

Package ‘harmonicmeanp’

January 7, 2019

Type Package

Title Harmonic Mean p-Values and Model Averaging by Mean Maximum Likelihood

Version 1.1

Date 2019-01-05

Author Daniel J. Wilson

Maintainer Daniel Wilson <hmp.R.package@gmail.com>

Depends FMStable, knitr, ape

Description The harmonic mean p-value (HMP) test combines p-values and corrects for multiple testing while controlling the family-wise error rate in a way that is more powerful than common alternatives including Bonferroni and Simes procedures, more stringent than controlling the false discovery rate, and is robust to positive correlations between tests and unequal weights. It is a multi-level test in the sense that a superset of one or more significant tests is almost certain to be significant and conversely when the superset is non-significant, the constituent tests are almost certain to be non-significant. It is based on MAMML (model averaging by mean maximum likelihood), a frequentist analogue to Bayesian model averaging, and is theoretically grounded in generalized central limit theorem. For detailed examples type `vignette("harmonicmeanp")` after installation.

License GPL-3

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2019-01-07 19:30:03 UTC

R topics documented:

harmonicmeanp	2
hmp.stat	3
mamml.stat	4
p.hmp	5
p.mamml	6
pharmonicmeanp	7
pLandau	8

Index**10**

harmonicmeanp	<i>Harmonic mean p-values and model averaging by mean maximum likelihood</i>
---------------	---

Description

The harmonic mean p -value (HMP) test combines p -values and corrects for multiple testing while controlling the family-wise error rate in a way that is more powerful than common alternatives including Bonferroni and Simes procedures, more stringent than controlling the false discovery rate, and is robust to positive correlations between tests and unequal weights. It is a multi-level test in the sense that a superset of one or more significant tests is almost certain to be significant and conversely when the superset is non-significant, the constituent tests are almost certain to be non-significant. It is based on MAMML (model averaging by mean maximum likelihood), a frequentist analogue to Bayesian model averaging, and is theoretically grounded in generalized central limit theorem.

Details

Package: harmonicmeanp
Type: Package
Version: 1.1
Date: 2019-01-05
License: GPL-3

The key function is `p.hmp` for combining p -values using the HMP. Type `vignette("harmonicmeanp")` for detailed examples.

Author(s)

Daniel J. Wilson

Maintainer: Daniel Wilson <hmp.R.package@gmail.com>

References

Daniel J. Wilson (2019) The harmonic mean p -value for combining dependent tests. *Proceedings of the National Academy of Sciences USA*, in press.

See Also

Package FMStable

Examples

```
# For detailed examples type vignette("harmonicmeanp")
```

```
# Example: simulate from a non-uniform distribution mildly enriched for small  $p$ -values.
# Compare the significance of the combined  $p$ -value for Bonferroni, Benjamini-Hochberg (i.e. Simes),
# HMP and (equivalently) MAMML with 2 degrees of freedom.
p = rbeta(1000,1/1.5,1)
min(p.adjust(p,"bonferroni"))
min(p.adjust(p,"BH"))
p.hmp(p)
p.mamml(1/p,2)
```

hmp.stat

Compute the harmonic mean p -value

Description

The harmonic mean p -value (HMP) is defined as the inverse of the (possibly weighted) arithmetic mean of the inverse p -values. When the HMP is small (e.g. less than 0.05), it is approximately well-calibrated, meaning that it can be directly interpreted. However, the function `p.hmp` calculates an exact p -value from the HMP and is preferred.

Usage

```
hmp.stat(p, w = NULL)
```

Arguments

<code>p</code>	A numeric vector of one or more p -values. Missing values (NAs) will cause a missing value to be returned.
<code>w</code>	An optional numeric vector of weights that can be interpreted as prior model probabilities for each of the alternative hypotheses represented by the individual p -values. The sum of the weights cannot exceed one but may be less than one, which is interpreted as meaning that some p -values have been excluded.

Value

The harmonic mean p -value is returned.

Author(s)

Daniel J. Wilson

References

Daniel J. Wilson (2019) The harmonic mean p -value for combining dependent tests. *Proceedings of the National Academy of Sciences USA*, in press.

See Also

`p.hmp`

Examples

```
# For detailed examples type vignette("harmonicmeanp")
p = rbeta(1000,1/1.5,1)
hmp.stat(p)
p.hmp(p)
```

mamml.stat

Compute the model-averaged mean maximized likelihood

Description

The model-averaged mean maximized likelihood (MAMML) is defined as the (possibly weighted) arithmetic mean of the maximized likelihood ratios from a series of likelihood ratio tests comparing mutually exclusive alternative hypotheses with the same nested null hypothesis based on the exact same data.

