Package 'Power2Stage'

April 20, 2019

11pm 20, 2017
Encoding UTF-8
Type Package
Title Power and Sample-Size Distribution of 2-Stage Bioequivalence Studies
Version 0.5.2
Date 2019-04-21
Description Contains functions to obtain the operational characteristics of bioequivalence studies with 2-stage designs (TSD) via simulations.
Imports PowerTOST, stats, mvtnorm, cubature
Suggests graphics
ByteCompile yes
<pre>URL https://github.com/Detlew/Power2Stage</pre>
<pre>BugReports https://github.com/Detlew/Power2Stage/issues</pre>
License GPL (>= 2)
LazyLoad yes
LazyData yes
NeedsCompilation no
Author Detlew Labes [aut, cre], Benjamin Lang [aut, ctb], Helmut Schütz [aut, ctb]
Maintainer Detlew Labes < Detlew Labes@gmx.de>
Repository CRAN
Date/Publication 2019-04-20 19:30:03 UTC
R topics documented:
final.tsd.in

2 final.tsd.in

final.tsd.in		ysis a nverse			_	of 2	-sta	age	22	ι2 ι	cro.	SSC	ove	r d	lesi	ign	b	ase	ed o	on
Index																				34
	sampleN2.TOST	 	 	 																31
	power.tsd.ssr	 	 	 																28
	power.tsd.p	 	 	 																25
	power.tsd.KM	 	 	 																23
	power.tsd.in	 	 	 																18
	power.tsd.GS	 	 	 																16
	power.tsd.fC	 	 	 																12

Description

Following the design scheme according to power.tsd. in the function performs the analysis after the second stage has been performed.

Usage

Arguments

alpha	If one element is given, the overall one-sided significance level (not the adjusted level for stage 2). If two elements are given, the adjusted one-sided alpha levels for stage 1 and stage 2, respectively. If missing, defaults to 0.05.
weight	Pre-defined weight(s) of stage 1. Note that using the notation from Maurer et al weight corresponds to information fraction, other literature may refer to sqrt(weight) as being the weight. weight must either contain one element (in case of max.comb.test = FALSE) or two elements (in case of max.comb.test = TRUE). If missing, defaults to 0.5 for max.comb.test = FALSE and to c(0.5, 0.25) for max.comb.test = TRUE.
max.comb.test	Logical; if TRUE (default) the maximum combination test will be used, otherwise the standard combination test.
GMR1	Observed ratio of geometric means (T/R) of stage 1 data (use e.g., 0.95 for 95%).
CV1	Observed coefficient of variation of the intra-subject variability of stage 1 (use $e.g.$, 0.3 for 30%).
n1	Sample size of stage 1.
df1	Optional; Error degrees of freedom of stage 1 that can be specified in addition to n1.

final.tsd.in 3

SEM1	Optional; Standard error of the difference of means of stage 1 that can be specified in addition to CV1. Must be on additive scale (i.e. usually log-scale).
GMR2	Observed ratio of geometric means (T/R) of (only) stage 2 data (use $e.g.$, 0.95 for 95%).
CV2	Observed coefficient of variation of the intra-subject variability of (only) stage 2 (use $e.g.$, 0.3 for 30%).
n2	Sample size of stage 2.
df2	Optional; Error degrees of freedom of (only) stage 2 that can be specified in addition to n2.
SEM2	Optional; Standard error of the difference of means of (only) stage 2 that can be specified in addition to CV2. Must be on additive scale (i.e. usually log-scale).
theta1	Lower bioequivalence limit. Defaults to 0.8.
theta2	Upper bioequivalence limit. Defaults to 1.25.

Details

The observed values GMR1, CV1, n1 must be obtained using data from stage 1 only, and GMR2, CV2, n2 must be obtained using data from stage 2 only. This may be done via the usual ANOVA approach.

The optional arguments df1, SEM1, df2 and SEM2 require a somewhat advanced knowledge (provided in the raw output from for example the software SAS, or may be obtained via emmeans::emmeans). However, it has the advantage that if there were missing data the exact degrees of freedom and standard error of the difference can be used, the former possibly being non-integer valued (e.g. if the Kenward-Roger method was used).

Value

Returns an object of class "evaltsd" with all the input arguments and results as components. As part of the input arguments a component cval is also presented, containing the critical values for stage 1 and 2 according to the input based on alpha, weight and max.comb.test. The class "evaltsd" has an S3 print method.

The results are in the components:

z1	Combination test statistic for first null hypothesis (standard combination test statistic in case of max.comb.test = FALSE or maximum combination test statistic in case of max.comb.test = TRUE)
z2	Combination test statistic for second null hypothesis (standard combination test statistic in case of max.comb.test = FALSE or maximum combination test statistic in case of max.comb.test = TRUE)
RCI	(Exact) repeated confidence interval for stage 2.
MEUE	Median unbiased point estimate as estimate for the final geometric mean ratio after stage 2.
stop_BE	Logical, indicating whether BE can be concluded after stage 2 or not.

Author(s)

B. Lang

References

König F, Wolfsegger M, Jaki T, Schütz H, Wassmer G.

Adaptive two-stage bioequivalence trials with early stopping and sample size re-estimation.

Vienna: 2014; 35th Annual Conference of the International Society for Clinical Biostatistics. Poster P1.2.88

doi: 10.13140/RG.2.1.5190.0967.

Patterson SD, Jones B. Bioequivalence and Statistics in Clinical Pharmacology.

Boca Raton: CRC Press; $2^{n\bar{d}}$ edition 2017.

Maurer W, Jones B, Chen Y. Controlling the type 1 error rate in two-stage sequential designs when testing for average bioequivalence.

Stat Med. 2018;1–21. doi: 10.1002/sim.7614.

Wassmer G, Brannath W. *Group Sequential and Confirmatory Adaptive Designs in Clinical Trials*. Springer 2016. doi: 10.1007/9783319325620.

See Also

```
power.tsd.in, interim.tsd.in
```

Examples

```
# Example from Maurer et al.
final.tsd.in(GMR1 = \exp(0.0424), CV1 = 0.3682, n1 = 20,
GMR2 = \exp(-0.0134), CV2 = 0.3644, n2 = 36)
```

interim.tsd.in

Interim analysis of first stage data of 2-stage 2x2 crossover designs based on the Inverse Normal method

Description

Following the design scheme according to power.tsd.in the function performs the interim analysis of the first stage data.

Usage

Arguments

alpha If one element is given, the overall one-sided significance level (not the adjusted

level for stage 1). In this case the adjusted alpha levels will be calcualted internally. If two elements are given, the argument refers to the two adjusted

one-sided alpha levels for stage 1 and stage 2, respectively.

If missing, defaults to 0.05.

weight Pre-defined weight(s) of stage 1, see 'Details' for more information. Note that

using the notation from Maurer et al, weight corresponds to information fraction, other literature may refer to sqrt(weight) as being the weight. weight must either contain one element (in case of max.comb.test = FALSE) or two elements (in

case of max.comb.test = TRUE).

If missing, defaults to 0.5 for max.comb.test = FALSE and to $c(0.5,\,0.25)$ for

max.comb.test = TRUE.

max.comb.test Logical; if TRUE (default) the maximum combination test will be used, other-

wise the standard combination test.

targetpower Desired (overall) target power to declare BE at the end of the trial.

GMR1 Observed ratio of geometric means (T/R) of stage 1 data (use *e.g.*, 0.95 for 95%).

n1 Sample size of stage 1.

CV1 Observed coefficient of variation of the intra-subject variability of stage 1 (use

e.g., 0.3 for 30%).

df1 Optional; Error degrees of freedom of stage 1 that can be specified in addition

to n1.

SEM1 Optional; Standard error of the difference of means of stage 1 that can be speci-

fied in addition to CV1. Must be on additive scale (i.e. usually log-scale).

theta1 Lower bioequivalence limit. Defaults to 0.8.

theta2 Upper bioequivalence limit. Defaults to 1.25.

GMR Assumed ratio of geometric means (T/R) to be used in power calculation for

stage 1 and sample size re-estimation for stage 2. If missing, defaults to 0.95.

usePE If TRUE the sample size re-estimation is done with the observed point estimate

(PE) of the treatment difference in stage 1.

Defaults to FALSE.

Note: The futility inspection via the power of stage 1 is always done with the

planning value GMR.

min.n2 Minimum sample size of stage 2. Defaults to 4.

If the sample size re-estimation step gives a sample size for stage 2 less than

min.n2, then min.n2 will be used for stage 2.

max.n Maximum overall sample size stage 1 + stage 2.

