

Package: SWSamp (via r-universe)

October 24, 2024

Type Package

Title Computes Sample Size for a Stepped Wedge Design, using Simulation-Based Calculations

Version 0.3.2

Date 2022-11-10

Imports foreach, doParallel, iterators, parallel, Matrix

Depends lme4

Suggests INLA

Additional_repositories <https://inla.r-inla-download.org/R/stable>

Author Gianluca Baio, Rosie Leech

Maintainer Gianluca Baio <g.baio@ucl.ac.uk>

Description Provides a suite of functions for the sample size calculations and power analysis in a Stepped Wedge Trial. Contains functions for closed-form sample size calculation (based on a set of specific models) and simulation-based procedures that can extend the basic framework.

License file LICENSE

NeedsCompilation no

RoxygenNote 7.2.1

Repository <https://giabaio.r-universe.dev>

RemoteUrl <https://github.com/giabaio/SWSamp>

RemoteRef HEAD

RemoteSha a42d1778b84ff866c93ec204ab4ed73beb629c8e

Contents

SWSamp-package	2
cluster.search	3
DE.woert	5
HH.binary	6

HH.count	8
HH.normal	9
make.swt	11
sim.power	12
sw.design.mat	15
Index	17

SWSamp-package

SWSamp

Description

Sample size calculations for a Stepped Wedge Trial

Details

Package: SWSamp
 Type: Package
 Version: 0.3.2
 Date: 2021-11-10
 License: GPL2
 LazyLoad: yes

The package provides a suite of function to compute the power for a Stepped Wedge Design under different assumptions. The package can generate power based on simulations or use closed-formulae based on Hussey et Hughes

Author(s)

Gianluca Baio

Maintainer: Gianluca Baio (g.baio@ucl.ac.uk)

References

Baio, G; Copas, A; Ambler, G; Hargreaves, J; Beard, E; and Omar, RZ Sample size calculation for a stepped wedge trial. *Trials*, 16:354. Aug 2015.

Hussey M and Hughes J. Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clinical Trials*. 28(2):182-91. Epub 2006 Jul 7. Feb 2007

<code>cluster.search</code>	<i>Finds the optimum number of clusters or time points</i>
-----------------------------	--

Description

Given some inputs determines the optimal combination of clusters/time points to get a set level of power.

Usage

```
cluster.search(
  target.power = NULL,
  I = NULL,
  J = NULL,
  H = NULL,
  K,
  design = "cross-sec",
  mu = 0,
  b.trt,
  b.time = NULL,
  sigma.y = NULL,
  sigma.e = NULL,
  rho = NULL,
  sigma.a = NULL,
  rho.ind = NULL,
  sigma.v = NULL,
  n.sims = 1000,
  formula = NULL,
  family = "gaussian",
  natural.scale = TRUE,
  sig.level = 0.05,
  n.cores = NULL,
  ...
)
```

Arguments

<code>target.power</code>	The target power (eg 0.8)
<code>I</code>	A vector specifying the range in which to search for the optimal number of clusters, eg <code>I=c(1,10)</code>
<code>J</code>	Number of time points
<code>H</code>	Number of units randomised at each time point
<code>K</code>	Average size of each cluster
<code>design</code>	type of design. Can be 'cross-sec' (default) or 'cohort' (repeated measurements)

mu	baseline outcome value
b.trt	Treatment effect
b.time	Time effect
sigma.y	total standard deviation
sigma.e	individual standard deviation
rho	Intra-class correlation coefficient
sigma.a	the sd of the the cluster-level intercept (default at NULL)
rho.ind	individual-level ICC (for cohorts)
sigma.v	the sd of the cluster-level slope (by intervention, default at NULL)
n.sims	Number of simulations to be used (default=1000)
formula	Formula describing the model to be used
family	The model family to be used. Default value is 'gaussian' and other possible choices are 'binomial' or 'poisson'
natural.scale	Indicator for whether the input is passed on the natural scale or on the scale of the linear predictor. By default is set to TRUE. In the case of family='gaussian' it does not have any effect, since the link for the linear predictor is the identity. But for family='binomial' or family='poisson', the user has to specify when the input is given on the logit or log scale
sig.level	Significance level (default=0.05)
n.cores	Specifies the number of processors to be used for the computation (default=NULL, which means that R will try and figure out)
...	Additional arguments

Value

Optimum_I	The value of the optimal number of clusters
power	The estimated power in correspondence of the optimal I
time2run	Computational time

