Package 'jocomo'

January 15, 2025

Type Package

Title Joint Comparison of Models

Version 0.2.0

Author Forrest C. Koch

Maintainer Forrest C. Koch <forrest.c.koch@gmail.com>

Description Provides extensions to McNemar's test which allows for the joint comparison of multiple predictive binary models.

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.3.1

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Depends R (>= 2.10)

R topics documented:

jocomo.chisq.test	*Title*

Description

Title

Title

Title

Title

Usage

```
jocomo.chisq.test(...)

## Default S3 method:
jocomo.chisq.test(
  x,
  y,
  samples,
  models,
  groups,
  correct = F,
  warn = getOption("warn"),
  ...
)

## S3 method for class 'formula'
jocomo.chisq.test(formula, data = parent.frame(), ...)

## S3 method for class 'xtabs'
jocomo.chisq.test(xt, data = parent.frame(), correct = FALSE, ...)
```

Arguments

`...`	TODO
`x`	TODO
`y`	TODO
`samples`	TODO
`models`	TODO
`groups`	TODO
`correct`	TODO
`warn`	TODO
`formula`	TODO
`data`	TODO
`xt`	TODO

Value

TODO

TODO

TODO

Examples

```
'TODO'
'TODO'
'TODO'
```

```
jocomo.ratio.test.default
```
Title

Description

Title

Usage

```
## Default S3 method:
jocomo.ratio.test(x, y, samples, models, groups)
```

Arguments

x	TODO
y	TODO
samples	TODO
models	TODO
groups	TODO

Value

TODO

Examples

```
'TODO'
```

```
jocomo.ratio.test.formula
```
Title

Description

Title

Usage

```
## S3 method for class 'formula'
jocomo.ratio.test(formula, data = parent.frame())
```

Arguments

formula	TODO
data	TODO

Value

TODO

Examples

```
'TODO'
```

```
jocomo.ratio.test.xtabs
```
Title

Description

Title

Usage

```
## S3 method for class 'xtabs'
jocomo.ratio.test(xt, data = parent.frame())
```

Arguments

xt	TODO
data	TODO

Value

TODO

Examples

```
'TODO'
```

```
multiclass.wu.statistic.default
```
Title

Description

Title

Usage

```
## Default S3 method:
multiclass.wu.statistic(x, y, correct = F, ...)
```

Arguments

x	An object which can be coerced to a 'matrix' of size $p * q$ where p is the number of samples and q is the number of models. The data should consist of two or more levels.
y	A 'vector' of length p where The data should consist of two or more levels.
correct	Add 0.5 to each cell of the 2x2 contingency table to adjust for 0 counts
...	Additional arguments passed on to methods.

```
multiclass.wu.statistic.xtabs
```
Title

Description

Title

Usage

```
## S3 method for class 'xtabs'
multiclass.wu.statistic(xt, correct = F, ...)
```

Arguments

xt	An 'xtabs' object of 3 or more dimensions indicating the cross-tabulation of model predictions and true labels. Each factor must have exactly two levels. The first dimension should refer to the true labels while the remaining dimensions refer to the model predictions.
correct	Add 0.5 to each cell of the 2x2 contingency table to adjust for 0 counts
...	Additional arguments passed on to methods.

---

multiclass.wu.test   *An extended version of Wu's Test allowing for multiclasses*

Description

An extended version of Wu's Test allowing for multiclasses

Title

Title

Title

An implementation of the extended McNemar statistic from Wu 2023 (doi: 10.1080/10543406.2022.2065500)

Usage

```
multiclass.wu.test(...)

## Default S3 method:
multiclass.wu.test(
  x,
  y,
  models,
  samples,
  correct = F,
  warn = getOption("warn"),
  ...
)

## S3 method for class 'formula'
multiclass.wu.test(formula, data = parent.frame(), ...)

## S3 method for class 'xtabs'
multiclass.wu.test(xt, ...)

multiclass.wu.statistic(...)
```

Arguments

...	Additional arguments passed on to methods.
x	Can be either a 'matrix', 'data.frame', or 'vector' of model predictions. As a 'matrix' or 'data.frame', 'x' should be a $p*q$ 'matrix' of binary predictions with p samples as rows and q models as columns. If 'x' is a 'vector', it should have length $p*q$ and both 'models' and 'samples' must be specified. The data must be able to be coerced to a 'factor' with two or more levels. Ignored if a formula is specified.
y	If 'x' is a 'matrix' or 'data.frame', then 'y' must be a 'vector' of length p indicating positive and negative cases. If 'x' is a 'vector', then 'y' must be a 'vector' of length $p*q$. The data must be able to be coerced to a 'factor' with the same levels as 'x'. Ignored if a formula is specified.

`models`	A 'vector' of length $p * q$ indicating which model the datum corresponds to. This should have q levels each occurring exactly p times. Ignored if 'x' is a 'matrix' or 'data.frame' or a formula is specified.
`samples`	A 'vector' of length $p * q$ indicating which subject the datum corresponds to. This should have p levels each occurring exactly q times. Ignored if 'x' is a 'matrix' or 'data.frame' or a formula is specified.
`correct`	TODO
`warn`	TODO
`formula`	A two-sided formula object describing predictions across multiple models and samples. Formulas may be specified in either wide, long, or cross-tabulated format. Refer to 'Details' for more information regarding formula specification.
`data`	an optional data frame containing the variables named in formula. By default the variables are taken from the environment from which 'multiclass.wu.test' is called. While data is optional, the package authors strongly recommend its use.
`xt`	An 'xtabs' object of 3 or more dimensions indicating the cross-tabulation of model predictions and true labels. Each factor must have the same number of levels. The first dimension should refer to the true labels while the remaining dimensions refer to the model predictions.

