Package 'SPRT'

April 15, 2015

Type Package

Title Wald's Sequential Probability Ratio Test

Version 1.0

Date 2015-04-14

Author Stephane Mikael Bottine

Maintainer Stephane Mikael Bottine <stephane.bottine@gmail.com>

Description Perform Wald's Sequential Probability Ratio Test on variables with a Normal, Bernoulli, Exponential and Poisson distribution. Plot acceptance and continuation regions, or create your own with the help of closures.

License BSD_2_clause + file LICENSE

NeedsCompilation no

Repository CRAN

Date/Publication 2015-04-15 01:01:11

R topics documented:

SPRT-package *The SPRT package*

Description

Perform Wald's Sequential Probability Ratio Test on variables with a Normal, Bernoulli, Exponential and Poisson distribution. Plot acceptance and continuation regions, or create your own with the help of closures.

Details

Package:	SPRT
Type:	Package
Version:	1.0
Date:	2015-04-13
License:	BSD_3_clause + file LICENSE

Author(s)

Stephane Mikael Bottine

References

Ghosh, B.K. and Sen, P.K. (1991). *Handbook of Sequential Analysis*, Marcel Dekker, New York.
Wald, A. (1947). *Sequential Analysis*, Dover, New York.

Examples

```
# SPRT on a normally distributed random variable
set.seed(123)
test <- SPRT(distribution = "normal", type1 = 0.05, type2 = 0.20,
h0 = 0, h1 = 1, values = rnorm(10))

# Test outcome
test

# Cum. sum of the random variable vs H0 and H1 boundaries
test$data.sum
plot(test)

# Log-likelihood ratio vs Wald's A and B boundaries
test$data.llr
plot(test, log = "y")

# Log-likelihood ratio across scalars or vectors
```

```
set.seed(123)
test$llr.fn(n = 10, sum(rnorm(10)))

set.seed(123)
test$llr.fn(n = seq(1,10,1), k = cumsum(rnorm(10)))

# H0 and H1 boundaries
test$h0.fn(n = seq(1,10,1))
test$h1.fn(n = seq(1,10,1))
```

`boundary.fn` *boundary.fn internal function*

Description

Return a numerical 2x2 matrix with the slope and intercept of H0 and H1 acceptance regions under the null and alternative hypotheses.

Usage

```
boundary.fn(distribution = "bernoulli", type1 = 0.05, type2 = 0.2, h0, h1)
```

Arguments

`distribution`	a character string specifying the distribution. Must be one of `"bernoulli"` (default), `"normal"`, `"exponential"` or `"poisson"`.
`type1`	the type I error. A number between 0 and 1.
`type2`	the type II error. A number between 0 and 1.
`h0`	the expected value of the random variable under the null hypothesis.
`h1`	the expected value of the random variable under the alternative hypothesis.

Details

This function is an internal function and need not be accessed directly. Its output is available directly from `SPRT` under the `k.boundaries` name.

Author(s)

Stephane Mikael Bottine

References

Ghosh, B.K. and Sen, P.K. (1991). *Handbook of Sequential Analysis*, Marcel Dekker, New York.
Wald, A. (1947). *Sequential Analysis*, Dover, New York.

See Also

`SPRT`

C.fn *C.fn internal function*

Description

Return a component of the log-likelood ratio (LLR).

Usage

```
C.fn(distribution, h)
```

Arguments

distribution	a character string specifying the distribution. Must be one of "bernoulli" (default), "normal", "exponential" or "poisson".
h	an instance of either h0 or h1.

Details

This function is an internal function and need not be accessed directly.

Author(s)

Stephane Mikael Bottine

References

Ghosh, B.K. and Sen, P.K. (1991). *Handbook of Sequential Analysis*, Marcel Dekker, New York.
Wald, A. (1947). *Sequential Analysis*, Dover, New York.

D.fn *D.fn internal function*

Description

Return a component of the log-likelood ratio (LLR).

Usage

```
D.fn(distribution, h)
```

Arguments

distribution	a character string specifying the distribution. Must be one of "bernoulli" (default), "normal", "exponential" or "poisson".
h	an instance of either h0 or h1.

Details

This function is an internal function and need not be accessed directly.

Author(s)

Stephane Mikael Bottine

References

Ghosh, B.K. and Sen, P.K. (1991). *Handbook of Sequential Analysis*, Marcel Dekker, New York.
Wald, A. (1947). *Sequential Analysis*, Dover, New York.