This is *not* a futility criterion regarding the maximum sample size! If max.n is set to a finite value and the sample size re-estimation gives a sample size for stage 2 (n2) such that n1 + n2 > max.n, then the sample size for stage 2 will be

set to n2 = max.n - n1.

Defaults to Inf, i.e., no constraint on the re-estimated sample size.

fCpower

Threshold for power monitoring step to decide on futility for cases where BE has not been achieved after stage 1: If BE has not been achieved after stage 1 and the power for stage 1 is greater than or equal to fCpower, then the study will be considered a failure.

See 'Details' for more information on the choice of fCpower.

fCrit

Futility criterion to use: "No" (no futility criterion regarding observed point estimate, confidence interval and maximum sample size), "PE" (observed point estimate of the geometric mean ratio from stage 1), "CI" (90% confidence interval of the geometric mean ratio from stage 1), "Nmax" (overall maximum sample size); or a combination thereof (concatenate abbreviations). Defaults to "CI".

fClower

Lower futility limit for the PE or CI of stage 1.

If the PE or CI is completely outside of fClower ... fCupper the study is to be stopped due to futility (not BE).

May be missing. If "PE" or "CI" is specified within fCrit, the default will be set to 0.8 for fCrit = "PE" or 0.95 for fCrit = "CI". If neither "PE" nor "CI" is specified within fCrit, there will be no futility constraint regarding point estimate or confidence interval from stage 1 (regardless of any specification of fClower and/or fCupper).

fCupper

Upper futility limit for the PE or CI of stage 1.

Analogous to fClower: Will be set to 1/fClower if missing.

fCNmax

Futility criterion regarding maximum sample size. If the determined sample size for stage 2 (n2) is such that n1 + n2 > fCNmax, the study will not continue to stage 2 and stopped due to futility (not BE).

If "Nmax" is specified within fCrit and argument fCNmax is missing, the value will be set to fCNmax = 4*n1. If "Nmax" is not specified within fCrit, then there will be no futility constraint regarding maximum sample size (regardless of any specification of fCNmax).

ssr.conditional

Method for sample size re-estimation step: "no" does not use conditional error rates nor the estimated conditional target power for the second stage, "error" uses conditional error rates for the second stage, and "error_power" uses both conditional error rates and the estimated conditional target power for the second stage.

Defaults to "error_power".

See also 'Details'.

pmethod

Power calculation method, also to be used in the sample size estimation for stage 2.

Implemented are "nct" (approximate calculations via non-central *t*-distribution, "exact" (exact calculations via Owen's Q), and "shifted" (approximate calculation via shifted central *t*-distribution like in the paper of Potvin *et al.*)

In contrast to power.tsd.in the default value here is "exact".

Details

The observed values of stage 1 (e.g. GMR1, n1, CV1) may be obtained based on the first stage data via the usual ANOVA approach.

CIOWCI

The optional arguments df1 and SEM1 require a somewhat advanced knowledge (provided in the raw output from for example the software SAS, or may be obtained via emmeans::emmeans). However, it has the advantage that if there were missing data the exact degrees of freedom and standard error of the difference can be used, the former possibly being non-integer valued (e.g. if the Kenward-Roger method was used).

The weight argument always refers to the first weight of a pair of weights. For example, in case of $\max.$ comb. test = FALSE the standard combination test requires two weights (w, 1-w) but only the first one, w, is required as input argument here because the second weight is automatically specified once the first is given. Similarly for $\max.$ comb. test = TRUE, w and w* need to be specified, which in turn define the two pairs of weights (w, 1-w) and (w*, 1-w*).

If ssr.conditional = "error_power", the design scheme generally calculates the estimated conditional target power of the second stage and uses this value as desired target power in the sample size reestimation process. If fCpower > targetpower, then the conditional target power may actually be negative. This does not seem sensible. Therefore, for such cases the desired target power for the sample size re-calculation will be set to targetpower.

Value

Returns an object of class "evaltsd" with all the input arguments and results as components. As part of the input arguments a component cval is also presented, containing the critical values for stage 1 and 2 according to the input based on alpha, weight and max.comb.test. The class "evaltsd" has an S3 print method.

The results are in the components:

p11	Observed p-value for first hypothesis.
p12	Observed p-value for second hypothesis.
z1	z statistic value for first null hypothesis.
z2	z statistic value for second null hypothesis.
RCI	(Exact) repeated confidence interval for stage 1.
futility	Three dimensional vector with either 0 or 1. The first component represents futility due to Power of first stage $>$ fCpower, the second futility due to CI (or PE) outside of fClower fCupper, the third futility due to n1 + n2 > fCNmax. Note that the futility rules can be applied in a non-binding manner.
CI90	90% Confidence interval for observed ratio of geometric means from stage 1. If fCrit $!=$ "CI" result will be NULL.
Power Stage 1	Calculated power of stage 1.
stop_s1	Logical, indicating whether to stop after stage 1 (due to BE or due to futility).
stop_fut	Logical, indicating whether study is recommended to be stopped after stage 1 due to futility.
stop_BE	Logical, indicating whether BE could be concluded after stage 1 or not (regardless of any futility criterion).

n2 Required (total) sample size for stage 2 (will be zero if BE has been shown after

stage 1).

alpha_ssr Only applicable if BE has not been shown after stage 1. Contains alpha values for

the two hypotheses required for sample size re-calculation. If ssr.conditional = "no" the result is equal to alpha, otherwise it contains the conditional error rates for the standard combination test (in case of max.comb.test = FALSE) or maxi-

 $mum\ combination\ test\ (in\ case\ of\ max.comb.test\ =\ TRUE).$

GMR_ssr Only applicable if BE has not been shown after stage 1. Contains the geometric

mean ratio used for sample size re-calculation (accounts for adaptive planning

step).

targetpower_ssr

Only applicable if BE has not been shown after stage 1. Contains the target power used for the sample size re-calculation (see also 'Details').

Author(s)

B. Lang

References

König F, Wolfsegger M, Jaki T, Schütz H, Wassmer G.

Adaptive two-stage bioequivalence trials with early stopping and sample size re-estimation.

Vienna: 2014; 35th Annual Conference of the International Society for Clinical Biostatistics. Poster P1.2.88

doi: 10.13140/RG.2.1.5190.0967.

Patterson SD, Jones B. Bioequivalence and Statistics in Clinical Pharmacology.

Boca Raton: CRC Press; $2^{n\bar{d}}$ edition 2017.

Maurer W, Jones B, Chen Y. Controlling the type 1 error rate in two-stage sequential designs when testing for average bioequivalence.

Stat Med. 2018;1-21. doi: 10.1002/sim.7614.

Wassmer G, Brannath W. *Group Sequential and Confirmatory Adaptive Designs in Clinical Trials*. Springer 2016. doi: 10.1007/9783319325620.

See Also

```
power.tsd.in, final.tsd.in
```

Examples

```
# Example from Maurer et al interim.tsd.in(GMR1 = exp(0.0424), CV1 = 0.3682, n1 = 20, max.n = 4000)
```

power.tsd 9

power.tsd	Power calculation of adaptive 2-stage BE studies in a 2x2 crossover design

Description

This function calculates the 'empiric' power of 2-stage BE studies according to Potvin *et al.* via simulations. The Potvin methods are modified to include a futility criterion Nmax and to allow the sample size estimation step to be done with the point estimate (PE) and MSE (calculated from CV) of stage 1.

