Package 'automatedtests'

May 5, 2025

Type Package

Title Automating Choosing Statistical Tests

Version 0.1.0

Maintainer Wouter Zeevat <wouterzeevat@gmail.com>

Description A tool for simplifying the selection of appropriate statistical tests based on input data, making statistical workflows easier for users of all experience levels.

License GPL-3

Encoding UTF-8

URL https://github.com/wouterzeevat/automatedtests

BugReports https://github.com/wouterzeevat/automatedtests/issues

RoxygenNote 7.3.2

VignetteBuilder knitr

Suggests knitr, rmarkdown

Imports R6, nnet, nortest, stats, RVAideMemoire

Depends R (>= 4.0.4)

NeedsCompilation no

Author Wouter Zeevat [aut, cre]

Repository CRAN

Date/Publication 2025-05-05 10:10:12 UTC

Contents

AutomatedTest *AutomatedTest class*

Description

The AutomatedTest class represents a result of a statistical test. It contains attributes such as the p-value, degrees of freedom, and more.

Methods

Public methods:

- AutomatedTest$new()
- AutomatedTest$getData()
- AutomatedTest$isPaired()
- AutomatedTest$getIdentifiers()
- AutomatedTest$getCompareTo()
- AutomatedTest$setCompareTo()
- AutomatedTest$getDatatypes()
- AutomatedTest$getParametricList()
- AutomatedTest$isParametric()
- AutomatedTest$getTest()
- AutomatedTest$getResult()
- AutomatedTest$isSignificant()
- AutomatedTest$print()
- AutomatedTest$clone()

Method `new()`**:** Initialize an instance of the AutomatedTest class

Usage:

```
AutomatedTest$new(data, identifiers, compare_to = NULL, paired = FALSE)
```

Arguments:

`data` A dataframe containing the data for the test.

`identifiers` A vector with the identifiers.

`compare_to` Numeric value to compare to for comparison in one-sample tests. Default is NULL.

`paired` Logical; if TRUE, the test will be performed as paired if applicable. Default is FALSE.

Method `getData()`**:** Get the data used in the test

Usage:

```
AutomatedTest$getData()
```

Returns: A dataframe with all features

Method `isPaired()`**:** Shows if the data is paired, if there are multiple rows with the same identifier, the data has more samples (TIDY DATA). Making the data paired.

Usage:

```
AutomatedTest$isPaired()
```

Returns: Whether the data is paired (TRUE/FALSE).

Method `getIdentifiers()`**:** A list of the identifiers used for the data

Usage:

```
AutomatedTest$getIdentifiers()
```

Returns: Returns the identifiers

Method `getCompareTo()`**:** Get the comparison value for one-sample tests

Usage:

```
AutomatedTest$getCompareTo()
```

Returns: A numeric value for comparison

Method `setCompareTo()`**:** Updates the compare_to variable. Is public because the compare value can get changed depending on the type of test.

Usage:

```
AutomatedTest$setCompareTo(compare_to)
```

Arguments:

`compare_to` Numeric value to compare to.

Returns: Updated object with comparison value set.

Method `getDatatypes()`**:** Get the data types of the features in the object

Usage:

```
AutomatedTest$getDatatypes()
```

Returns: A list of data types (e.g., Quantitative or Qualitative)

Method `getParametricList()`**:** Get the parametric test results of the features

Usage:

```
AutomatedTest$getParametricList()
```

Returns: A list of parametric test results

Method `isParametric()`**:** Check if the data meets parametric assumptions

Usage:

```
AutomatedTest$isParametric()
```

Returns: TRUE if parametric assumptions are met, otherwise FALSE

Method `getTest()`**:** Get the statistical test that was chosen

Usage:

```
AutomatedTest$getTest()
```

Returns: The name of the statistical test

Method `getResult()`**:** Get the result of selected statistical test

Usage:

```
AutomatedTest$getResult()
```

Returns: The result of the statistical test

Method `isSignificant()`**:** Whether the test results are significant or not.

Usage:

```
AutomatedTest$isSignificant()
```

Returns: TRUE / FALSE depending on the significance of the test.

Method `print()`**:** Print a summary of the test object

Usage:

```
AutomatedTest$print()
```

Method `clone()`**:** The objects of this class are cloneable with this method.

Usage:

```
AutomatedTest$clone(deep = FALSE)
```

Arguments:

`deep` Whether to make a deep clone.

`automatical_test` *Automatically Run a Statistical Test*

Description

Automatically choose the best fitting statistical test for your data, and returns an easily readable `AutomatedTest` object from either a data frame or individual vectors. This object contains the executed test together with all statistics and properties.

Usage

```
automatical_test(..., compare_to = NULL, identifiers = FALSE, paired = FALSE)
```

Arguments

`...`	Either a single data frame or multiple equal-length vectors representing columns of data.
`compare_to`	A numeric value to compare against during a one-sample test. If the data is categorical, the value will default to $1/k$, where k is the number of categories, assuming a uniform distribution. If numeric, the default will be 0.
`identifiers`	Logical; if TRUE, the first column/vector is treated as identifiers and excluded from testing.
`paired`	Logical; if TRUE, the test will be performed as paired if applicable, regardless of whether identifiers are provided. This applies to paired tests like McNemar's or the Cochran Q test.

Details

The `automatical_test` function automatically selects and runs the most fitting statistical test based on the data provided. It can accept data as either a single data frame or multiple individual vectors, provided the vectors are of equal length.

If `identifiers` is set to TRUE, the first column will be treated as identifiers and excluded from the test, supporting TIDY data.

When a multiple group test is selected (i.e., more than two groups, columns, or variables are used), the first non-identifier column will be used as the grouping or target variable, meaning all other variables will be tested against it.

The `paired` parameter can be used to force paired testing for supported tests (such as McNemar's test or Cochran's Q), even if identifiers are not explicitly included in the input.

If you want to override the defaults, you can change the `compare_to` value to specify one-sample tests.

Once the test has been executed, you can use the method `$getResult()` on the resulting object to get more detailed information about the test's execution, including a summary of the test used and all statistics.

Value

An object of class `AutomatedTest`. The object contains the results of the statistical test performed on the data. You can use the method `$getResult()` to obtain more detailed information about the execution of the test.

Author(s)

Wouter Zeevat

See Also

`AutomatedTest` for the class used by this function.

Examples

```
# Example 1: Using individual vectors
test1 <- automatical_test(iris$Species, iris$Sepal.Length, identifiers = FALSE)

# Example 2: Forcing a paired test
before <- c(200, 220, 215, 205, 210)
after <- c(202, 225, 220, 210, 215)
paired_data <- data.frame(before, after)
test2 <- automatical_test(before, after, paired = TRUE)

# Retrieve more detailed information about the test
# test1$getResult()
```

Index