Package 'SurrogateSeq'

June 18, 2024

Type Package

Title Group Sequential Testing of a Treatment Effect Using a Surrogate Marker

Version 1.0

Author Layla Parast, Jay Bartroff

Maintainer Layla Parast <parast@austin.utexas.edu>

Description
Provides functions to implement group sequential procedures that allow for early stopping to declare efficacy using a surrogate marker and the possibility of futility stopping. More details will be available in the future in: Parast, L. and Bartroff, J (2024) ``Group Sequential Testing of a Treatment Effect Using a Surrogate Marker".

License GPL

Imports stats, MASS, ggplot2

NeedsCompilation no

Depends R (>= 3.5.0)

Contents

bdr.gs.mc.fut *Calculates the boundaries for the group sequential tests with futility stopping*

Description

Returns the boundaries for any group sequential test of the null vs. 2-sided alternative whose boundaries take the form of a single constant times a known weight vector, which is w.vec; allows for futility stopping. These include Pocock (w.vec=(1,1,..)), O'Brien-Fleming (w.vec=(sqrt(n.stg/1), sqrt(n.stg/2), ..., 1)), etc. It does this by returning quantiles of the sample paths of the (null) test statistic paths in mc.paths.

Usage

```
bdr.gs.mc.fut(c1 = NULL, c2 = NULL, pp = 0.4, n.stg, j.star = 1, alpha = 0.05,
alpha0 = (j.star/n.stg) * alpha, mc.paths, inf.fraction = (1:n.stg)/n.stg,
N.iter.max = 100, alpha.tol = 0.02 * alpha)
```

Arguments

c1	c1 and c2 are the constants determining the outer boundary $b[j] = c1*(j/J)^{pp-1/2}$ and futility boundaries $a[j] = (c1 + c2) * (j/J)^{1/2} - c2 * (j/J)^{pp-1/2}$ for j >= j.star, where J is the max no of stages (AKA n.stg). If c1 is null, it is found as the upper alpha0 quantile of the max over the first j.star stages.
c2	see description in c1
pp	power parameter for Wang-Tsiatis boundaries; default is 0.4
n.stg	maximum number of analyses
j.star	earliest stage at which futility stopping is allowed. Should be <= n.stg-1 (there is already "futility stopping" at the n.stg-th stage anyway). Default is 1.
alpha	desired rejection probability of the test; default is 0.05
alpha0	the part of alpha that c1 is chosen to spend in first j.star stages; default is (j.star/n.stg) * alpha
mc.paths	matrix of sample paths, each row being a sample path, no. of columns is number of stages
inf.fraction	information fraction vector of the same length as n.stg which reflects the fraction of information at each analysis, should be positive, non-decreasing, and the last entry should be 1; default is (1:n.stg)/n.stg, user may want to specify a different vector for unequal time points
N.iter.max	max no. of iterations for finding c2
alpha.tol	the tolerance for stopping search for c2

Value

Returns a list:

a	the futility boundary vectors
b	the null-rejection boundary vectors
prej	prob. of rejecting the null (at any stage)

EM	expected stopping stage number
se.M	standard error of stopping time
c1	constants used in boundaries a, b
c2	constants used in boundaries a, b

Author(s)

Jay Bartroff

`bdr.gs.mc.gen` *Calculates the boundaries for the group sequential tests*

Description

Returns the boundaries for any group sequential test of the null vs. 2-sided alternative whose boundaries take the form of a single constant times a known weight vector, which is w.vec. These include Pocock (w.vec=(1,1,..)), O'Brien-Fleming (w.vec=(sqrt(n.stg/1), sqrt(n.stg/2), ..., 1)), etc. It does this by returning quantiles of the sample paths of the (null) test statistic paths in mc.paths.

Usage

```
bdr.gs.mc.gen(alpha = 0.05, mc.paths, w.vec)
```

Arguments

alpha	desired rejection probability of the test; default is 0.05
mc.paths	matrix of sample paths, each row being a sample path, no. of columns is number of stages
w.vec	weight vector corresponding to desired test

Value

Returns a list:

cons	the constant in the boundary vector cons*w.vec
bndry.vec	the boundary vector cons*w.vec

Author(s)

Jay Bartroff

`cov.surr.gs` *Computes variances and standardized covariance matrix for the group sequential statistic*

Description

Computes variances and standardized covariance matrix for the group sequential statistic

Usage

```
cov.surr.gs(s0.4.est, s1.4.est, sa.0, ya.0, nb.0, nb.1, full.matrix = TRUE,
naive = FALSE)
```