Details

Three formats are availabe for formula specification.

Wide format: The term on the left of the ~ operator should refer to a factor with two or more levels indicating the true labels for each subject. Terms should be separated by + operators and refer to predictions from each model.

Long format: The term on the left of the ~ operator should refer to a factor with two or more levels indicating the true labels for each subject. The formula should take the form 'x:y|z' where 'x' are the model predictions, 'y' describes the model which yielded the prediction, and 'z' describes the subject which the prediction is for.

Cross-tabulated format: The term on the left of the ~ operator should refer to the frequency, or total counts, for that stratum. Terms should be separated by + operators and refer to predictions from each model. An additional term, separated from the others by '||' should be included on the right hand side of the formula. e.g with 3 models, x1, x2, and x3, and true labels y, the formula should follow 'Freq ~ x1 + x2 + x3 || y'

`wu.statistic.default`	*Title*

Description

Title

Usage

```
## Default S3 method:
wu.statistic(x, y, correct = F, ...)
```

Arguments

x	An object which can be coerced to a 'matrix' of size $p * q$ where p is the number of samples and q is the number of models. The data should consist of two ordered levels and is coerced into a logical 'matrix'.
y	A 'vector' of length p where The data should consist of two ordered levels and is coerced into a logical 'vector'.
correct	Add 0.5 to each cell of the 2x2 contingency table to adjust for 0 counts
...	Additional arguments passed on to methods.

`wu.statistic.xtabs`	*Title*

Description

Title

Usage

```
## S3 method for class 'xtabs'
wu.statistic(xt, correct = F, ...)
```

Arguments

xt	An 'xtabs' object of 3 or more dimensions indicating the cross-tabulation of model predictions and true labels. Each factor must have exactly two levels. The first dimension should refer to the true labels while the remaining dimensions refer to the model predictions.
correct	Add 0.5 to each cell of the 2x2 contingency table to adjust for 0 counts
...	Additional arguments passed on to methods.

`wu.test`	*An implementation of Wu's test from Wu 2023 (doi: 10.1080/10543406.2022.2065500)*

Description

An implementation of Wu's test from Wu 2023 (doi: 10.1080/10543406.2022.2065500)

Title

Title

An implementation of the extended McNemar statistic from Wu 2023 (doi: 10.1080/10543406.2022.2065500)

Usage

```
wu.test(...)

## Default S3 method:
wu.test(x, y, models, samples, ...)

## S3 method for class 'formula'
wu.test(formula, data = parent.frame(), ...)

## S3 method for class 'xtabs'
wu.test(xt, ...)

wu.statistic(...)
```

Arguments

`...`	Additional arguments passed on to methods.
`x`	Can be either a 'matrix', 'data.frame', or 'vector' of model predictions. As a 'matrix' or 'data.frame', 'x' should be a $p * q$ 'matrix' of binary predictions with p samples as rows and q models as columns. If 'x' is a 'vector', it should have length $p * q$ and both 'models' and 'samples' must be specified. The data must be able to be coerced to a 'factor' with two levels. Ignored if a formula is specified.
`y`	If 'x' is a 'matrix' or 'data.frame', then 'y' must be a 'vector' of length p indicating positive and negative cases. If 'x' is a 'vector', then 'y' must be a 'vector' of length $p * q$. The data must be able to be coerced to a 'factor' with the same two levels as 'x'. Ignored if a formula is specified.
`models`	A 'vector' of length $p * q$ indicating which model the datum corresponds to. This should have q levels each occurring exactly p times. Ignored if 'x' is a 'matrix' or 'data.frame' or a formula is specified.
`samples`	A 'vector' of length $p * q$ indicating which subject the datum corresponds to. This should have p levels each occurring exactly q times. Ignored if 'x' is a 'matrix' or 'data.frame' or a formula is specified.
`formula`	A two-sided formula object describing predictions across multiple models and samples. Formulas may be specified in either wide, long, or cross-tabulated format. Refer to 'Details' for more information regarding formula specification.
`data`	an optional data frame containing the variables named in formula. By default the variables are taken from the environment from which 'wu.test' is called. While data is optional, the package authors strongly recommend its use.
`xt`	An 'xtabs' object of 3 or more dimensions indicating the cross-tabulation of model predictions and true labels. Each factor must have exactly two levels. The first dimension should refer to the true labels while the remaining dimensions refer to the model predictions.

Details

Three formats are availabe for formula specification.

Wide format: The term on the left of the ~ operator should refer to a factor with two levels indicating the true labels for each subject. Terms should be separated by + operators and refer to predictions from each model.

Long format: The term on the left of the ~ operator should refer to a factor with two levels indicating the true labels for each subject. The formula should take the form 'x:y|z' where 'x' are the model predictions, 'y' describes the model which yielded the prediction, and 'z' describes the subject which the prediction is for.

Cross-tabulated format: The term on the left of the ~ operator should refer to the frequency, or total counts, for that stratum. Terms should be separated by + operators and refer to predictions from each model. An additional term, separated from the others by '||' should be included on the right hand side of the formula. e.g with 3 models, x1, x2, and x3, and true labels y, the formula should follow 'Freq ~ x1 + x2 + x3 || y'

Index