlo-based method to compute power for binomial responses: “skpr” and “skprJMP.” IDA released the R package “skpr” in 2017, and since then the package has supported developmental and operational test planning across dozens of programs, and has become a core piece of analytical test infrastructure for many organizations. However, some organizations have been unable to use “skpr” due to strict information technology constraints that disallow the installation of the R programming language. The new “skprJMP” add-in implements a subset of the capabilities provided by skpr entirely within JMP, allowing these organizations to calculate power for these designs.

The poster explains why Monte Carlo techniques are particularly well suited for calculating statistical power for a given design, and provides an overview of the features and capabilities of these two software packages.

Tyler Morgan-Wall

Institute for Defense Analyses

Background

As part of IDA's efforts to drive innovation by improving the capabilities, methods, and approaches used in our own work and in the test community, IDA released the R package skpr for Design of Experiments (DoE) in 2017. The skpr package provides a unified set of tools for the generation and evaluation of experimental designs. The skpr package's capabilities include performing exact power calculations using Monte Carlo methods, generating optimal designs, and providing a code-based interface to support a reproducible DoE workflow. Since skpr was released, the tool has supported developmental and operational test planning across dozens of programs, and has become a core piece of analytical test infrastructure for many organizations. Most importantly, this tool enables analysts to calculate exact power estimates for test designs with non-linear responses, which in practice are common. These include designs based on logistic regression, a capability which none of the commercial statistical software suites (JMP, Minitab, StatEase) currently provide.

Problem Statement

Some organizations have been unable to use skpr due to strict information technology constraints that disallow the installation of the R programming language, and thus they have been unable to generate accurate power estimates for designs with non-linear responses.

skprJMP

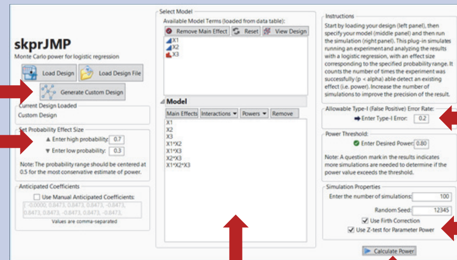
skprJMP is a software add-in for the JMP statistical software application that provides organizations using JMP with highly accurate Monte Carlo-based power calculation tools. The skprJMP add-in provides a subset of skpr's capabilities implemented entirely within JMP, enabling these organizations to calculate power for these designs within the constraints of their computing environment.

Live demo: Monte Carlo power evaluation with "skpr" and "skprJMP"

User Interface for skprJMP

Load or generate designs directly in JMP

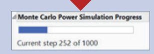
Set probability range to automatically calculate effect size and anticipated coefficients



Specify the model (main effects and interaction terms)

Run the power calculation

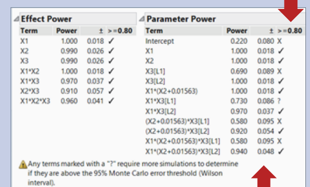
Observe simulation progress



Specify the Type-I error rate

Specify the simulation parameters

Whether the term passes the user-specified power threshold



Effect Power

Term	Power	±	>=0.80
X1	1.000	0.018	✓
X2	0.990	0.026	✓
X3	0.990	0.026	✓
X1*X2	1.000	0.018	✓
X1*X3	0.970	0.037	✓
X2*X3	0.910	0.057	✓
X1*X2*X3	0.960	0.041	✓

Parameter Power

Term	Power	±	>=0.80
Intercept	0.220	0.080	X
X1	1.000	0.018	✓
X2	1.000	0.018	✓
X3[L1]	0.690	0.089	X
X3[L2]	1.000	0.018	✓
X1*(X2+0.01563)	1.000	0.018	✓
X1*X3[L1]	0.730	0.086	?
X1*X3[L2]	0.970	0.037	✓
(X2+0.01563)*X3[L1]	0.580	0.095	X
(X2+0.01563)*X3[L2]	0.920	0.054	✓
X1*(X2+0.01563)*X3[L1]	0.580	0.095	X
X1*(X2+0.01563)*X3[L2]	0.940	0.048	✓

Any terms marked with a "?" require more simulations to determine if they are above the 95% Monte Carlo error threshold (Wilson interval).