Author(s)

Rosie Leach

References

- Baio, G; Copas, A; Ambler, G; Hargreaves, J; Beard, E; and Omar, RZ Sample size calculation for a stepped wedge trial. *Trials*, 16:354. Aug 2015.
- Hussey M and Hughes J. Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clinical Trials*. 28(2):182-91. Epub 2006 Jul 7. Feb 2007

Examples

```
#cluster.search(I=c(4,10),target.power=.8,J=6,K=30,mu=1.5,b.trt=.8,rho=0,
#family="poisson",n.sims=10)
```

DE.woert

*Computes the Design Effect for a Stepped Wedge Trial***Description**

Sample size calculations for a SWT using a cross-sectional design. This is based on (the correct version) of Woertman et al (2013), as described in Baio et al (2015).

Usage

```
DE.woert(
  outcome = "cont",
  input,
  K,
  J,
  B = 1,
  T = 1,
  rho,
  sig.level = 0.05,
  power = 0.8
)
```

Arguments

outcome	String. Type of outcome. Options are cont, bin or count
input	input = a list containing the arguments. This differs depending on the type of outcome, as follows: - continuous outcome: 1) delta (treatment effect) 2) sd (standard deviation) - binary outcome: 1) p1 (baseline probability of outcome) 2) either p2 (treatment probability of outcome), or OR (treatment effect as OR) - count outcome: 1) r1 (baseline rate of outcome) 2) either r2 (treatment rate of outcome), or RR (treatment effect as RR)
K	average cluster size
J	number of time points (excluding baseline)
B	number of baseline measurement times
T	number of measurement times during each crossover
rho	ICC
sig.level	significance level (default = 0.05)
power	Power (default = 0.8)

Value

n.cls.swt	Number of clusters required to reach the pre-specified power with the given significance level.
n.pts	The total number of participants required.

DE.woert	The resulting Design Effect.
CF	The resulting Correction Factor.
n.rct	The original individual RCT sample required to reach the pre-specified power with the given significance level.

Author(s)

Gianluca Baio

References

Baio, G; Copas, A; Ambler, G; Hargreaves, J; Beard, E; and Omar, RZ Sample size calculation for a stepped wedge trial. *Trials*, 16:354. Aug 2015.

Examples

```
# Continuous outcome
input <- list(delta=-0.3875,sd=1.55)
K <- 20
J <- 5
rho <- .2
DE.woert(input=input,K=K,J=J,rho=rho)
#
# Binary outcome
input <- list(OR=.53,p1=.26)
DE.woert(outcome="bin",input=input,K=K,J=J,rho=rho)
#
# Count outcome
input <- list(RR=.8,r1=1.5)
DE.woert(outcome="count",input=input,K=K,J=J,rho=rho)
```

HH.binary

*Power calculation for binary outcome based on analytic formula of
Hussey and Hughes*

Description

Sample size calculations for binary outcomes based on the formula provided by Hussey and Hughes (2007)

Usage

```
HH.binary(  
  p1,  
  OR,  
  I,  
  J,
```

```

    K,
    rho = 0,
    sig.level = 0.05,
    which.var = "within",
    X = NULL
  )

```

Arguments

p1	Baseline probability of the outcome (for the controls)
OR	Value of the expected Odds Ratio (for the intervention vs control)
I	Number of clusters
J	Number of time points
K	Average size of each cluster
rho	Intra-class correlation coefficient (default=0)
sig.level	Significance level (default=0.05)
which.var	String character specifying which variance to be considered (options are the default value 'within' or 'total')
X	A design matrix for the stepped wedge design, indicating the time at which each of the clusters should switch the active intervention. By default is NULL and automatically computed, but can be passed as an extra argument as a user-defined matrix with I rows and (J+1) columns

Value

power	The resulting power
sigma.y	The estimated total (marginal) sd for the outcome
sigma.e	The estimated residual sd
sigma.a	The resulting cluster-level sd
setting	A list including the following values: - n.clusters = The number of clusters - n.time.points = The number of 'active' time points - avg.cluster.size = The average cluster size - design.matrix = The design matrix for the SWT under consideration

Author(s)

Gianluca Baio

References

- Baio, G; Copas, A; Ambler, G; Hargreaves, J; Beard, E; and Omar, RZ Sample size calculation for a stepped wedge trial. *Trials*, 16:354. Aug 2015.
- Hussey M and Hughes J. Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clinical Trials*. 28(2):182-91. Epub 2006 Jul 7. Feb 2007

