Package 'rmass2'

May 9, 2026

Type Package

Title Repeated Measures with Attrition: Sample Sizes and Power Levels for 2 Groups

Version 0.0.0.2

Depends stats (>= 3.6.2)

Description For the calculation of sample size or power in a two-group repeated measures design, accounting for attrition and accommodating a variety of correlation structures for the repeated measures; details of the method can be found in the scientific paper: Donald Hedeker, Robert D. Gibbons, Christine Waternaux (1999) <doi:10.3102/10769986024001070>.

License GPL-2 | GPL-3

Encoding UTF-8

RoxygenNote 7.2.3

Suggests knitr, rmarkdown, testthat (>= 3.0.0), utils (>= 4.3.1)

Config/testthat/edition 3

VignetteBuilder knitr

NeedsCompilation no

Author Yiheng Wei [aut] (ORCID: <https://orcid.org/0009-0007-2211-8850>), Soumya Sahu [aut, cre], Donald Hedeker [aut]

Maintainer Soumya Sahu <ssahu6@uic.edu>

Repository CRAN

Date/Publication 2023-09-20 01:00:02 UTC

Contents

rmass2 2

Index **6**

rmass2 *Repeated Measures with Attrition: Sample Sizes for 2 Groups*

Description

RMASS2 calculates the sample size or power for a two-group repeated measures design. It allows for attrition and a variety of correlation structures for the repeated measures.

Usage

```
rmass2(mode = 1, n = 2, N11 = NULL, alpha = 0.05,
nside = 2, power = NULL, ratio = 1,
attrit = 0, estype = 0, es = 0.5,
ctype = 1, corr = 0.5, sigma = rep(1, n))
```

Arguments

mode	an integer indicating mode of use. If mode = 1, one must provide the power level and the sample size is estimated. if mode = 2, one must provide the sample size and the power level is estimated. (default=1)
n	an integer giving the number of timepoints (default=2)
N11	the number of subjects in group 1 at the first timepoint. If mode = 1, N11 is not necessary and any value entered for N11 will not be used to calculate the sample size. If the mode = 2, N11 is required to estimate the power level. (default=100)
alpha	an optional numerical value giving the alpha level for statistical test. Should be a numeric value between 0 and 1. (default=0.05)
nside	an optional integer indicating 1- or 2-sided test. Should be 1 or 2. (default=2)
power	an optional numerical value giving the level of power. Should be a numeric value between 0 and 1. If mode = 1, power is required to estimate the sample size. If mode = 2, power is not necessary and any value entered for power will not be used to calculate the power level. (default=0.9)
ratio	an optional numerical value giving the ratio of sample sizes of group 1 to group 2. (default=1)
attrit	an optional vector giving the attrition rates between each timepoint. If 0 is entered, no attrition is assumed. (default=0)
estype	an optional value indicating the type of expected group differences over time. Three types are possible: 0=constant, 1=linear trend, 2=user defined. (default=0)
es	an optional numerical value or vector giving the effect size (mean difference divided by a common standard deviation) of expected group difference(s). If estype=0, the value of the expected constant group effect size is given. If estype=1, the value of the expected group effect size at the last timepoint is given. The program assumes a 0 effect size at the first timepoint that changes linearly to the effect size at the last timepoint. If estype=2, a vector of the expected group effect size at each of the n timepoints is given. (default=0.5)

ctype	an integer indicating the correlation structure of the repeated measures. Can only take values of 1, 2, or 3 (1 = all correlations equal; 2 = AR1 structure; 3 = toeplitz or banded correlation matrix). (default=1)
corr	a numerical value or vector giving the values of correlations for different correlation structures. If ctype = 1, a numerical value of the assumed equal value of all pairwise correlations is given. If ctype = 2, a numerical value of first-lag correlation is given. If ctype = 3, a vector of n-1 lag correlations is given. (default=0.5)
sigma	an optional vector of size n giving the variance for each timepoint. (default: sigma=1 at all timepoints)

Details

This package utilizes a longitudinal design that accommodates subject attrition, enabling the estimation of sample sizes based on specified levels of statistical power for significance tests. Conversely, it can also estimate the power levels for significance tests based on different sample sizes set by the user. The package allows for the modification of attrition rates, expected group differences over time, effect sizes, and covariance structure, which influence the estimated sample sizes and power levels. For more details on estimation formulas used, please refer to Hedeker, Gibbons, & Waternaux (1999).

Value

When mode = 1 (estimate the sample size), the following information will be printed row by row:

1.Correlation matrix of the repeated outcome across time, determined by the type of correlation structure (ctype) and values of correlations (corr).

2.Number of timepoints, alpha level, power levels (without attrition and with attrition), sample size ratio, input by the function's arguments.

3.Retention rates across time, calculated by the input attrition rate (attrit).

4.Effect sizes across time, determined by the type of expected group differences over time (estype) and values of effect sizes for each timepoint (es).