`llr.fn` *LLR internal function*

Description

The log-likelihood ratio can be expressed as a function of `k` and `n`. This internal function returns a numeric vector with the coefficients of `k` and `n`.

Usage

```
llr.fn(distribution, h0, h1)
```

Arguments

`distribution`	a character string specifying the distribution. Must be one of `"bernoulli"` (default), `"normal"`, `"exponential"` or `"poisson"`.
`h0`	the expected value of the random variable under the null hypothesis.
`h1`	the expected value of the random variable under the alternative hypothesis.

Details

This function is an internal function and need not be accessed directly. Its output is available directly from `SPRT` under the `llr.coefficients` name.

Author(s)

Stephane Mikael Bottine

References

Ghosh, B.K. and Sen, P.K. (1991). *Handbook of Sequential Analysis*, Marcel Dekker, New York.
Wald, A. (1947). *Sequential Analysis*, Dover, New York.

See Also

`SPRT`

`plot.SPRT` *Plot SPRT boundaries and acceptance regions*

Description

Plot SPRT boundaries against the cumulative sum of the random variable (default) or the log-likelihood ratio (when you set `log = "y"`).

Usage

```
## S3 method for class 'SPRT'
plot(x = SPRT, y = NULL, type = NULL, xlim = NULL, ylim = NULL, log = "",
    main = "SPRT", sub = NULL, xlab = "Observations", ylab = NULL,
    ann = par("ann"), axes = TRUE, frame.plot = axes, panel.first = NULL,
    panel.last = NULL, asp = NULL, col = 1, lty = 2, lwd = 1,
    ...)
```

Arguments

`x`	an object with class `"SPRT"`; the output of the `SPRT` function.
`y`	the y coordinates of the points in the plot, automatically populated when an object with class `"SPRT"` is provided.
`type`	a 1-character string for the type of plot. The following values are possible: `"p"` for points (default), `"l"` for lines, `"b"` for points and lines, `"c"` for empty points joined by lines, `"o"` for overplotted points and lines, `"s"` and `"S"` for stair steps and `"h"` for histogram-like vertical lines. Finally, "n" hides all points.
`xlim`	numeric vector of length 2, giving the x coordinates ranges.
`ylim`	numeric vector of length 2, giving the y coordinates ranges.
`log`	set `log` equal to "y" to plot the random variable's log-likelihood ratio against the natural logarithm of Wald's A and B boundaries. Otherwise, the plot returns the cumulative sum of the random variable on the y-axis, against H1 and H0 acceptance boundaries.
`main`	an overall title for the plot: see `title`.
`sub`	a sub title for the plot: see `title`.
`xlab`	a title for the x axis: see `title`.
`ylab`	a title for the y axis: see `title`.
`ann`	a logical value indicating whether to show plot and axis titles. Defaults to the current setting.
`axes`	a logical value indicating whether both axes should be drawn on the plot.
`frame.plot`	a logical indicating whether a box should be drawn around the plot.
`panel.first`	an 'expression' to be evaluated after the plot axes are set up but before any plotting takes place. This can be useful for drawing background grids or scatterplot smooths. Note that this works by lazy evaluation: passing this argument from other `plot` methods may well not work since it may be evaluated too early.

panel.last	an expression to be evaluated after plotting has taken place but before the axes, title and box are added. See the comments about panel.first.
asp	a numeric scalar giving the aspect ratio y/x.
col	a vector up to 3 elements in length. The first colour is applied to the dots, the second is applied to the H0 boundary (or B boundary if applicable) and the third is applied to the H1 boundary (or A boundary if applicable). If the vector holds 2 colours, the first colour is applied to the dots and the second is applied across boundaries.
lty	the line type. Specify the line type either as an integer (0=blank, 1=solid (default), 2=dashed, 3=dotted, 4=dotdash, 5=longdash, 6=twodash) or as a character string ("blank", "solid", "dashed", "dotted", "dotdash", "longdash" or "twodash"). This parameter accepts a vector up to 2 elements in length. The first element is applied to the H0 boundary (or B boundary if applicable). The second element is applied to the H1 boundary (or A boundary if applicable).
lwd	the line width, a positive number, defaulting to 1. The interpretation is device-specific, and some devices do not implement line widths less than one. This parameter accepts a vector up to 2 elements in length. The first element is applied to the H0 boundary (or B boundary if applicable). The second element is applied to the H1 boundary (or A boundary if applicable).
...	other graphical parameters, such as cex, pch, family and font (see par).