Examples

```
HH.binary(p1=.26,OR=.55,I=10,J=5,K=20,rho=.2)
```

HH.count	<i>Power calculation for count outcome based on analytic formula of Hussey and Hughes</i>
----------	---

Description

Sample size calculations for count outcomes based on the formula provided by Hussey and Hughes (2007)

Usage

```
HH.count(
  lambda1,
  RR,
  I,
  J,
  K,
  rho = 0,
  sig.level = 0.05,
  which.var = "within",
  X = NULL
)
```

Arguments

lambda1	Baseline value for the rate at which the outcome occurs
RR	Relative risk (of the intervention vs the control)
I	Number of clusters
J	Number of time points
K	Average size of each cluster
rho	Intra-class correlation coefficient (default=0)
sig.level	Significance level (default=0.05)
which.var	String character specifying which variance to be considered (options are the default value 'within' or 'total')
X	A design matrix for the stepped wedge design, indicating the time at which each of the clusters should switch the active intervention. By default is NULL and automatically computed, but can be passed as an extra argument as a user-defined matrix with I rows and (J+1) columns

Value

power	The resulting power
sigma.y	The estimated total (marginal) sd for the outcome
sigma.e	The estimated residual sd
sigma.a	The resulting cluster-level sd
setting	A list including the following values: - n.clusters = The number of clusters - n.time.points = The number of 'active' time points - avg.cluster.size = The average cluster size - design.matrix = The design matrix for the SWT under consideration

Author(s)

Gianluca Baio

References

Baio, G; Copas, A; Ambler, G; Hargreaves, J; Beard, E; and Omar, RZ Sample size calculation for a stepped wedge trial. *Trials*, 16:354. Aug 2015.

Hussey M and Hughes J. Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clinical Trials*. 28(2):182-91. Epub 2006 Jul 7. Feb 2007

Examples

```
HH.count(lambda1=1.55, RR=.87, I=10, J=5, K=20, rho=.2)
```

HH.normal	<i>Power calculation for normal outcome based on analytic formula of Hussey and Hughes</i>
-----------	--

Description

Sample size calculations for normal outcomes based on the formula provided by Hussey and Hughes (2007)

Usage

```
HH.normal(  
  mu,  
  b.trt,  
  sigma,  
  I,  
  J,  
  K,  
  rho = 0,  
  sig.level = 0.05,
```

```

    which.var = "within",
    X = NULL
  )

```

Arguments

mu	Mean value of the outcome for the controls
b.trt	Treatment effect against controls
sigma	Value of the standard deviation (if which.var='within' then it's assumed to be the residual sd. If which.var='total', then it's assumed to be the total sd)
I	Number of clusters
J	Number of time points
K	Average size of each cluster
rho	Intra-class correlation coefficient (default=0)
sig.level	Significance level (default=0.05)
which.var	String character specifying which variance to be considered (options are the default value 'within' or 'total')
X	A design matrix for the stepped wedge design, indicating the time at which each of the clusters should switch the active intervention. By default is NULL and automatically computed, but can be passed as an extra argument as a user-defined matrix with I rows and (J+1) columns

Value

power	The resulting power
sigma.y	The estimated total (marginal) sd for the outcome
sigma.e	The estimated residual sd
sigma.a	The resulting cluster-level sd
setting	A list including the following values: - n.clusters = The number of clusters - n.time.points = The number of 'active' time points - avg.cluster.size = The average cluster size - design.matrix = The design matrix for the SWT under consideration

Author(s)

Gianluca Baio

References

- Baio, G; Copas, A; Ambler, G; Hargreaves, J; Beard, E; and Omar, RZ Sample size calculation for a stepped wedge trial. *Trials*, 16:354. Aug 2015.
- Hussey M and Hughes J. Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clinical Trials*. 28(2):182-91. Epub 2006 Jul 7. Feb 2007

Examples

```
HH.normal(mu=.3,b.trt=-.3875,I=10,J=5,K=20,rho=.2,sigma=1.55)
```

make.swt

Simulates a 'virtual' Stepped Wedge trial

Description

Simulates trial data for a SWT with normally distributed outcome

Usage

```
make.swt(
  I = NULL,
  J = NULL,
  H = NULL,
  K,
  design = "cross-sec",
  mu = NULL,
  b.trt,
  b.time = NULL,
  sigma.y = NULL,
  sigma.e = NULL,
  rho,
  sigma.a = NULL,
  rho.ind = NULL,
  sigma.v = NULL,
  X = NULL,
  family = "gaussian",
  natural.scale = TRUE
)
```