5.Standard deviations for each timepoint (sigma).

6.Time-related contrasts for statistical tests across time, based on the effect size. If the type of effect size is constant (estype = 0), the contrasts across time will be set as a constant unit vector. If the type of effect size is linear (estype = 1), the contrasts across time will be set as a linear unit vector. If the type of effect size is user-defined (estype = 2), the contrasts will be set as 1 for all the timepoints.

7.Mean differences across time, determined by the effect size and standard deviation. If the standard deviations vary across time, these mean differences incorporate this heterogeneity.

8.Calculated composite mean difference and composite variances (without attrition and with attrition).

9.Estimated sample sizes (without attrition and with attrition) for two groups.

When mode = 2 (estimate the level of power), the following information will be printed row by row:

1.Correlation matrix of each subject across time, determined by the type of correlation structure (ctype) and values of correlations (corr).

2.Number of timepoints, alpha level, sample size ratio, input by the function's arguments.

3.Estimated power levels (without attrition and with attrition)

4.Retention rates across time, calculated by the input attrition rate (attrit).

5.Effect sizes across time, determined by the type of expected group differences over time (estype) and values of effect sizes for each timepoint (es).

6.Standard deviations for each timepoint (sigma).

7.Time-related contrasts for statistical tests across time, based on the effect size. If the type of effect size is constant (estype = 0), the contrasts across time will be set as a constant unit vector. If the type of effect size is linear (estype = 1), the contrasts across time will be set as a linear unit vector. If the type of effect size is user-defined (estype = 2), the contrasts will be set as 1 for all the timepoints.

8.Mean differences across time, determined by the effect size and standard deviation. If the standard deviations vary across time, these mean differences incorporate this heterogeneity.

9.Calculated composite mean difference and composite variances (without attrition and with attrition).

10.Sample sizes (without attrition and with attrition) for two groups, determined by the number of subjects in group 1 at the first timepoint (N11), the ratio of sample sizes of group 1 to group 2 (ratio), and the attrition rates between each timepoint (attrit).

rmass2 can also return a list which allows users to extract some arguments. This list contains the following components:

`corr.matrix` the correlation matrix of the repeated outcome across time

`power` a dataframe of the inputted power level (mode = 1); a dataframe of the estimated power level without attrition and the estimated power level with attrition (mode = 2)

`retention.rate` a vector of the retention rates across time

`effect.size` a vector of the effect sizes across time

`stand.dev` a vector of the standard deviations across time

`contrast` a vector of the contrasts across time

`mean.diff` a vector of the mean differences across time

`composite.mean.diff` a value of the composite mean difference

`composite.var` a dataframe of the composite variance without attrition and the composite variance with attrition

`composite.effect.size` a dataframe of the composite effect size without attrition and the composite effect size with attrition

`N11` a dataframe of the estimated sample size for group 1 at the first timepoint without attrition and the estimated sample size for group 1 at the first timepoint with attrition (mode = 1); a dataframe of the calculated sample size for group 1 at the first timepoint without attrition and the calculated sample size for group 1 at the first timepoint with attrition (mode = 2)

`sample.size.group1` a dataframe of the estimated sample sizes for group 1 without attrition and the estimated sample sizes for group 1 with attrition (mode = 1); a dataframe of the calculated sample size for group 1 without attrition and the calculated sample size for group 1 with attrition (mode = 2)

`sample.size.group2` a dataframe of the estimated sample sizes for group 2 without attrition and the estimated sample sizes for group 2 with attrition (mode = 1); a dataframe of the calculated sample size for group 1 without attrition and the calculated sample size for group 1 with attrition (mode = 2)

Author(s)

This R implementation of RMASS2 was written by Yiheng Wei and Soumya Sahu. The design was based on the FORTRAN program by Donald Hedeker with the same name.

References

Hedeker, Gibbons, & Waternaux (1999). Sample size estimation for longitudinal designs with attrition. Journal of Educational and Behavioral Statistics, 24:70-93

Examples

```
#Estimate the sample size for four timepoints.
#Use the default values for other variables.
#Extract the estimated sample sizes for the first group at the first timepoint.
output <- rmass2(n = 4)
output$N11
#Estimate the sample size for two timepoints.
#Set the level of power as 0.8.
#Set the attrition rate between the 1st timepoint and the 2nd timepoint as 0.2.
#Set the correlation between the 1st timepoint and the 2nd timepoint as 0.6.
#Use the default values for other variables.
rmass2(power = 0.8, attrit = c(0.2), corr = 0.6)
#Estimate the power level with a sample size of 60 subjects in group 1 at the 1st timepoint.
#Set the attrition rate between the 1st timepoint and the 2nd timepoint as 0.2.
#Use the default values for other variables.
rmass2(mode = 2, N = 60, attrit = c(0.2))
```

Index

rmass2, 2