Arguments

`s0.4.est`	surrogate marker in the control group which is used for estimating means and covariances of S0, S1 in the Study B data. For designing tests (e.g., finding boundaries) these may come from Study A data, but for analyzing tests these may come from Study B data. Number of columns is the number of stages, number of rows may differ from rows in sa.0
`s1.4.est`	surrogate marker in the treated group which is used for estimating means and covariances of S0, S1 in the Study B data. For designing tests (e.g., finding boundaries) these may come from Study A data, but for analyzing tests these may come from Study B data. Number of columns is the number of stages, number of rows may differ from rows in sa.0
`sa.0`	surrogate marker in the control group in Study A
`ya.0`	primary outcome in the control group in Study A
`nb.0`	sample size for the control group in Study B
`nb.1`	sample size for the treated group in Study B
`full.matrix`	if TRUE, the standardized covariance matrix is provided; default is TRUE
`naive`	user should set to TRUE to compute covariance for "cumulative" test statistic, FALSE for naive statistic that only uses study B data from timepoint J at the J-th analysis; default is FALSE

Value

Returns a list:

`var.vec.del`	variance vector computed by the delta method
`cov.stand.del`	if full.matrix = TRUE, covariance matrix of the standardized test statistic computed by the delta method
`var.vec.samp`	variance vector computed by the sample mean and covariance of s0.4.est and s1.4.est
`cov.stand.samp`	if full.matrix = TRUE, covariance matrix of the standardized test statistic computed by the sample mean and covariance of s0.4.est and s1.4.est

Author(s)

Jay Bartroff

`delta.e.estimate` *Tests for a treatment effect on the primary outcome using surrogate marker information*

Description

Nonparametric test for a treatment effect on the primary outcome using surrogate marker information. This test borrows information from a prior study (Study A) about the relationship between the surrogate and the primary outcome to test for a treatment effect in the current study (Study B).

Usage

```
delta.e.estimate(sone = NULL, szero = NULL, szerop, yzerop, extrapolate = TRUE,
mat = NULL, n1 = NULL, n0 = NULL)
```

Arguments

`sone`	surrogate marker in the treated group in Study B
`szero`	surrogate marker in the control group in Study B
`szerop`	surrogate marker in the control group in Study A
`yzerop`	primary outcome in the control group in Study A
`extrapolate`	TRUE or FALSE; extrapolate for values outside of the support in Study A
`mat`	for Study B, the user can either provide sone and szero or can provide a vector, mat, where the first n1 values are the surrogate marker in the treated group in the Study B, and the remaining values are the surrogate marker in the control group in Study B
`n1`	sample size of treated group in Study B; only needed if mat is provided instead of sone and szero
`n0`	sample size of control group in Study B; only needed if mat is provided instead of sone and szero

Value

`delta.e`	estimated treatment effect using surrogate marker information
`sd.closed`	estimated standard error of treatment effect estimate
`delta.e.z`	test statistic
`delta.e.p`	p-value of test statistic

Author(s)

Layla Parast

References

Parast, Cai, and Tian (2023). Using a Surrogate with Heterogeneous Utility to Test for a Treatment Effect. Statistics in Medicine, 42(1): 68-88.

Parast and Bartroff (2024+). Group Sequential Testing of a Treatment Effect Using a Surrogate Marker. Under Review.

Examples

```
data(example.data)
delta.e.estimate(sone = example.data$s1, szero = example.data$s0, szerop = example.data$s0.p,
yzerop = example.data$y0.p)

data(StudyA.aids)
data(StudyB.aids)
s1.studyb = StudyB.aids$s1
s0.studyb = StudyB.aids$s0
s0.studya = StudyA.aids$s0

#24 weeks

delta.e.vec = delta.e.estimate(sone=s1.studyb$CD4_24weeks[!is.na(s1.studyb$CD4_24weeks)],
szero=s0.studyb$CD4_24weeks[!is.na(s0.studyb$CD4_24weeks)], szerop = s0.studya$CD4_24weeks,
yzerop = StudyA.aids$y0, extrapolate = TRUE)
delta.e.vec
```

---

example.data *Example data*

Description

Example data

Usage

```
data("example.data")
```

Format

A list with 9 elements:

`w0.p` the baseline covariate in the control group in the prior study (Study A)

`s0.p` the surrogate marker in the control group in the prior study (Study A

`y0.p` the primary outcome in the control group in the prior study (Study A

`w1` a baseline covariate in the treatment group in the current study (Study B)

`w0` a baseline covariate in the control group in the current study (Study B)

`s1` the surrogate marker in the treatment group in the current study (Study B)

`s0` the surrogate marker in the control group in the current study (Study B)

`y1` the primary outcome in the treatment group in the current study (Study B)

`y0` the primary outcome in the control group in the current study (Study B)

Examples

```
data(example.data)
names(example.data)
```

gs.boundaries *Computes group sequential boundaries*

Description

Computes group sequential (and naive) boundaries for the nonparametric test for a treatment effect on the primary outcome using surrogate marker information. The boundaries and test statistic borrow information from a prior study (Study A) about the relationship between the surrogate and the primary outcome to test for a treatment effect in the current study (Study B).