Arguments

I	Number of clusters
J	Number of time points
H	Number of units randomised at each time point
K	Average size of each cluster
design	type of design. Can be 'cross-sec' (default) or 'cohort' (repeated measurements)
mu	baseline outcome value
b.trt	Treatment effect
b.time	Time effect

sigma.y	total standard deviation
sigma.e	individual standard deviation
rho	Intra-class correlation coefficient
sigma.a	the sd of the the cluster-level intercept (default at NULL)
rho.ind	individual-level ICC (for cohorts)
sigma.v	the sd of the cluster-level slope (by intervention, default at NULL)
X	A design matrix for the SWT. Default at NULL (will be computed automatically)
family	The model family to be used. Default value is 'gaussian' and other possible choices are 'binomial' or 'poisson'
natural.scale	Indicator for whether the input is passed on the natural scale or on the scale of the linear predictor. By default is set to TRUE. In the case of family='gaussian' it does not have any effect, since the link for the linear predictor is the identity. But for family='binomial' or family='poisson', the user has to specify when the input is given on the logit or log scale

Value

data	A data frame containing the resulting simulated dataset
------	---

Author(s)

Gianluca Baio, Rosie Leach

References

Baio, G; Copas, A; Ambler, G; Hargreaves, J; Beard, E; and Omar, RZ Sample size calculation for a stepped wedge trial. *Trials*, 16:354. Aug 2015.

See Also

See Also [sim.power](#)

 sim.power

Power calculations based on a simulation approach

Description

Simulation-based power calculations for a SWT with normally distributed outcome

Usage

```

sim.power(
  I,
  J,
  H = NULL,
  K,
  design = "cross-sec",
  mu = 0,
  b.trt,
  b.time = NULL,
  sigma.y = NULL,
  sigma.e = NULL,
  rho = NULL,
  sigma.a = NULL,
  rho.ind = NULL,
  sigma.v = NULL,
  n.sims = 1000,
  formula = NULL,
  family = "gaussian",
  natural.scale = TRUE,
  sig.level = 0.05,
  n.cores = NULL,
  method = "lme",
  plot = FALSE,
  ...
)

```

Arguments

I	Number of clusters
J	Number of time points
H	Number of units randomised at each time point
K	Average size of each cluster
design	type of design. Can be 'cross-sec' (default) or 'cohort' (repeated measurements)
mu	baseline outcome value
b.trt	Treatment effect
b.time	Time effect
sigma.y	total standard deviation
sigma.e	individual standard deviation
rho	Intra-class correlation coefficient
sigma.a	the sd of the the cluster-level intercept (default at NULL)
rho.ind	individual-level ICC (for cohorts)
sigma.v	the sd of the cluster-level slope (by intervention, default at NULL)

n.sims	Number of simulations to be used (default=1000)
formula	Formula describing the model to be used
family	The model family to be used. Default value is 'gaussian' and other possible choices are 'binomial' or 'poisson'
natural.scale	Indicator for whether the input is passed on the natural scale or on the scale of the linear predictor. By default is set to TRUE. In the case of family='gaussian' it does not have any effect, since the link for the linear predictor is the identity. But for family='binomial' or family='poisson', the user has to specify when the input is given on the logit or log scale
sig.level	Significance level (default=0.05)
n.cores	Specifies the number of processors to be used for the computation (default=NULL, which means that R will try and figure out)
method	A string specifying the method to be used for the calculation. The default value is lme, indicating a standard frequentist analysis, based on (generalised) linear model and including structured (random) effects, when necessary. An alternative version is framed in a Bayesian setting and uses Integrated Nested Laplace Approximation (INLA) to analyse the data and obtain the relevant estimates for the posterior distributions of the model parameters. This can be performed by setting method='inla'
plot	Shows a plot of the moving average of the resulting power after 10 of the estimate towards some common value. The default is FALSE, in which case a graph is not shown.
...	Additional optional arguments. The user can specify a function named data, which defines the simulation of the 'virtual trial data'. These can be in any given form, with the only constraint that it should return a data frame containing the relevant variables. The function data can have no arguments at all, but this can be relaxed and there can be suitable inputs to this function. In this case, the user also needs to specify a list inpts including all the values for the arguments to be used with the user-defined function data. When using user-defined data generating processes, the user must be also pass a relevant formula, depending on what the model used (for both generation of the dataset and analysis) is. Another additional argument that can be passed to the call to sim.sw.cont is treatment, a string specifying the name of the treatment variable (if not present, SWSamp assumes that this is exactly 'treatment').