Author(s)

Stephane Mikael Bottine

References

Ghosh, B.K. and Sen, P.K. (1991). *Handbook of Sequential Analysis*, Marcel Dekker, New York.
Wald, A. (1947). *Sequential Analysis*, Dover, New York.

See Also

par plot.default SPRT

Examples

```
## Run the SPRT() function
test <- SPRT(distribution = "normal", type1 = 0.05, type2 = 0.20,
  h0 = 0, h1 = 1, values = rnorm(10))

## Plot k vs. H0 and H1 boundaries
plot(test)

## Plot the random variable's log-likelihood ratio vs. Wald's A and B constants
plot(test, log = "y")

## Draw a line through through dots
plot(test, log = "y", type = "b")
```

```
## Change the size and appearance of the dots
plot(test, log = "y", pch = 20, cex = 1.25)

## Use difference colours for the dots and boundaries
plot(test, log = "y", pch = 20, cex = 1.25, col = c(1,2,3))
```

`print.SPRT` *Print SPRT boundaries and outcome*

Description

Returns a user-friendly summary of the test's outcome, including Wald boundaries.

Usage

```
## S3 method for class 'SPRT'
print(x = SPRT, ...)
```

Arguments

`x`	an object with class `"SPRT"`; the output of the `SPRT` function.
`...`	further arguments passed to or from other methods.

Author(s)

Stephane Mikael Bottine

References

Ghosh, B.K. and Sen, P.K. (1991). *Handbook of Sequential Analysis*, Marcel Dekker, New York.
Wald, A. (1947). *Sequential Analysis*, Dover, New York.

See Also

`print.default` `SPRT`

SPRT.default *Wald's Sequential Probability Ratio Test*

Description

Perform Wald's Sequential Probability Ratio Test on variables with a Normal, Bernoulli, Exponential or Poisson distribution. Returns an object with support for `print` and `plot` methods.

Usage

```
## Default S3 method:
SPRT(distribution = "bernoulli", type1 = 0.05, type2 = 0.2, h0, h1,
values = NULL, n = NULL, k = NULL)
```

Arguments

`distribution`	a character string specifying the distribution. Must be one of `"bernoulli"` (default), `"normal"`, `"exponential"` or `"poisson"`.
`type1`	the type I error. A number between 0 and 1.
`type2`	the type II error. A number between 0 and 1.
`h0`	the expected value of the random variable under the null hypothesis.
`h1`	the expected value of the random variable under the alternative hypothesis.
`values`	an optional vector containing values of the random variable. A logical vector when distribution is `"bernoulli"` and a numerical vector otherwise.
`n`	an optional numerical scalar for the number of observations of the random variable. `n` is optional and can be used as an alternative to `values`.
`k`	an optional numerical scalar for the cumulative sum of the random variable. `k` is optional and can be used as an alternative to `values`.

Details

Perform Wald's Sequential Probability Test on a simple hypothesis test of the null against the alternative.

The null hypothesis tested is that the expected value of the random variable is equal to `h0`. The alternative hypothesis tested is that the expected value of the random variable is equal to `h1`.

The expected value of the variable is the probability of success of a Bernoulli variable, the mean of a Normal variable, the mean (as well as the variance) of a Poisson variable, and the mean (as well as the standard deviation and scale parameter) of an Exponential distribution.

Optionally, specify `values`, a vector with observations of the random variable in the order in which they occurred. `values` is a logical vector of TRUE or FALSE observations in the case of a Bernoulli variable, and a numerical vector otherwise.

Or specify `n` and `k` as an alternative to `values`. `n` is the number of observations, and `k` is the cumulative sum of the random variable across observations (or the number of successes in the case of a Bernoulli variable). When `values` is given, SPRT infers `n` and `k`, and removes any `NA` values in the process.