Usage

```
power.tsd(method = c("B", "C", "B0"), alpha0 = 0.05, alpha = c(0.0294, 0.0294), n1, GMR, CV, targetpower = 0.8, pmethod = c("nct", "exact", "shifted"), usePE = FALSE, Nmax = Inf, min.n2 = 0, theta0, theta1, theta2, npct = c(0.05, 0.5, 0.95), nsims, setseed = TRUE, details = FALSE)
```

Arguments

rguments	
method	Decision schemes according to Potvin <i>et.al.</i> (defaults to "B"). Potvin's 'Method D' can be obtained by choosing "C" but setting alpha=c(0.028, 0.028). Montague's 'Method D' can be obtained by choosing "C" but setting alpha=c(0.028, 0.028) and GMR=0.9. method="B0" uses the decision scheme of Zheng <i>et al.</i> MSDBE (modified sequential design for BE studies) which differs from B in case of different alphas w.r.t. power monitoring and BE decision in case of power >= target power.
alpha0	Alpha value for the first step(s) in Potvin "C", the power inspection and BE decision if power > targetpower. Defaults to 0.05. Only observed if method="C".
alpha	Vector (two elements) of the nominal alphas for the two stages. Defaults to Pocock's setting alpha=c(0.0294, 0.0294). Common values together with other arguments are: rep(0.0294, 2): Potvin <i>et al.</i> 'Method B' rep(0.0269, 2): Fulgsang 'Method C/D' (method="C", GMR=0.9, targetpower=0.9) rep(0.0274, 2): Fuglsang 'Method C/D' (method="C", targetpower=0.9) rep(0.0280, 2): Montague <i>et al.</i> 'Method D' (method="C", GMR=0.9) rep(0.0284, 2): Fulgsang 'Method B' (GMR=0.9, targetpower=0.9) rep(0.0304, 2): Kieser & Rauch c(0.01, 0.04): Zheng <i>et al.</i> 'MSDBE' (method="B0")
n1	Sample size of stage 1.
GMR	Ratio T/R to be used in decision scheme (power calculations in stage 1 and sample size estimation for stage 2).
CV	Coefficient of variation of the intra-subject variability (use e.g., 0.3 for 30%).
targetpower	Power threshold in the power monitoring steps and power to achieve in the sam-

ple size estimation step.

10 power.tsd

pmethod Power calculation method, also to be used in the sample size estimation for stage

2.

Implemented are "nct" (approximate calculations via non-central *t*-distribution, "exact" (exact calculations via Owen's Q), and "shifted" (approximate calculations via Owen's Q), and "shifted" (approximate calculations) and "shifted" (approx

tion via shifted central *t*-distribution like in the paper of Potvin *et al*.

Defaults to "nct" as a reasonable compromise between speed and accuracy in

the sample size estimation step.

usePE If TRUE the sample size estimation step is done with MSE **and** PE of stage 1.

Defaults to FALSE i.e., the sample size is estimated with GMR and MSE (calcu-

lated from CV) of stage 1 analogous to Potvin et. al.

NB: The power inspection steps in the Potvin methods are always done with the

GMR argument and MSE (CV) of stage 1.

Nmax Futility criterion. If set to a finite value, all studies simulated in which a sample

size >Nmax is obtained will be regarded as BE=FAIL.

Set this argument to Inf, the default, to work without that futility criterion.

min.n2 Minimum sample size of stage 2. Defaults to zero.

If the sample size estimation step gives N < n1 the sample size for stage 2 will

be forced to min.n2, *i.e.*, the total sample size to n1+min.n2.

theta0 True ratio of T/R for simulating. Defaults to the GMR argument if missing.

theta1 Lower bioequivalence limit. Defaults to 0.8.

theta2 Upper bioequivalence limit. Defaults to 1.25.

npct Percentiles to be used for the presentation of the distribution of n(total)=n1+n2.

Defaults to c(0.05, 0.5, 0.95) to obtain the 5% and 95% percentiles and the

median.

nsims Number of studies to simulate.

If missing, nsims is set to 1E+05 = 100,000 or to 1E+06 = 1 Mio if estimating the empiric Type I Error ('alpha'), *i.e.*, with theta0 at the border or outside the

acceptance range theta1 ... theta2.

setseed Simulations are dependent on the starting point of the (pseudo) random number

generator. To avoid differences in power for different runs a set . seed (1234567)

is issued if setseed=TRUE, the default.

Set this argument to FALSE to view the variation in power between different

runs.

ulation steps. Defaults to FALSE.

Details

The calculations follow in principle the simulations as described in Potvin et al.

The underlying subject data are assumed to be evaluated after log-transformation. But instead of simulating subject data, the statistics pe1, mse1 and pe2, SS2 are simulated via their associated distributions (normal and χ^2 distributions).

power.tsd 11

Value

Returns an object of class "pwrtsd" with all the input arguments and results as components.

The class "pwrtsd"" has an S3 print method.

The results are in the components:

pBE Fraction of studies found BE.

pBE_s1 Fraction of studies found BE in stage 1.
pct_s2 Percentage of studies continuing to stage 2.

nmean Mean of n(total), aka average total sample size (ASN).

nrange Range (min, max) of n(total).

nperc Vector of percentiles of the distribution of n(total).

ntable Object of class "table" summarizing the discrete distribution of n(total) via its

distinct values and counts of occurences of these values.

This component is only given back if usePE==FALSE or otherwise if is.finite(Nmax),

i.e., a futility criterion is used.

Author(s)

D. Labes

References

Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ, Smith RA. Sequential design approaches for bioequivalence studies with crossover designs.

Pharm Stat. 2008; 7(4):245-62. doi: 10.1002/pst.294

Montague TH, Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ. Additional results for 'Sequential design approaches for bioequivalence studies with crossover designs'.

Pharm Stat. 2011; 11(1):8–13. doi: 10.1002/pst.483

Fuglsang A. Controlling type I errors for two-stage bioequivalence study designs.

Clin Res Reg Aff. 2011; 28(4):100–5. doi: 10.3109/10601333.2011.631547

Fuglsang A. Sequential Bioequivalence Trial Designs with Increased Power and Controlled Type I Error Rates.

AAPS J. 2013; 15(3):659–61. doi: 10.1208/s1224801394755

Fuglsang A. Futility Rules in Bioequivalence Trials with Sequential Designs.

AAPS J. 2014; 16(1):79-82. doi: 10.1208/s1224801395400

Schütz H. Two-stage designs in bioequivalence trials.

Eur J Clin Pharmacol. 2015; 71(3):271–81. doi: 10.1007/s0022801518062

Kieser M, Rauch G. Two-stage designs for cross-over bioequivalence trials.

Stat Med. 2015; 34(16):2403–16. doi: 10.1002/sim.6487

Zheng Ch, Zhao L, Wang J. Modifications of sequential designs in bioequivalence trials.

Pharm Stat. 2015; 14(3):180-8. doi: 10.1002/pst.1672

See Also

```
power.tsd.p for analogous calculations for 2-group parallel design.
```

power.tsd.fC for analogous calculations with futility check based on point estimate of stage 1.

Examples

```
# using all the defaults and 24 subjects in stage 1, CV of 25%
power.tsd(n1=24, CV=0.25)
# computation time ~ 1 sec
# as above, but save results for further use
res <- power.tsd(n1=24, CV=0.25)
## Not run:
# representation of the discrete distribution of n(total)
# via plot method of object with class "table" which creates a
# 'needle' plot
plot(res$ntable/sum(res$ntable), ylab="Density",
     xlab=expression("n"[total]), las=1,
    main=expression("Distribution of n"[total]))
# If you prefer a histogram instead (IMHO, not the preferred plot):
# reconstruct the ntotal values from the ntable
ntot <- rep.int(as.integer(names(res$ntable)),</pre>
                times=as.integer(res$ntable))
# annotated histogram
hist(ntot, freq=FALSE, breaks=res$nrange[2]-res$nrange[1],
     xlab=expression("n"[total]), las=1,
     main=expression("Histogram of n"[total]))
abline(v=c(res$nmean, res$nperc[["50%"]]), lty=c(1, 3))
legend("topright", box.lty=0, legend=c("mean", "median"),
       lty=c(1, 3), cex=0.9)
## End(Not run)
```

power.tsd.fC

Power calculation of adaptive 2-stage BE studies (2x2 crossover) with a futility criterion for the point estimate of T/R or its 90% CI

Description

This function calculates the 'empiric' power of 2-stage BE studies according to Potvin *et al.* 'method B/C' via simulations. The Potvin methods are modified to include a futility criterion for the point estimate or for its 90%CI and to allow the sample size estimation step to be done with the point estimate (PE) and MSE of stage 1.

Usage

Arguments

method Decision schemes according to Potvin et.al. (defaults to "B").

Montague's 'Method D' can be obtained by choosing "C" but setting alpha=c(0.028, 0.028).

'Method E' of Xu et al. can be obtained by choosing "B" and setting alphas, fu-

tility criterion "CI", max.n, and n1 according to the reference.

'Method F' can be obtained choosing "C" with the appropriate design setting according to the reference.

method="B0" uses the decision scheme of Zheng et al. MSDBE (modified sequential design for BE studies) which differs from B in case of different alphas w.r.t. power monitoring and BE decision in case of power >= target power.

alpha0 Alpha value for the first step(s) in Potvin "C", the power inspection and BE

decision if power > targetpower. Defaults to 0.05.

Only observed if method="C"

alpha Vector (two elements) of the nominal alphas for the two stages. Defaults to

Pocock's setting alpha=c(0.0294, 0.0294).