Usage

```
mamml.stat(R, w = NULL)
```

Arguments

R A numeric vector of one or more maximized likelihood ratios. Missing values (NAs) will cause a missing value to be returned.

w An optional numeric vector of weights that can be interpreted as prior model probabilities for each of the alternative hypotheses represented by the individual p -values. The sum of the weights cannot exceed one but may be less than one, which is interpreted as meaning that some p -values have been excluded.

Value

The model-averaged mean maximized likelihood ratio is returned.

Author(s)

Daniel J. Wilson

References

Daniel J. Wilson (2019) The harmonic mean p -value for combining dependent tests. *Proceedings of the National Academy of Sciences USA*, in press.

See Also

p.mamml

Examples

```
# For detailed examples type vignette("harmonicmeanp")
nu = 3
R = exp(0.5*rchisq(1000,nu))
mamml.stat(R)
p.mamml(R,nu)
```

p.hmp

Compute a combined p-value via the harmonic mean p-value

Description

The harmonic mean p -value (HMP) test combines p -values and corrects for multiple testing while controlling the family-wise error rate in a way that is more powerful than common alternatives including Bonferroni and Simes procedures, more stringent than controlling the false discovery rate, and is robust to positive correlations between tests and unequal weights. It is a multi-level test in the sense that a superset of one or more significant tests is almost certain to be significant and conversely when the superset is non-significant, the constituent tests are almost certain to be non-significant. It is based on MAMML (model averaging by mean maximum likelihood), a frequentist analogue to Bayesian model averaging, and is theoretically grounded in generalized central limit theorem.

Usage

```
p.hmp(p, w = NULL)
```

Arguments

p	A numeric vector of one or more p -values. Missing values (NAs) will cause a missing value to be returned.
w	An optional numeric vector of weights that can be interpreted as prior model probabilities for each of the alternative hypotheses represented by the individual p -values. The sum of the weights cannot exceed one but may be less than one, which is interpreted as meaning that some p -values have been excluded.

Value

A combined p -value is returned.

Author(s)

Daniel J. Wilson

References

Daniel J. Wilson (2019) The harmonic mean p -value for combining dependent tests. *Proceedings of the National Academy of Sciences USA*, in press.

See Also

hmp.stat

Examples

```
# For detailed examples type vignette("harmonicmeanp")
# Example: simulate from a non-uniform distribution mildly enriched for small \emph{p}-values.
# Compare the significance of the combined p-value for Bonferroni, Benjamini-Hochberg (i.e. Simes),
# HMP and (equivalently) MAMML with 2 degrees of freedom.
p = rbeta(1000,1/1.5,1)
min(p.adjust(p,"bonferroni"))
min(p.adjust(p,"BH"))
p.hmp(p)
p.mamml(1/p,2)
```

p.mamml	<i>Compute a combined p-value via the model-averaged mean maximized likelihood ratio</i>
---------	--

Description

The model averaging by mean maximum likelihood (MAMML) test combines likelihood ratio tests and corrects for multiple testing while controlling the family-wise error rate in a way that is more powerful than common alternatives including Bonferroni and Simes procedures, more stringent than controlling the false discovery rate, and is robust to positive correlations between tests and unequal weights. It is a frequentist analogue to Bayesian model averaging, is theoretically grounded in generalized central limit theorem, and motivates the simpler and better-calibrated harmonic mean p -value (HMP) test. The model-averaged mean maximized likelihood (MAMML) is defined as the (possibly weighted) arithmetic mean of the maximized likelihood ratios from a series of likelihood ratio tests comparing mutually exclusive alternative hypotheses with the same nested null hypothesis based on the exact same data.

Usage

```
p.mamml(R, nu, w = NULL)
```

Arguments

R	A numeric vector of one or more maximized likelihood ratios. Missing values (NAs) will cause a missing value to be returned.
nu	A numeric scalar or vector for the degrees of freedom corresponding to all or each of the maximized likelihood ratios respectively.
w	An optional numeric vector of weights that can be interpreted as prior model probabilities for each of the alternative hypotheses represented by the individual p -values. The sum of the weights cannot exceed one but may be less than one, which is interpreted as meaning that some p -values have been excluded.

Value

The model-averaged mean maximized likelihood ratio is returned.