Usage

```
gs.boundaries(szerop, sonep, yzerop, nzero, none, n.stg, B.norm = 1e+06,
alpha = 0.05, pp = 0.4, inf.fraction = (1:n.stg)/n.stg, plot=FALSE)
```

Arguments

szerop	surrogate marker in the control group in Study A
sonep	surrogate marker in the treated group in Study A
yzerop	primary outcome in the control group in Study A
nzero	sample size of control group in Study B
none	sample size of treated group in Study B
n.stg	maximum number of analyses
B.norm	number of multivariate normal vectors to use in simulation for boundaries; default is 1e+06
alpha	desired rejection probability of the test; default is 0.05
pp	power parameter for Wang-Tsiatis boundaries; default is 0.4
inf.fraction	information fraction vector of the same length as n.stg which reflects the fraction of information at each analysis, should be positive, non-decreasing, and the last entry should be 1; default is (1:n.stg)/n.stg, user may want to specify a different vector for unequal time points
plot	TRUE or FALSE if a plot of the boundaries is desired; default is FALSE

Value

Returns a list of boundaries:

Naive	Naive boundaries
Bonf	Bonferroni boundaries
Pocock	Pocock boundaries
OBrien_Fleming	O'Brien-Fleming boundaries
Wang_Tsiatis	Wang-Tsiatis boundaries

Author(s)

Layla Parast and Jay Bartroff

References

Parast and Bartroff (2024+). Group Sequential Testing of a Treatment Effect Using a Surrogate Marker. Under Review.

Examples

```
data(example.data)
data(StudyA.aids)
data(StudyB.aids)
s0.studya = StudyA.aids$s0
s1.studya = StudyA.aids$s1

bound = gs.boundaries(szerop = s0.studya, sonep = s1.studya, yzerop=StudyA.aids$y0,
nzero = nrow(StudyB.aids$s0),none = nrow(StudyB.aids$s1), n.stg=4, B.norm=1e6,
alpha=0.05)

bound
```

`gs.boundaries.fut` *Computes group sequential boundaries with futility stopping*

Description

Computes group sequential (and naive) boundaries for the nonparametric test for a treatment effect on the primary outcome using surrogate marker information. The boundaries and test statistic borrow information from a prior study (Study A) about the relationship between the surrogate and the primary outcome to test for a treatment effect in the current study (Study B). The group sequential boundaries allow for futility stopping (bounds given).

Usage

```
gs.boundaries.fut(szerop, sonep, yzerop, nzero, none, n.stg, B.norm = 1e+06,
alpha = 0.05, pp = 0.4, inf.fraction = (1:n.stg)/n.stg, j.star=1,
alpha0=(j.star/n.stg)*alpha,
plot = FALSE)
```

Arguments

`szerop`	surrogate marker in the control group in Study A
`sonep`	surrogate marker in the treated group in Study A
`yzerop`	primary outcome in the control group in Study A
`nzero`	sample size of control group in Study B
`none`	sample size of treated group in Study B
`n.stg`	maximum number of analyses
`B.norm`	number of multivariate normal vectors to use in simulation for boundaries; default is 1e+06
`alpha`	desired rejection probability of the test; default is 0.05
`pp`	power parameter for Wang-Tsiatis boundaries; default is 0.4

inf.fraction	information fraction vector of the same length as n.stg which reflects the fraction of information at each analysis, should be positive, non-decreasing, and the last entry should be 1; default is (1:n.stg)/n.stg, user may want to specify a different vector for unequal time points
j.star	earliest stage at which futility stopping is allowed. Should be <= n.stg-1 (there is already "futility stopping" at the n.stg-th stage anyway). Default is 1.
alpha0	the part of alpha that c1 is chosen to spend in first j.star stages; default is (j.star/n.stg)*alpha
plot	TRUE or FALSE if a plot of the boundaries is desired; default is FALSE

Value

Returns a list of boundaries:

Naive	Naive boundaries
Bonf	Bonferroni boundaries
Pocock.futility	Pocock futility boundaries
Pocock.nullrejection	Pocock null rejection boundaries
OBrien_Fleming.futility	O'Brien-Fleming futility boundaries
OBrien_Fleming.nullrejection	O'Brien-Fleming null rejection boundaries
Wang_Tsiatis.futility	Wang-Tsiatis futility boundaries
Wang_Tsiatis.nullrejection	Wang-Tsiatis null rejection boundaries

Author(s)

Layla Parast and Jay Bartroff

References

Parast and Bartroff (2024+). Group Sequential Testing of a Treatment Effect Using a Surrogate Marker. Under Review.