Common values together with other arguments are:

rep(0.0294, 2): Potvin et al. 'Method B' (fCrit="CI", fCupper=Inf)

rep(0.0269, 2): Fulgsang 'Method C/D' (method="C", GMR=0.9, targetpower=0.9, fCrit="CI", fCupper=Inf)

rep(0.0274, 2): Fuglsang 'Method C/D' (method="C", targetpower=0.9, fCrit="CI", fCupper=Inf)

rep(0.0280, 2): Montague et al. 'Method D' (method="C", GMR=0.9, fCrit="CI", fCup-

rep(0.0284, 2): Fulgsang 'Method B' (GMR=0.9, targetpower=0.9, fCrit="CI", fCupper=Inf)

rep(0.0304, 2): Kieser & Rauch (fCrit="CI", fCupper=Inf)

c(0.01, 0.04): Zheng et al. 'MSDBE' (method="B0", fCrit="CI", fCupper=Inf)

c(0.0249, 0.0357): Xu et al. 'Method E' for CV 10–30% (fCrit="CI", fClower=0.9374, max.n=42)

c(0.0254, 0.0363): Xu et al. 'Method E' for CV 30-55% (fCrit="CI", fClower=0.9305, max.n=42)

c(0.0248, 0.0364): Xu et al. 'Method F' for CV 10-30% (method="C", fCrit="CI", fClower=0.9492, max c(0.0259, 0.0349): Xu et al. 'Method F' for CV 30–55% (method="C", fCrit="CI", fClower=0.9305, max

Sample size of stage 1. For Xu's methods the recommended sample size should be at least 18 (if CV 10-30%) or 48 (if CV 30-55%).

Coefficient of variation of the intra-subject variability (use e.g., 0.3 for 30%).

GMR Ratio T/R to be used in decision scheme (power calculations in stage 1 and

sample size estimation for stage 2).

Power threshold in the power monitoring steps and power to achieve in the samtargetpower

ple size estimation step.

Power calculation method, also to be used in the sample size estimation for stage

Implemented are "nct" (approximate calculations via non-central t-distribution, "exact" (exact calculations via Owen's Q), and "shifted" (approximate calculation via shifted central t-distribution like in the paper of Potvin et al.

Defaults to "nct" as a reasonable compromise between speed and accuracy in

the sample size estimation step.

n1

C۷

pmethod

usePE If TRUE the sample size estimation step is done with MSE and PE of stage 1.

Defaults to FALSE, *i.e.*, the sample size is estimated with anticipated (fixed) GMR given as argument and MSE of stage 1 (analogous to Potvin *et. al.*).

GMR given as argument and MSE of stage 1 (analogous to Potvin et. al.).

If TRUE (the default) the interim power monitoring step in stage 1 evaluation of

'method B' will be done as described in Potvin *et.al*. Setting this argument to FALSE will omit this step.

Has no effect if method="C" is choosen.

min.n2 Minimum sample size of stage 2. Defaults to zero.

If the sample size estimation step gives N < n1 the sample size for stage 2 will

be forced to min.n2, *i.e.*, the total sample size to n1+min.n2.

max.n If max.n is set to a finite value the re-estimated total sample size (N) is set to

min(max.n, N).

Defaults to Inf which is equivalent to not constrain the re-estimated sample size. Attention! max.n here is **not** a futility criterion like Nmax in other functions of

this package.

fCrit Futility criterion.

powerstep

If set to "PE" the study stops after stage 1 if not BE and if the point estimate (PE) of stage 1 evaluation is outside the range defined in the next two arguments

"fClower" and "fCupper".

If set to "CI" the study stops after stage 1 if not BE and if the confidence interval of stage 1 evaluation is outside the range defined in the next two arguments.

Defaults to "PE". Futility criterion to use for PE or CI.

fClower Lower futility limit for the PE or CI of stage 1.

If the PE or CI is outside fClower ... fCupper the study is stopped in the interim

with the result FAIL (not BE).

May be missing. Defaults then to 0.8 if fCrit="PE" or 0.925 if fCrit="CI".

fCupper Upper futility limit for the PE or CI of stage 1.

Will be set to 1/fClower if missing.

theta0 Assumed ratio of geometric means (T/R) for simulations. If missing, defaults to

GMR.

theta1 Lower bioequivalence limit. Defaults to 0.8. theta2 Upper bioequivalence limit. Defaults to 1.25.

npct Percentiles to be used for the presentation of the distribution of n(total)=n1+n2.

Defaults to c(0.05, 0.5, 0.95) to obtain the 5% and 95% percentiles and the

median.

nsims Number of studies to simulate.

If missing, nsims is set to 1E+05 = 100,000 or to 1E+06 = 1 Mio if estimating the empiric Type I Error ('alpha'), *i.e.*, with theta0 at the border or outside the

acceptance range theta1 ... theta2.

setseed Simulations are dependent on the starting point of the (pseudo) random number

generator. To avoid differences in power for different runs a set. seed (1234567)

is issued if setseed=TRUE, the default.

Set this argument to FALSE to view the variation in power between different

runs.

ulation steps. Defaults to FALSE.

Details

The calculations follow in principle the simulations as described in Potvin et al.

The underlying subject data are assumed to be evaluated after log-transformation. But instead of simulating subject data, the statistics pe1, mse1 and pe2, SS2 are simulated via their associated distributions (normal and χ^2 distributions).

Value

Returns an object of class "pwrtsd" with all the input arguments and results as components.

The class "pwrtsd" has an S3 print method.

The results are in the components:

pBE Fraction of studies found BE.

pBE_s1 Fraction of studies found BE in stage 1.

pct_s2 Percentage of studies continuing to stage 2.

nmean Mean of n(total), aka average total sample size (ASN).

nrange Range (min, max) of n(total).

nperc Percentiles of the distribution of n(total).

ntable Object of class "table" summarizing the discrete distribution of n(total) via its

distinct values and counts of occurences of these values.

This component is only given back if usePE==FALSE or usePE==TRUE & fClower>0 & is.finite(fCupp

i.e., a futility range is used.

Author(s)

D. Labes

References

Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ, Smith RA. Sequential design approaches for bioequivalence studies with crossover designs.

Pharm Stat. 2008; 7(4):245–62. doi: 10.1002/pst.294

Montague TH, Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ. Additional results for 'Sequential design approaches for bioequivalence studies with crossover designs'.

Pharm Stat. 2011; 11(1):8–13. doi: 10.1002/pst.483

Fuglsang A. Sequential Bioequivalence Trial Designs with Increased Power and Controlled Type I Error Rates.

AAPS J. 2013; 15(3):659-61. doi: 10.1208/s1224801394755

Schütz H. Two-stage designs in bioequivalence trials.

Eur J Clin Pharmacol. 2015; 71(3):271-81. doi: 10.1007/s0022801518062

Kieser M, Rauch G. Two-stage designs for cross-over bioequivalence trials.

Stat Med. 2015; 34(16):2403–16. doi: 10.1002/sim.6487

Zheng Ch, Zhao L, Wang J. Modifications of sequential designs in bioequivalence trials.

Pharm Stat. 2015; 14(3):180-8. doi: 10.1002/pst.1672

power.tsd.GS

Xu J, Audet C, DiLiberti CE, Hauck WW, Montague TH, Parr TH, Potvin D, Schuirmann DJ. *Optimal adaptive sequential designs for crossover bioequivalence studies*. Pharm Stat. 2016;15(1):15–27. doi: 10.1002/pst.1721

See Also

```
power.tsd
```

Examples

power.tsd.GS

Power calculation for non-adaptive group sequential (2-stage) BE studies

Description

This function calculates the 'empiric' power of group sequential 2-stage BE in 2×2 crossover designs via simulations.

The number of subjects in both stages has to be prespecified (non-adaptive).

Usage

```
power.tsd.GS(alpha = c(0.0294, 0.0294), n, CV, theta0, theta1, theta2, fCrit = c("CI", "PE"), fClower, fCupper, nsims, setseed = TRUE, details = FALSE)
```

Arguments

alpha	Vector of the two nominal alpha values to be used in the $100(1-2\alpha)$ confidence interval calculations in the two stages. Use something like package ldbounds for choosing the nominal alphas.
n	Vector of the two sample sizes in stage 1 and stage 2. $n(total)$ is $n[1]+n[2]$ if a second stage is necessary. Otherwise it is $n[1]$.
CV	Coefficient of variation of the intra-subject variability (use <i>e.g.</i> , 0.3 for 30%).

power.tsd.GS 17

theta0 Assumed ratio of geometric means (T/R) for simulations. If missing, defaults to

GMR.

theta1 Lower bioequivalence limit. Defaults to 0.8.
theta2 Upper bioequivalence limit. Defaults to 1.25.

fCrit Futility criterion.