Author(s)

Daniel J. Wilson

References

Daniel J. Wilson (2019) The harmonic mean p -value for combining dependent tests. *Proceedings of the National Academy of Sciences USA*, in press.

See Also

mamml.stat, hmp.stat, p.hmp

Examples

```
# For detailed examples type vignette("harmonicmeanp")
nu = 3
R = exp(0.5*rchisq(1000,nu))
mamml.stat(R)
p.mamml(R,nu)
```

pharmonicmeanp	<i>The (asymptotic) probability that a harmonic mean p-value (the test statistic of the HMP procedure) has value less than or equal to x.</i>
----------------	---

Description

The recommended usage of the HMP procedure is via the function `p.hmp` which internally computes the harmonic mean p -value test statistic and outputs an asymptotically exact p -value for the test statistic. If the test statistic (x) has already been computed, this function calculates the asymptotically exact cumulative distribution function ($\Pr(X < x)$), i.e. the p -value).

Usage

```
pharmonicmeanp(x, L, log=FALSE, lower.tail=TRUE)
```

Arguments

<code>x</code>	The previously computed harmonic mean p -value test statistic, for example using function <code>hmp.stat</code> .
<code>L</code>	The number of constituent p -values used in calculating x .
<code>log</code>	If true the log probability is output.
<code>lower.tail</code>	If true (the default) the lower tail probability is returned. Otherwise the upper tail probability.

Value

A tail probability is returned.

Author(s)

Daniel J. Wilson

References

Daniel J. Wilson (2019) The harmonic mean p -value for combining dependent tests. *Proceedings of the National Academy of Sciences USA*, in press.

See Also

p.hmp

Examples

```
# For detailed examples type vignette("harmonicmeanp")
# Example: simulate from a non-uniform distribution mildly enriched for small \emph{p}-values.
# Compare the significance of the combined p-value for Bonferroni, Benjamini-Hochberg (i.e. Simes),
# HMP and (equivalently) MAMML with 2 degrees of freedom.
p = rbeta(1000,1/1.5,1)
min(p.adjust(p,"bonferroni"))
min(p.adjust(p,"BH"))
x = hmp.stat(p)
pharmonicmeanp(x,length(p))
p.hmp(p)
p.mamml(1/p,2)
```

pLandau

The cumulative distribution function of a Landau distribution.

Description

If X has a Landau(μ,σ) distribution, the function produces the probability $\Pr(X < x)$ when `lower.tail=TRUE` or $\Pr(X > x)$ when `lower.tail=FALSE`. Read the cited paper for the parameterization of the distribution.

Usage

```
pLandau(x,mu=log(pi/2),sigma=pi/2,log=FALSE,lower.tail=TRUE)
```


Arguments

x	The value of a Landau-distributed random variable.
mu	The location parameter of the Landau distribution. Defaults to $\log(\pi/2)$ to give Landau's original distribution.
sigma	The scale parameter of the Landau distribution. Defaults to $\pi/2$ to give Landau's original distribution.
log	If true the log probability is output.
lower.tail	If true (the default) the lower tail probability is returned. Otherwise the upper tail probability.

Value

A tail probability is returned.

Author(s)

Daniel J. Wilson

References

Daniel J. Wilson (2019) The harmonic mean p -value for combining dependent tests. *Proceedings of the National Academy of Sciences USA*, in press.

See Also

p.hmp

Examples

```
# For detailed examples type vignette("harmonicmeanp")
# Example: simulate from a non-uniform distribution mildly enriched for small \emph{p}-values.
# Compare the significance of the combined p-value for Bonferroni, Benjamini-Hochberg (i.e. Simes),
# HMP and (equivalently) MAMML with 2 degrees of freedom.
p = rbeta(1000,1/1.5,1)
min(p.adjust(p,"bonferroni"))
min(p.adjust(p,"BH"))
x = hmp.stat(p)
pLandau(1/x,log(length(p))+(1 + digamma(1) - log(2/pi)),pi/2,lower.tail=FALSE)
p.hmp(p)
p.mamml(1/p,2)
```

Index

*Topic **\textasciitildeMAMML**

mamml.stat, 4

p.mamml, 6

*Topic **\textasciitildehmp**

hmp.stat, 3

p.hmp, 5

pharmonicmeanp, 7

pLandau, 8

harmonicmeanp, 2

hmp (harmonicmeanp), 2

hmp.stat, 3

mamml.stat, 4

p.hmp, 5

p.mamml, 6

pharmonicmeanp, 7

pLandau, 8