Value

power	The resulting estimated power, for the given configuration. If the model does not include random effects, this is based on the p-value computed by lm, which is used to analyse the simulated datasets. If the model does include random effects (which is the case for a SWT), then SWSample assesses whether the 'true' effect is correctly detected by computing the (1-alpha) whether it is entirely above or below 0. This is because it is difficult to assess the correct degrees of freedom of the resulting (linear) mixed model. The p-value could be computed using the Satterthwaite approximation, or by using a rougher Normal approximation,
-------	---

	but in line with suggestions by Pinheiro, J. C., and D. M. Bates. 2000. Mixed-effects models in S and S-PLUS. Springer, New York, we sidestep the problem by focussing on estimation, rather than hypothesis testing for this.
time2run	Running time, in seconds
ci.power	Estimated 95% confidence interval for the power - based on normal approximation
theta	Estimated treatment effect with standard error
rnd.eff.sd	Estimated variance components
setting	A list summarising the assumptions in terms of number of clusters, time points, type of model, formula used

Author(s)

Gianluca Baio

References

Baio, G; Copas, A; Ambler, G; Hargreaves, J; Beard, E; and Omar, RZ Sample size calculation for a stepped wedge trial. *Trials*, 16:354. Aug 2015.

Examples

```
mu1=0.3
b.trt=-0.3875
sigma.e=1.55
J=5
K=20
sig.level=0.05
n.sims=10
rho=0.5
pow.cont <- sim.power(I=14, J=J, H=NULL, K=K, rho=rho, mu=mu1, sigma.e=sigma.e, b.trt=b.trt,
                      formula=NULL, n.sims=n.sims, sig.level=sig.level, n.cores=2)
pow.cont$power
pow.cont$ci
pow.cont$time2run.sec
pow.cont$rnd.eff.sd^2
```

sw.design.mat

*Creates a design matrix for a Stepped Wedge Trial***Description**

Constructs a basic SWT design matrix

Usage

```
sw.design.mat(I, J, H = NULL)
```

Arguments

I	Number of clusters
J	Number of time points
H	Number of units randomised at each time point

Value

Returns a design matrix X

Author(s)

Gianluca Baio

References

Baio, G; Copas, A; Ambler, G; Hargreaves, J; Beard, E; and Omar, RZ Sample size calculation for a stepped wedge trial. *Trials*, 16:354. Aug 2015.

Index

- * **Continuous**
 - make.swt, 11
 - sim.power, 12
- * **Design**
 - DE.woert, 5
 - SWSamp-package, 2
- * **Effect**
 - DE.woert, 5
- * **Hughes**
 - HH.binary, 6
 - HH.count, 8
 - HH.normal, 9
- * **Hussey**
 - HH.binary, 6
 - HH.count, 8
 - HH.normal, 9
- * **Sample**
 - SWSamp-package, 2
- * **Simulation-based**
 - sim.power, 12
- * **Stepped**
 - cluster.search, 3
 - make.swt, 11
 - sim.power, 12
 - SWSamp-package, 2
- * **Trial**
 - make.swt, 11
- * **Wedge**
 - SWSamp-package, 2
- * **and**
 - HH.binary, 6
 - HH.count, 8
 - HH.normal, 9
- * **calculations**
 - sim.power, 12
 - SWSamp-package, 2
- * **data**
 - make.swt, 11
- * **design**
 - cluster.search, 3
 - make.swt, 11
 - sim.power, 12
- * **formula**
 - HH.binary, 6
 - HH.count, 8
 - HH.normal, 9
- * **outcome**
 - make.swt, 11
 - sim.power, 12
- * **sample**
 - sim.power, 12
- * **simulation**
 - make.swt, 11
- * **size**
 - sim.power, 12
 - SWSamp-package, 2
- * **wedge**
 - cluster.search, 3
 - make.swt, 11
 - sim.power, 12
- cluster.search, 3
- DE.woert, 5
- HH.binary, 6
- HH.count, 8
- HH.normal, 9
- make.swt, 11
- sim.power, 12, 12
- sw.design.mat, 15
- SWSamp (SWSamp-package), 2
- SWSamp-package, 2