If set to "PE" the study stops after stage 1 if not BE and if the point estimate (PE) of stage 1 evaluation is outside the range defined in the next two arguments

"fClower" and "fCupper".

If set to "CI" the study stops after stage 1 if not BE and if the 90% confidence interval of stage 1 evaluation is outside the range defined in the next two argu-

ments.

Defaults to "CI".

fClower Lower limit of the futility criterion. Defaults to 0 if missing, i.e., no futility

criterion.

fCupper Upper limit of the futility criterion. Defaults to 1/fClower if missing.

nsims Number of studies to simulate.

If missing, nsims is set to 1E+05 = 100,000 or to 1E+06 = 1 Mio if estimating the empiric Type I Error ('alpha'), *i.e.*, with theta0 at the border or outside the

acceptance range theta1 ... theta2.

setseed Simulations are dependent on the starting point of the (pseudo) random number

generator. To avoid differences in power for different runs a set. seed (1234567)

is issued if setseed=TRUE, the default.

Set this argument to FALSE to view the variation in power between different

runs.

ulation steps. Defaults to FALSE.

Details

The calculations follow in principle the simulations as described in Potvin *et al.* for adaptive designs, but with no interim power inspection and pre-specified (fixed) number of subjects in stage 2. The underlying subject data are assumed to be evaluated after log-transformation. But instead of simulating subject data, the statistics pe1, mse1 and pe2, SS2 are simulated via their associated distributions (normal and χ^2 distributions).

Value

Returns an object of class "pwrtsd" with all the input arguments and results as components.

The class "pwrtsd" has a S3 print method.

The results are in the components:

pBE Fraction of studies found BE.

pBE_s1 Fraction of studies found BE in stage 1.

pct_s2 Percentage of studies continuing to stage 2.

Note

The code is reasonable fast. 1E6 sims take ~ 1 sec on my machine. Even 1E7 sims are meanwhile possible without too much beer. :-)

Author(s)

D. Labes

References

```
Gould AL. Group sequential extensions of a standard bioequivalence testing procedure.
J Pharmacokin Biopharm. 1995; 23(1):57–86 doi: 10.1007/BF02353786

Patterson SD, Jones B. Bioequivalence and Statistics in Clinical Pharmacology.
Boca Raton: CRC Press; 2<sup>nd</sup> edition 2016. Chapter 5.6 Optional Designs.

Schütz H. Two-stage designs in bioequivalence trials.
```

Eur J Clin Pharmacol. 2015; 71(3):271–81. doi: 10.1007/s0022801518062

Kieser M, Rauch G. Two-stage designs for cross-over bioequivalence trials.

Stat Med. 2015; 34(16):2403–16. doi: 10.1002/sim.6487

Zheng Ch, Zhao L, Wang J. *Modifications of sequential designs in bioequivalence trials*. Pharm Stat. 2015; 14(3):180–8. doi: 10.1002/pst.1672

See Also

```
power.tsd and power.tsd.p for adaptive sequential designs.
```

Examples

```
# using the Pocock alpha settings and no futility rule # (defaults), a CV of 20% and 12 subjects in both stages, # midway interim power.tsd.GS(CV=0.2, n=c(12,12))
```

power.tsd.in

Power calculation of adaptive 2-stage BE studies with 2x2 crossover design, based on the Inverse Normal method

Description

The function calculates the 'empirical' power of 2-stage BE studies based on the Inverse-Normal combination method. The main design scheme is according to Maurer *et al.* (Maximum Combination Test), but it may also be used for other/modified designs, e.g. standard combination method, usage of the observed treatment difference after stage 1 in the sample size re-estimation step or different futility criteria.

Usage

Arguments

alpha If one element is given, the overall one-sided significance level. In this case the

adjusted alpha levels will be calculated internally. If two elements are given, the argument refers to the two adjusted one-sided alpha levels for stage 1 and stage

2, respectively.

If missing, defaults to 0.05.

weight Pre-defined weight(s) of stage 1, see 'Details' for more information. Note that

using the notation from Maurer et al, weight corresponds to information fraction, other literature may refer to sqrt(weight) as being the weight. weight must either contain one element (in case of max.comb.test = FALSE) or two elements (in

case of max.comb.test = TRUE).

If missing, defaults to 0.5 for max.comb.test = FALSE and to c(0.5, 0.25) for

max.comb.test = TRUE.

max.comb.test Logical; if TRUE (default) the maximum combination test will be used, other-

wise the standard combination test.

n1 Sample size of stage 1.

CV Coefficient of variation of the intra-subject variability (use e.g., 0.3 for 30%).

targetpower Desired (overall) target power to declare BE at the end of the trial.

theta0 Assumed ratio of geometric means (T/R) for simulations. If missing, defaults to

GMR.

theta1 Lower bioequivalence limit. Defaults to 0.8.

theta2 Upper bioequivalence limit. Defaults to 1.25.

GMR Assumed ratio of geometric means (T/R) to be used in power calculation for

stage 1 and sample size re-estimation for stage 2. If missing, defaults to 0.95.

usePE If TRUE the sample size re-estimation is done with the observed point estimate

(PE) of the treatment difference in stage 1.

Defaults to FALSE.

Note: The futility inspection via the power of stage 1 is always done with the

planning value GMR.

min.n2 Minimum sample size of stage 2. Defaults to 4.

If the sample size re-estimation step gives a sample size for stage 2 less than

min.n2, then min.n2 will be used for stage 2.

max.n Maximum overall sample size stage 1 + stage 2.

This is *not* a futility criterion regarding the maximum sample size! If max.n is set to a finite value and the sample size re-estimation gives a sample size for stage 2 (n2) such that n1 + n2 > max.n, then the sample size for stage 2 will be

set to n2 = max.n - n1.

Defaults to Inf, i.e., no constraint on the re-estimated sample size.

fCpower

Threshold for power monitoring step to decide on futility for cases where BE has not been achieved after stage 1: If BE has not been achieved after stage 1 and the power for stage 1 is greater than or equal to fCpower, then the study will be considered a failure.

See 'Details' for more information on the choice of fCpower.

fCrit

Futility criterion to use: "No" (no futility criterion regarding observed point estimate, confidence interval and maximum sample size), "PE" (observed point estimate of the geometric mean ratio from stage 1), "CI" (90% confidence interval of the geometric mean ratio from stage 1), "Nmax" (overall maximum sample size); or a combination thereof (concatenate abbreviations; see 'Examples'). Defaults to "CI".

fClower

Lower futility limit for the PE or CI of stage 1.

If the PE or CI is completely outside of fClower ... fCupper the study is stopped due to futility (not BE).

May be missing. If "PE" or "CI" is specified within fCrit, the default will be set to 0.8 for fCrit = "PE" or 0.95 for fCrit = "CI". If neither "PE" nor "CI" is specified within fCrit, there will be no futility constraint regarding point estimate or confidence interval from stage 1 (regardless of any specification of fClower and/or fCupper).

fCupper

Upper futility limit for the PE or CI of stage 1.

Analogous to fClower: Will be set to 1/fClower if missing.

fCNmax

Futility criterion regarding maximum sample size. If the determined sample size for stage 2 (n2) is such that n1 + n2 > fCNmax, the study will not continue to stage 2 and stopped due to futility (not BE).

If "Nmax" is specified within fCrit and argument fCNmax is missing, the value will be set to fCNmax = 4*n1. If "Nmax" is not specified within fCrit, then there will be no futility constraint regarding maximum sample size (regardless of any specification of fCNmax).

ssr.conditional

Method for sample size re-estimation step: "no" does not use conditional error rates nor the estimated conditional target power for the second stage, "error" uses conditional error rates for the second stage, and "error_power" uses both conditional error rates and the estimated conditional target power for the second stage.

Defaults to "error_power".

See also 'Details'.

pmethod

Power calculation method, also to be used in the sample size estimation for stage 2.

Implemented are "nct" (approximate calculations via non-central *t*-distribution, "exact" (exact calculations via Owen's Q), and "shifted" (approximate calculation via shifted central *t*-distribution like in the paper of Potvin *et al.*)

Defaults to "nct" as it is a reasonable compromise between speed and accuracy in the sample size estimation step.

npct Percentiles to be used for the presentation of the distribution of n(total)=n1+n2.