Value

A list with class "SPRT" containing the following components:

distribution	equal to distribution.
n	equal to n if given, or the length of the values vector otherwise.
k	equal to k if given, or the cumulative sum of the values vector otherwise.
h0	equal to h0.
h1	equal to h1.
wald.A	the natural logarithm of Wald's A boundary (see waldBoundary).
wald.B	the natural logarithm of Wald's B boundary (see waldBoundary).
k.boundaries	a numerical 2x2 matrix with the slope and intercept of H0 and H1 acceptance regions under the null and alternative hypotheses.
llr.coefficients	a numerical vector of the n and k coefficients behind the random variable's log-likelihood ratio.
llr	a numerical scalar of the random variable's log-likelihood ratio.
decision	the outcome of the Sequential Probability Ratio Test. Returns FALSE when llr >= wald.A, TRUE when llr <= wald.B and NA otherwise.
interpretation	a character vector interpreting the outcome of the Sequential Probability Ratio Test. Returns "Accept H1" when decision is FALSE, "Accept H0" when decision is TRUE and "Continue testing" otherwise.
data.llr	a data frame comparing the random variable's log-likelihood ratio against the natural logarithm of Wald's A and B boundaries. The data frame's columns are named n, values, k, wald.B, wald.A and llr. It contains a row for every observation of the random variable. This output is only available when you specify values.
data.sum	a data frame comparing the random variable's cumulative sum, k, against acceptance boundaries for the null and alternative hypotheses. The data frame's columns are named n, values, k, h0 and h1. It contains a row for every observation of the random variable. This output is only available when you specify values.
llr.fn	a function that returns the value of the random variable's log-likelihood ratio for different n and k. This function is a closure; it encapsulates your distribution, type1, type2, h0 and h1 settings.
h0.fn	a function that returns the k acceptance boundary for the null hypothesis for different n. This function is a closure; it encapsulates your distribution, type1, type2, h0 and h1 settings.
h1.fn	a function that returns the k acceptance boundary for the alternative hypothesis for different n. This function is a closure; it encapsulates your distribution, type1, type2, h0 and h1 settings.

Note

This function returns an object with support for print and plot methods.

Author(s)

Stephane Mikael Bottine

References

Ghosh, B.K. and Sen, P.K. (1991). *Handbook of Sequential Analysis*, Marcel Dekker, New York.
Wald, A. (1947). *Sequential Analysis*, Dover, New York.

See Also

plot.SPRT, print.SPRT, waldBoundary

Examples

```
# SPRT on a normally distributed random variable
set.seed(123)
test <- SPRT(distribution = "normal", type1 = 0.05, type2 = 0.20,
  h0 = 0, h1 = 1, values = rnorm(10))

# Test outcome
test

# Cumulative sum of the random variable vs H0 and H1 boundaries
test$data.sum
plot(test)

# Sequential log-likelihood ratio vs Wald's A and B constants
test$data.llr
plot(test, log = "y")

# Calculate the log-likelihood ratio across scalars or vectors
set.seed(123)
test$llr.fn(n = 10, sum(rnorm(10)))

set.seed(123)
test$llr.fn(n = seq(1,10,1), k = cumsum(rnorm(10)))

# Calculate H0 and H1 boundaries
test$h0.fn(n = seq(1,10,1))
test$h1.fn(n = seq(1,10,1))
```

---

waldBoundary *Wald Boundaries*

Description

Calculate Wald's A and B boundaries, such that A and B satisfy $0 < B < A < Inf$. Note that the function returns the natural logarithm of the boundaries by default.

Usage

```
waldBoundary(type1 = 0.05, type2 = 0.2, boundary = NULL, log = TRUE)
```

Arguments

type1	type I error rate.
type2	type II error rate.
boundary	an optional character string specifying the boundary. Must be one of "A" or "B".
log	a logical variable indicating whether to return the boundary's natural logarithm.

Value

A numeric vector of length 1 if boundary is set to either "A" or "B", or a numeric vector of length 2 if boundary is set to NULL.

Author(s)

Stephane Mikael Bottine

References

Ghosh, B.K. and Sen, P.K. (1991). *Handbook of Sequential Analysis*, Marcel Dekker, New York.
Wald, A. (1947). *Sequential Analysis*, Dover, New York.

See Also

SPRT

Examples

```
## Return the logarithm of Wald's A and B boundaries
waldBoundary()

# Return Wald's A and B boundaries
waldBoundary(log = FALSE)

# Return only Wald's A boundary
waldBoundary(boundary = "A", log = FALSE)

## Boundary for different type I and II error rates
waldBoundary(type1 = 0.01, type2 = 0.01, boundary = "A", log = FALSE)
```

Index