Monte Carlo error (decreases with more simulations)

User Interface for skpr

Generate optimal design

Specify design traits (number of trials, model, and number of factors) along with factor types and levels

Generated Design

Calculate power

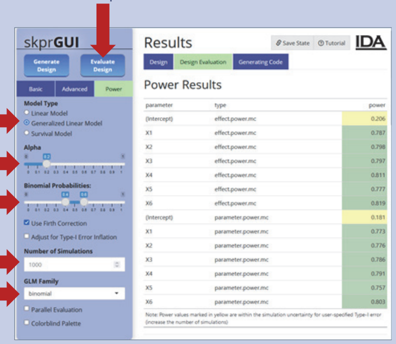
Specify power calculation type

Specify Type-I error

Specify effect size

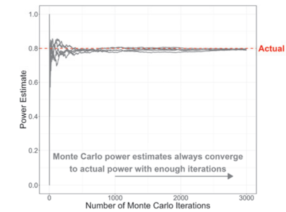
Specify number of simulations

Specify response type



R logo obtained from https://commons.wikimedia.org/wiki/File:R_logo.svg
JMP logo obtained from https://commons.wikimedia.org/wiki/File:JMP_Logo.png

Monte Carlo Power



Monte Carlo computation converges to the correct power for any sample size and test configuration

Monte Carlo techniques are a broad class of computational algorithms that use repeated random sampling to obtain numerical results. Calculating statistical power for a given design is a problem particularly well-suited for this simulation approach, for two major reasons.

First, the user can specify exactly how the data from the test is generated. This means that for a complex experimental design, a user can exactly specify how each factor influences the simulated response. Thus, the user has fine control over the effect size and the number of effects contributing to the simulated outcome.

Secondly, the user can generate the power estimate with the same methods and techniques they plan to use when analyzing the real data from the experiment. If the user wants to fit a generalized linear model, they can calculate the power using the same libraries and methods they intend to use during the data analysis.

The exact equivalence between the method used to generate the power and that used during the data analysis portion provides a strong analytic basis for the generated power values.

Acknowledgments

Thank you to Breeana Anderson, George Khoury, Jason Sheldon, and Heather Wojton for reviewing skprJMP.

All graphs, data, and designs shown on this poster were made using simulated data

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
04/16/2024

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
DATAWorks 2024: Monte Carlo Power Evaluation with skpr and skprJMP Live Demo Poster

5a. CONTRACT NUMBER
HQ0034-19-D-0001

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Tyler Morgan-Wall

5d. PROJECT NUMBER
BD-9-2299(37)

5e. TASK NUMBER
229937

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
730 East Glebe Road
Alexandria, Virginia 22305

8. PERFORMING ORGANIZATION REPORT NUMBER
3002094

3002094

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Director, Operational Test and Evaluation
1700 Defense Pentagon
Washington D.C. 20301

10. SPONSOR/MONITOR'S ACRONYM(S)
DOT&E

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release: distribution is unlimited.

13. SUPPLEMENTARY NOTES
Project Leader: Robert W Atkins

14. ABSTRACT
skprJMP is a software add-in for the JMP statistical software application that provides organizations using JMP with highly accurate Monte Carlo-based power calculation tools. The skpr R package provides a unified set of tools for the generation and evaluation of experimental designs. The skprJMP add-in implements a subset of the capabilities provided by skpr entirely within JMP, allowing organizations who rely on JMP to calculate power for these designs. This poster will be presented alongside a live demonstration of both skpr and skprJMP.

15. SUBJECT TERMS
JMP; operational testing; logistic regression; statistical power; signal-to-noise ratio; monte carlo

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified	Unclassified	Unclassified	SAR	TBD	Robert W Atkins
					19b. TELEPHONE NUMBER *(Include area code)* 703-933-6546