Defaults to c(0.05, 0.5, 0.95) to obtain the 5% and 95% percentiles and the

median.

nsims Number of studies to simulate.

If missing, nsims is set to 1E+05 = 100,000 or to 1E+06 = 1 Mio if estimating the empiric Type I Error ('alpha'), *i.e.*, with theta0 at the border of the acceptance

range theta1 ... theta2.

setseed Simulations are dependent on the starting point of the (pseudo) random number

generator. To avoid differences in power for different runs a set. seed (1234567)

is issued if setseed=TRUE, the default.

Set this argument to FALSE to view the variation in power between different

runs

details If set to TRUE the function prints the results of time measurements of the sim-

ulation steps. Default to FALSE.

Details

The underlying subject data are assumed to be evaluated after log-transformation. But instead of simulating subject data, the statistics pe1, mse1 and pe2, mse2 are simulated via their associated distributions (Normal and χ^2 distribution).

The weight argument always refers to the first weight of a pair of weights. For example, in case of $\max. comb. test = FALSE$ the standard combination test requires two weights (w, 1-w) but only the first one, w, is required as input argument here because the second weight is automatically specified once the first is given. Similarly for $\max. comb. test = TRUE$, w and w* need to be specified, which in turn define the two pairs of weights (w, 1-w) and (w*, 1-w*).

If ssr.conditional = "error_power", the design scheme generally calculates the estimated conditional target power of the second stage and uses this value as desired target power in the sample size reestimation process. If fCpower > targetpower, then the conditional target power may actually be negative. This does not seem sensible. Therefore, for such cases the desired target power for the sample size re-calculation will be set to targetpower.

Value

Returns an object of class "pwrtsd" with all the input arguments and results as components. As part of the input arguments a component cval is also presented, containing the critical values for stage 1 and 2 according to the input based on alpha, weight and max.comb.test.

The class "pwrtsd" has an S3 print method.

The results are in the components:

pBE Fraction of studies found BE.

pBE_s1 Fraction of studies found BE in stage 1.

pct_stop_s1 Percentage of studies stopped after stage 1 (due to BE or due to futility).

pct_stop_fut Percentage of studies stopped after stage 1 due to futility.

pct_s2 Percentage of studies continuing to stage 2.

nmean Mean of n(total).

nrange Range (min, max) of n(total).

nperc Vector of percentiles of the distribution of n(total).

Author(s)

B. Lang

References

König F, Wolfsegger M, Jaki T, Schütz H, Wassmer G.

Adaptive two-stage bioequivalence trials with early stopping and sample size re-estimation.

Vienna: 2014; 35^{th} Annual Conference of the International Society for Clinical Biostatistics. Poster P1.2.88

doi: 10.13140/RG.2.1.5190.0967.

Patterson SD, Jones B. Bioequivalence and Statistics in Clinical Pharmacology.

Boca Raton: CRC Press; 2^{nd} edition 2017.

Kieser M, Rauch G. Two-stage designs for cross-over bioequivalence trials.

Stat Med. 2015; 34(16):2403–16. doi: 10.1002/sim.6487.

Maurer W, Jones B, Chen Y. Controlling the type 1 error rate in two-stage sequential designs when testing for average bioequivalence.

Stat Med. 2018;1–21. doi: 10.1002/sim.7614.

Wassmer G, Brannath W. *Group Sequential and Confirmatory Adaptive Designs in Clinical Trials*. Springer 2016. doi: 10.1007/9783319325620.

See Also

```
interim.tsd.in, final.tsd.in
```

Examples

```
# 12 subjects in stage 1, CV of 20%, no futility rule, otherwise all defaults
# except too low number of sims due to CRAN policy for run-time of examples
# This setting resembles values in Table 6.8 and 6.10 in Patterson and Jones
# if nsims=1e5 (default)
power.tsd.in(n1 = 12, CV = 0.2, fCrit = "No", nsims=1e4)
# Note that actual replication would require the following additional
# arguments (extremely long run-time)
## Not run:
power.tsd.in(n1 = 12, CV = 0.2, fCrit = "No", max.n = 4000,
             pmethod = "exact", nsims = 1E6)
## End(Not run)
# Table 8 in Maurer et al
power.tsd.in(n1 = 36, CV = 0.4, max.n = 4000)
# Same example as before but with additional futility criterion regarding
# maximum sample size (using the default 4*n1 as maximum bound)
power.tsd.in(n1 = 36, CV = 0.4, \max.n = 4000, fCrit = c("CI", "Nmax"))
```

23 power.tsd.KM

power.tsd.KM	Power calculation of adaptive 2-stage BE studies (2x2 crossover)
	(variant of power.2stage to obtain the results of Karalis / Macheras)

Description

This function calculates the 'empiric' power of 2-stage BE studies according to Potvin et al. via simulations. The Potvin methods are modified as described by Karalis & Macheras to include a futility criterion Nmax and to perform the power calculation steps and the sample size estimation step in the decision schemes with the MSE (calculated from CV) and the point estimate (PE) of T/R from stage 1.

Usage

```
power.tsd.KM(method = c("C", "B"), alpha0 = 0.05, alpha = c(0.0294, 0.0294),
             n1, CV, targetpower = 0.8, pmethod = c("nct", "exact"),
             Nmax = 150, theta0, theta1, theta2, npct = c(0.05, 0.5, 0.95),
             nsims, setseed = TRUE, details = FALSE)
```

Arg

Nmax

rguments	
method	Decision schemes according to Potvin <i>et al</i> . Default is "C" aka TSD in the paper of Karalis & Macheras if setting alpha=c(0.0294, 0.0294). TSD-1 of Karalis can be obtained by choosing "C" but setting alpha=c(0.028, 0.028). TSD-2 of Karalis can be obtained by choosing "B" and setting alpha=c(0.0294, 0.0294).
alpha0	Alpha value for the first step(s) in Potvin C aka TSD of Karalis & Macheras or TSD-1 of Karalis, the power inspection and BE decision if power > targetpower. Defaults to 0.05.
alpha	Vector (two elements) of the nominal alphas for the two stages. Defaults to Pocock's alpha setting alpha=c(0.0294, 0.0294) as in TSD of Karalis & Macheras.
n1	Sample size of stage 1.
CV	Coefficient of variation of the intra-subject variability (use e.g., 0.3 for 30%).
targetpower	Power threshold in the first step of Potvin "C" and power to achieve in the sample size estimation step.
pmethod	Power calculation method, also to be used in the sample size estimation for stage 2.
	Implemented are ""nct" (approximate calculations via non-central <i>t</i> -distribution and "exact" (exact calculations via Owen's Q).
	Defaults to "nct" as a reasonable compromise between speed and accuracy in
	the sample size estimation step.

Futility criterion. If set to a finite value all studies simulated in which a sample size >Nmax is obtained will be regarded as BE=FAIL. Defaults to 150, as rec-

Set this argument to Inf, to work without that futility criterion.

ommended by Karalis & Macheras.

power.tsd.KM

theta0 Assumed ratio of geometric means (T/R) for simulations. If missing, defaults to

GMR.

theta1 Lower bioequivalence limit. Defaults to 0.8. theta2 Upper bioequivalence limit. Defaults to 1.25.

npct Percentiles to be used for the presentation of the distribution of n(total)=n1+n2.

Defaults to c(0.05, 0.5, 0.95) to obtain the 5% and 95% percentiles and the

median.

nsims Number of studies to simulate.

If missing, nsims is set to 1E+05 = 100,000 or to 1E+06 = 1 Mio if estimating the empiric Type I Error ('alpha'), *i.e.*, with theta0 at the border or outside the

acceptance range theta1 ... theta2.

setseed Simulations are dependent on the starting point of the (pseudo) random number

generator. To avoid differences in power for different runs a set. seed (1234567)

is issued if setseed=TRUE, the default.

Set this argument to FALSE to view the variation in power between different

runs.

details If set to TRUE the function prints the results of time measurements of the sim-

ulation steps. Defaults to FALSE.

Details

The calculations follow in principle the simulations as described in Potvin et al.

The underlying subject data are assumed to be evaluated after log-transformation. But instead of simulating subject data, the statistics pe1, mse1 and pe2, SS2 are simulated via their associated distributions (normal and χ^2 distributions).

In contrast to Potvin *et al.* the power calculation steps as well as the sample size adaption step of the decision schemes are done using the MSE (calculated from CV) **and** the point estimate from stage 1.

This resembles the methods described in Karalis & Macheras and Karalis.

Value

Returns an object of class "pwrtsd" with all the input arguments and results as components.

The class "pwrtsd" has a S3 print method.