Examples

```
data(example.data)
data(StudyA.aids)
data(StudyB.aids)
s0.studya = StudyA.aids$s0
s1.studya = StudyA.aids$s1

bound = gs.boundaries.fut(szerop = s0.studya, sonep = s1.studya, yzerop=StudyA.aids$y0,
nzero = nrow(StudyB.aids$s0),none = nrow(StudyB.aids$s1), n.stg=4, B.norm=1e6,
alpha=0.05)

bound
```

kern.estJ *Computes kernel density estimate*

Description

Computes kernel density estimate

Usage

```
kern.estJ(sb.arg, band.h, sa.vec, ya.vec)
```

Arguments

sb.arg	surrogate marker from Study B
band.h	bandwidth
sa.vec	surrogate marker from Study A
ya.vec	primary outcome from Study A

Value

kernel density estimate

Author(s)

Jay Bartroff

Kern.FUN *Calculates kernel matrix*

Description

Helper function; this calculates the kernel matrix

Usage

```
Kern.FUN(zz, zi, bw)
```

Arguments

zz	zz
zi	zi
bw	bandwidth

Value

the kernel matrix

Author(s)

Layla Parast

op.char.gs.fut — *Compute the operating characteristics on the group sequential test with futility stopping statistics in paths*

Description

Compute the operating characteristics on the group sequential test with futility stopping statistics in paths: The expected stopping stage no., plus the probability of rejecting the null in favor of the 2-sided alternative. This is for a general GS test which uses the boundaries in bndry.vec.

Usage

```
op.char.gs.fut(b.vec, a.vec, paths)
```

Arguments

b.vec	"null-rejection" boundaries, should be >= 0, and a.vec[n.stg] = b.vec[n.stg].
a.vec	futility boundaries, should be >= 0, and a.vec[n.stg] = b.vec[n.stg];a.vec[j]=0 means no futility stopping at stage j.
paths	matrix of test statistic sample paths, each row being a sample path, no. of columns is max number

Value

Returns a list:

EM	expected stopping stage number
se.M	standard error of stopping time
prej	prob. of rejecting the null (at any stage)

pred.smooth.2 — *Calculates the conditional mean function*

Description

Helper function; calculates the condition mean of Y given S, based on Study A data

Usage

```
pred.smooth.2(kernel.use,kernel.apply, bw,outcome)
```

Arguments

kernel.use	surrogate values in the control group in Study A
kernel.apply	surrogate values in Study B
bw	bandwidth
outcome	outcome in the control group in Study A

Value

expected outcome for each surrogate value

Author(s)

Layla Parast

`StudyA.aids` *ACTG 320 clinical trial data*

Description

Primary outcome and surrogate marker measurements over time from the ACTG 320 clinical trial data

Usage

```
data("StudyA.aids")
```

Format

A list with 4 elements:

`y1` the primary outcome in the treatment group in Study A; the primary outcome is defined as -1 times (log of RNA at 40 weeks - log of RNA at baseline) because a DECREASE in RNA is better

`y0` the primary outcome in the control group in Study A

`s1` a dataframe of the surrogate markers at different time points in the treatment group in Study A; the surrogate marker is change in CD4 cell count from baseline to 4 weeks (CD4_4weeks), 8 weeks (CD4_8weeks), 24 weeks (CD4_24weeks), and 40 weeks (CD4_40weeks). Note that higher values indicate increasing CD4 cell count which is "better".

`s0` a dataframe of the surrogate markers at different time points in the control group in Study A

Examples

```
data(StudyA.aids)
```

`StudyB.aids` *ACTG 193A clinical trial data*

Description

Surrogate marker measurements over time from the ACTG 193A clinical trial data. Note that the time points do not exactly match up to ACTG 320. In the paper, we use Study A surrogate data at 24 weeks to construct the conditional mean function applied to Study B at 16 weeks. Also note that some subjects are missing values of the surrogate at one or more time points. The naive estimate of the treatment effect using the surrogates uses all non-missing data available at each time point.

Usage

```
data("StudyB.aids")
```

Format

A list with 2 elements:

`s1` a dataframe of the surrogate markers at different time points in the treatment group in Study B; the surrogate marker is change in CD4 cell count from baseline to 8 weeks (CD4_8weeks), 16 weeks (CD4_16weeks), 24 weeks (CD4_24weeks), and 40 weeks (CD4_40weeks). Note that higher values indicate increasing CD4 cell count which is "better".

`s0` a dataframe of the surrogate markers at different time points in the control group in Study B

Examples

```
data(StudyB.aids)
```

VTM *Repeats a row.*

Description

Helper function; this function creates a matrix that repeats vc, dm times where each row is equal to the vc vector.

Usage

```
VTM(vc, dm)
```

Arguments

`vc`	the vector to repeat.
`dm`	number of rows.

Value

a matrix that repeats vc, dm times where each row is equal to the vc vector

Index