

# missalpha: An R package for computing bounds of Cronbach's alpha with missing data

Feng Ji<sup>1</sup> and Biying Zhou<sup>1</sup>

<sup>1</sup> Department of Applied Psychology & Human Development, University of Toronto, Toronto, Canada  
Corresponding author

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

## Software

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

Editor: [Open Journals](#)

## Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

## License

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

## Summary

Cronbach's alpha is a widely used index of internal consistency and scale reliability in psychological and educational measurement. Despite its popularity, standard implementations often fail to account for missing data appropriately, leading researchers to either use ad-hoc methods or rely on listwise deletion. In practice, this can result in biased reliability estimates.

To address this, we developed missalpha, an R package that estimates the upper and lower bounds of Cronbach's alpha under arbitrary missingness mechanisms. Our approach is inspired by the concept of *Manski bounds* ([Manski, 2003](#)), offering researchers a robust, agnostic summary of reliability when the missing data mechanism is unknown or not easily modeled. missalpha implements both exact enumeration (for small problems) and optimization-based algorithms (for larger datasets), enabling principled worst-case scenario analysis for reliability.

## Statement of Need

In applied research, Cronbach's alpha is often reported as a point estimate and compared against conventional thresholds (e.g., 0.7 or 0.8) to judge scale adequacy ([Nunnally, 1978](#)). However, in the presence of missing data, particularly when the missingness mechanism is unclear, standard point estimation may over- or under-estimate the true internal consistency of a scale.

Existing packages like psych ([Revelle, 2017](#)) and ltm ([Rizopoulos, 2007](#)) compute alpha but assume complete data or impute missing entries without evaluating uncertainty in reliability caused by missingness. To our knowledge, no current package offers a general framework to compute bounds on Cronbach's alpha that remain valid under arbitrary missing data patterns.

The missalpha package fills this gap by providing tools to: - Compute sharp lower and upper bounds of Cronbach's alpha under any missing data mechanism; - Perform sensitivity analysis via enumeration, Monte Carlo approximation, and global optimization; - Support both discrete (Likert-type) and continuous response formats.

The package is useful when researchers seek to evaluate how missing data may affect conclusions about scale reliability, and when no strong assumptions about the missingness mechanism can be made.

## Package Features

missalpha provides the following main functionalities:

- `cronbachs_alpha()`: Unified wrapper function for computing alpha bounds via different methods.

- `compute_alpha_min()` / `compute_alpha_max()`: Core functions using binary search with optimization (e.g., GA, DEoptim, nloptr) to solve for alpha bounds.
- `cronbach_alpha_enum()`: Exhaustive enumeration of all missing value configurations for exact bound computation.
- `cronbach_alpha_rough()`: Monte Carlo approximation of alpha bounds for large-scale problems.
- `display_all()`: Function to compare and visualize results across all methods.

Internally, all methods formulate the alpha bound problem as a constrained nonlinear program and apply black-box solvers from GA (Scrucca, 2013), DEoptim (Mullen et al., 2011), and nloptr (Ypma et al., 2018). These solvers identify imputations of missing entries that minimize or maximize the alpha value, thus constructing the global worst-case bounds.

## Availability

The R package missalpha is publicly available on [Github](#) (latest development version):

## Github

```
devtools::install_github("Feng-Ji-Lab/missalpha")  
library(missalpha)
```

## References

- Manski, C. F. (2003). *Partial identification of probability distributions*. Springer.
- Mullen, K. M., Ardia, D., Gil, D. L., Windover, D., & Cline, J. (2011). DEoptim: An r package for global optimization by differential evolution. *Journal of Statistical Software*, 40, 1–26.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
- Revelle, W. R. (2017). *Psych: Procedures for personality and psychological research*.
- Rizopoulos, D. (2007). Ltm: An r package for latent variable modeling and item response analysis. *Journal of Statistical Software*, 17, 1–25.
- Scrucca, L. (2013). GA: A package for genetic algorithms in r. *Journal of Statistical Software*, 53, 1–37.
- Ypma, J., Borchers, H. W., Eddelbuettel, D., & Ypma, M. J. (2018). Package “nloptr.” *R Package Version*, 1(1).