The results are in the components:

pBE Fraction of studies found BE.

pBE_s1 Fraction of studies found BE in stage 1.
pct_s2 Percentage of studies continuing to stage 2.

nmean Mean of n(total).

nrange Range (min, max) of n(total).

nperc Percentiles of the distribution of n(total).

ntable Object of class "table" summarizing the discrete distribution of n(total) via its

distinct values and counts of occurences of these values. This component is only given back if is.finite(Nmax).

power.tsd.p 25

Author(s)

D. Labes

References

Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ, Smith RA. Sequential design approaches for bioequivalence studies with crossover designs.

Pharm Stat. 2008; 7(4):245–62. doi: 10.1002/pst.294

Karalis V, Macheras P. An Insight into the Properties of a Two-Stage Design in Bioequivalence Studies.

Pharm Res. 2013; 30(7):1824–35. doi: 10.1007/s1109501310263

Karalis V. The role of the upper sample size limit in two-stage bioequivalence designs.

Int J Pharm. 2013; 456(1):87–94. doi: 10.1016/j.ijpharm.2013.08.013

Fuglsang A. Futility Rules in Bioequivalence Trials with Sequential Designs.

AAPS J. 2014; 16(1):79-82. doi: 10.1208/s1224801395400

Schütz H. Two-stage designs in bioequivalence trials.

Eur J Clin Pharmacol. 2015; 71(3):271-81. doi: 10.1007/s0022801518062

Examples

```
# using all the defaults
# but too low number of sims to complain with the CRAN policy:
# "check time only a few seconds per example"
# minimum number of sims should be 1E5 for power, 1E6 sims for 'alpha'
power.tsd.KM(n1=16, CV=0.2, nsims=1E4)
# ~3 sec if nsims=1E5
```

power.tsd.p

Power calculation of adaptive 2-stage BE studies in 2-group parallel designs

Description

This functions calculate the 'empirical' power of 2-stage BE studies with 2 parallel groups according to Fuglsang via simulations. The Potvin decision schemes are modified to include a futility criterion Nmax and to allow the sample size estimation step to be done with point estimate and variabilities from stage 1 (fully adaptive).

Function power.tsd.pAF() performes exactly as described in Fuglsang's paper, namely the power monitoring steps and the sample size estimation are based always on the pooled *t*-test.

Function power.tsd.p() with argument test="welch" on the other hand uses the genuine power of Welch's test. Moreover it accepts unequal stage 1 treatment groups.

26 power.tsd.p

Usage

```
power.tsd.p(method = c("B", "C"), alpha0 = 0.05, alpha = c(0.0294, 0.0294),
            n1, GMR, CV, targetpower = 0.8,
            pmethod = c("nct", "exact", "shifted"),
            usePE = FALSE, Nmax = Inf, test = c("welch", "t-test", "anova"),
            theta0, theta1, theta2, npct = c(0.05, 0.5, 0.95), nsims,
            setseed = TRUE, details = FALSE)
power.tsd.pAF(method = c("B", "C"), alpha0 = 0.05, alpha = c(0.0294, 0.0294),
              n1, GMR, CV, targetpower = 0.8,
              pmethod = c("shifted", "nct", "exact"),
              usePE = FALSE, Nmax = Inf, test = c("welch", "t-test", "anova"),
              theta0, theta1, theta2, npct = c(0.05, 0.5, 0.95), nsims,
              setseed = TRUE, details = FALSE)
```

Arguments

method Decision schemes according to Potvin et.al. (defaults to "B").

Potvin's 'method D' can be obtained by choosing "C" but setting alpha=c(0.028, 0.028).

alpha0 Alpha value for the first step(s) in Potvin "C", the power inspection and BE

decision if power > targetpower.

Defaults to 0.05.

alpha Vector (two elements) of the nominal alphas for the two stages.

Defaults to Pocock's alpha setting alpha=c(0.0294, 0.0294).

Sample size of stage 1. n1

> Function power.tsd.p() accepts also a vector of stage 1 sample sizes with two elements, where the number of subjects in the in treatment group T should be given in the first element and the number of subjects in the threatment group Rin the second.

If given with one element, the total n1 should be even.

GMR Ratio T/R to be used in decision scheme (power calculations in stage 1 and

sample size estimation for stage 2).

C۷ Coefficient of variation of the total variability (use e.g., 0.3 for 30%)

Can be a vector with two elements. In that case CV[1] is for Test and CV[2] for

the Reference treatment.

targetpower Power threshold in the power monitoring steps and power to achieve in the sam-

ple size estimation step.

pmethod Power calculation method, also to be used in the sample size estimation for stage

Implemented are "nct" (approximate calculations via non-central t-distribution, "exact" (exact calculations via Owen's Q), and "shifted" (approximate calcula-

tion via shifted central t-distribution

Defaults to "nct" for speed reasons in function power.tsd.p() as a reasonable compromise between speed and accuracy in the sample size estimation step.

Defaults to "shifted" in function power.tsd.pAF() for consistancy with Fuglsang.

power.tsd.p 27

usePE If TRUE the sample size estimation step is done with MSE **and** PE of stage 1.

Defaults to FALSE i.e., the sample size is estimated with GMR and MSE (calcu-

lated from CV) of stage 1 analogous to Potvin et. al.

NB: The power inspection steps in the Potvin methods are always done with the

GMR argument and MSE (CV) of stage 1.

Nmax Futility criterion. If set to a finite value, all studies simulated in which a sample

size >Nmax is obtained will be regarded as BE=FAIL.

Set this argument to Inf, the default, to work without that futility criterion.

test Test on which the CI calculations are based on.

Defaults to "welch" = Welch's t-test accounting for heteroscedasticity in the

variabilities of Test and Reference, but neglecting stage effects.

"anova" calculates the $100(1-2\alpha)$ confidence interval based on an ANOVA with

treatment and stage in the model.

"t-test" calculates the $100(1-2\alpha)$ confidence interval based on the t-test assum-

ing equal variabilities of Test and Reference and neglecting stage effects.

theta0 Assumed ratio of geometric means (T/R) for simulations. If missing, defaults to

GMR.

theta1 Lower bioequivalence limit. Defaults to 0.8.

theta2 Upper bioequivalence limit. Defaults to 1.25.

npct Percentiles to be used for the presentation of the distribution of n(total)=n1+n2.

Defaults to c(0.05, 0.5, 0.95) to obtain the 5% and 95% percentiles and the

median.

nsims Number of studies to simulate.

If missing, nsims is set to 1E+05 = 100,000 or to 1E+06 = 1 Mio if estimating

the empiric Type I Error ('alpha'), i.e., with theta0 at the border or outside the

acceptance range theta1 ... theta2.

setseed Simulations are dependent on the starting point of the (pseudo) random number

generator. To avoid differences in power for different runs a set. seed (1234567)

is issued if setseed=TRUE, the default.

Set this argument to FALSE to view the variation in power between different

runs.

details If set to TRUE the function prints the results of time measurements of the sim-

ulation steps. Defaults to FALSE.

Details

The calculations follow in principle the simulations as described by Fuglsang.

The underlying subject data are assumed to be evaluated after log-transformation. But instead of simulating subject data the statistics (mean and variance of Test and Reference of stage 1 and stage 2) are simulated via their associated distributions (normal and χ^2).

Value

Returns an object of class "pwrtsd" with all the input arguments and results as components.

The class "pwrtsd" has an S3 print method.

The results are in the components:

28 power.tsd.ssr

pBE	Fraction of studies found BE.
pBE_s1	Fraction of studies found BE in stage 1.
pct_s2	Percentage of studies continuing to stage 2.
nmean	Mean of n(total).
nrange	Range (min, max) of n(total).
nperc	Percentiles of the distribution of n(total).
ntable	Object of class "table" summarizing the discrete distribution of n(total) via its distinct values and counts of occurences of these values. This component is only given back if usePE==FALSE or otherwise if is.finite(Nmax), <i>i.e.</i> , a futility

Author(s)

D. Labes

References

Fuglsang A. Sequential Bioequivalence Approaches for Parallel Design.

AAPS J. 2014; 16(3):373-8. doi: 10.1208/s1224801495711

criterion is used.

Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ, Smith RA. Sequential design approaches for bioequivalence studies with crossover designs.

Pharm Stat. 2008; 7(4):245–62. doi: 10.1002/pst.294

See Also

power.2stage for analogous calculations for the 2×2 crossover.

Examples

```
# using all the defaults
power.tsd.p(n1=48, CV=0.25)
```

power.tsd.ssr

Power of 2-stage BE studies in 2x2 crossover designs with interim sample size re-estimation

Description

This function calculates the 'empiric' power (via simulations) of 2-stage BE studies with interim sample size re-estimation (*i.e.*, but no BE decision after stage 1). The sample size re-estimation can be done blinded or unblinded.

power.tsd.ssr 29

Usage

Arguments

alpha Nominal type I error. Has to be adjusted in case of inflation of the Type I Error.

n1 Sample size of stage 1.

GMR Ratio T/R to be used in the sample size re-estimation.

Defaults to 0.95 if missing.

CV Coefficient of variation of the intra-subject variability (use e.g., 0.3 for 30%).

Anticipated population value.

targetpower Power to achieve in the sample size estimation step.

pmethod Power calculation method to be used in the sample size re-estimation for stage

2..

Implemented are "nct" (approximate calculations via non-central *t*-distribution, "exact" (exact calculations via Owen's Q), and "shifted" (approximate calculations via Owen's Q), and "shifted" (approximate calculations via Owen's Q).

tion via shifted central *t*-distribution.

Also implemented is the large sample approximation as used in the references. Defaults to "nct" as a reasonable compromise between speed and accuracy in

the sample size estimation step.

blind If TRUE the blinded estimate of the intra-subject variance, *i.e.*, the estimate from

the period differences, is used in sample size estimation.

If FALSE the usual MSE from stage 1 is used. Defaults to FALSE since most BE studies are open.

usePE If TRUE the point estimate from the interim analysis is used in the sample size

re-estimation. Defaults to FALSE.

usePE = TRUE doesn't make sense if blind = TRUE. In that case the function

issues a warning and usePE is reset to usePE = FALSE.

min.n If $\min.n>n1$, the re-estimated sample size (N) is set to $\max(\min.n,N)$.

If min.n=0 (the default), no minimal sample size is applied.

max.n If max.n is set to a finite value the re-estimated sample size (N) is set to min(max.n,N).

Defaults to Inf which is equivalent to not constrain the re-estimated sample size. Attention! max.n is here **not** a futility criterion like Nmax in other functions of

this package.

theta0 Assumed ratio of geometric means (T/R) for simulations. If missing, defaults to

GMR.

theta1 Lower bioequivalence limit. Defaults to 0.8.

theta2 Upper bioequivalence limit. Defaults to 1.25.

npct Percentiles to be used for the presentation of the distribution of n(total)=n1+n2.

Defaults to c(0.05, 0.5, 0.95) to obtain the 5% and 95% percentiles and the

median.

30 power.tsd.ssr

nsims Number of studies to simulate.

If missing, nsims is set to 1E+05 = 100,000 or to 1E+06 = 1 Mio if estimating the empiric Type I Error ('alpha'), *i.e.*, with theta0 at the border or outside the

acceptance range theta1 ... theta2.

setseed Simulations are dependent on the starting point of the (pseudo) random number

generator. To avoid differences in power for different runs a set. seed (1234567)

is issued if setseed=TRUE, the default.

Set this argument to FALSE to view the variation in power between different

runs.

details If set to TRUE the function prints the results of time measurements of the sim-

ulation steps. Defaults to FALSE.

Details

The calculations follow in principle the simulations as described in Potvin et al.

The underlying subject data are assumed to be evaluated after log-transformation. But instead of simulating subject data, the statistics pe1, mse1 and pe2, SS2 are simulated via their associated distributions (normal and χ^2 distributions).

Value

Returns an object of class class "pwrtsd" with all the input arguments and results as components. The class class "pwrtsd" has a S3 print method.

The results are in the components:

pBE Fraction of studies found BE.

pct_s2 Percentage of studies continuing to stage 2.

nmean Mean of n(total).

nrange Range (min, max) of n(total).

nperc Percentiles of the distribution of n(total).

ntable Object of class "table" summarizing the discrete distribution of n(total) via its

unique values and counts of occurences of these values.

ntable is only given back if usePE = FALSE

Note

The computation time is in the magnitude of a few seconds for 100,000 sim's on my machine (Intel core i7 2.5 GHz, 12GB RAM) if the non-central *t* approximation is used. Thus be a bit patient if you simulate for the Tpye I Error 'alpha' with 1 Mio sim's.

Using the crude pmethod="ls" on the other hand results in a nearly immediate sample size reestimation.

Author(s)

D. Labes

sampleN2.TOST 31

References

```
Golkowski D, Friede T, Kieser M. Blinded sample size re-estimation in crossover bioequivalence trials.

Pharm Stat. 2014; 13(3):157–62. doi: 10.1002/pst.1617

Jones B, Kenward MG. Design and Analysis of Cross-Over Trials.

Boca Raton: CRC Press; 3<sup>rd</sup> edition 2014. Chapter 12.

Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ, Smith RA. Sequential design approaches for bioequivalence studies with crossover designs.

Pharm Stat. 2008; 7(4):245–62. doi: 10.1002/pst.294
```

See Also

power. tsd for 2-stage studies with interim BE decision.

Examples

sampleN2.TOST

Sample size re-estimation of adaptive 2-stage BE studies in 2x2 crossover and parallel designs based on power of TOST

Description

This function estimates the necessary sample size of stage 2 to have at least a given power.

Usage

32 sampleN2.TOST

Arguments

alpha Alpha value for the final analysis of pooled data. Defaults to Pocock's alpha

setting alpha=0.0294.

CV Coefficient of variation of the intra-subject variability as ratio.

n1 Sample size of stage 1.

theta0 True ratio of T/R for simulating. Defaults to 0.95 argument if missing.

theta1 Lower bioequivalence limit. Defaults to 0.8.
theta2 Upper bioequivalence limit. Defaults to 1.25.
targetpower Power to achieve at least. Must be >0 and <1.

design Character string describing the study design. Implemented are "2x2" and "par-

allel".

method Method for calculation of the power. Implemented are "exact" (exact calculation

via Owen's Q), "nct" (approximate calculation via non-central *t*-distribution, and "shifted" (approximate calculation via shifted central *t*-distribution like in the

paper of Potvin *et al*. Defaults to "exact".

imax Maximum number of steps in sample size search.

Defaults to 100. Adaption only in rare cases needed.

Details

The sample size is calculated via iterative evaluation of power of the TOST procedure.

Start value for the sample size search is taken from a large sample approximation according to Zhang, modified.

Value

A data.frame with the input and results will be returned.

The "Sample size" column contains the sample size for the second stage.

Note

Since in the final analysis one degree of freedom less than in a fixed sample design (or in stage 1) is used, power will be slightly lower than the one obtained with function sampleN.TOST of package PowerTOST. However, different *sample sizes* are extremely unlikely.

Author(s)

H. Schütz, D. Labes

References

Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ, Smith RA. Sequential design approaches for bioequivalence studies with crossover designs.

Pharm Stat. 2008; 7(4):245–62. doi: 10.1002/pst.294

Zhang P. A Simple Formula for Sample Size Calculation in Equivalence Studies.

J Biopharm Stat. 2003; 13(3):529–38. doi: 10.1081/BIP120022772

sampleN2.TOST 33

Examples

```
# using all the defaults, CV of 25% and 12 subjects in stage 1
print(sampleN2.TOST(CV=0.25, n1=12), row.names=FALSE)
# should give a stage 2 sample of 22 and achieved power ~0.814
# CV 10% and 12 subjects in stage 1
print(sampleN2.TOST(CV=0.1, n1=12), row.names=FALSE)
# should give a sample size of 0 (second stage not reasonable
# since power ~0.973 already achieved in stage 1
```

Index

```
final.2stage.in(final.tsd.in), 2
final.tsd.in, 2, 8, 22
interim.2stage.in(interim.tsd.in), 4
interim.tsd.in, 4, 4, 22
power.2stage, 28
power.2stage (power.tsd), 9
power.2stage.fC (power.tsd.fC), 12
power.2stage.GS (power.tsd.GS), 16
power.2stage.in (power.tsd.in), 18
power.2stage.KM (power.tsd.KM), 23
power.2stage.p (power.tsd.p), 25
power.2stage.pAF (power.tsd.p), 25
power.2stage.ssr(power.tsd.ssr), 28
power.tsd, 9, 16, 18, 31
power.tsd.fC, 11, 12
power.tsd.GS, 16
power.tsd.in, 4, 8, 18
power.tsd.KM, 23
power.tsd.p, 11, 18, 25
power.tsd.pAF (power.tsd.p), 25
power.tsd.ssr, 28
sampleN2.TOST, 31
```