Package: nortest (via r-universe)

August 24, 2024

Title Tests for Normality

Version 1.0-4

Date 2015-07-29

Description Five omnibus tests for testing the composite hypothesis of
normality.

License GPL (>= 2)

Imports stats

Author Juergen Gross [aut], Uwe Ligges [aut, cre]

Maintainer Uwe Ligges <ligges@statistik.tu-dortmund.de>

NeedsCompilation no

Date/Publication 2015-07-30 00:14:57

Repository https://uligges.r-universe.dev

RemoteUrl https://github.com/cran/nortest

RemoteRef HEAD

RemoteSha a558613013b841d217ffdb89b59476b8384f5d63

Contents

---

ad.test   *Anderson-Darling test for normality*

Description

Performs the Anderson-Darling test for the composite hypothesis of normality, see e.g. Thode (2002, Sec. 5.1.4).

Usage

```
ad.test(x)
```

Arguments

x   a numeric vector of data values, the number of which must be greater than 7. Missing values are allowed.

Details

The Anderson-Darling test is an EDF omnibus test for the composite hypothesis of normality. The test statistic is

$$A = -n - \frac{1}{n} \sum_{i=1}^{n} [2i - 1][\ln(p_{(i)}) + \ln(1 - p_{(n-i+1)})],$$

where $p_{(i)} = \Phi([x_{(i)} - \overline{x}]/s)$. Here, Φ is the cumulative distribution function of the standard normal distribution, and $\overline{x}$ and s are mean and standard deviation of the data values. The p-value is computed from the modified statistic $Z = A(1.0 + 0.75/n + 2.25/n^2)\backslash$ according to Table 4.9 in Stephens (1986).

Value

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

statistic   the value of the Anderson-Darling statistic.

p.value   the p-value for the test.

method   the character string "Anderson-Darling normality test".

data.name   a character string giving the name(s) of the data.

Note

The Anderson-Darling test is the recommended EDF test by Stephens (1986). Compared to the Cramer-von Mises test (as second choice) it gives more weight to the tails of the distribution.

Author(s)

Juergen Gross

References

Stephens, M.A. (1986): Tests based on EDF statistics. In: D'Agostino, R.B. and Stephens, M.A., eds.: Goodness-of-Fit Techniques. Marcel Dekker, New York.

Thode Jr., H.C. (2002): Testing for Normality. Marcel Dekker, New York.

See Also

shapiro.test for performing the Shapiro-Wilk test for normality. cvm.test, lillie.test, pearson.test, sf.test for performing further tests for normality. qqnorm for producing a normal quantile-quantile plot.

Examples

```
ad.test(rnorm(100, mean = 5, sd = 3))
ad.test(runif(100, min = 2, max = 4))
```

---

cvm.test   *Cramer-von Mises test for normality*

Description

Performs the Cramer-von Mises test for the composite hypothesis of normality, see e.g. Thode (2002, Sec. 5.1.3).

Usage

```
cvm.test(x)
```

Arguments

x   a numeric vector of data values, the number of which must be greater than 7. Missing values are allowed.

Details

The Cramer-von Mises test is an EDF omnibus test for the composite hypothesis of normality. The test statistic is

$$W = \frac{1}{12n} + \sum_{i=1}^{n} \left(p_{(i)} - \frac{2i-1}{2n} \right)^2,$$

where $p_{(i)} = \Phi([x_{(i)} - \overline{x}]/s)$. Here, Φ is the cumulative distribution function of the standard normal distribution, and $\overline{x}$ and s are mean and standard deviation of the data values. The p-value is computed from the modified statistic $Z = W(1.0 + 0.5/n)$ according to Table 4.9 in Stephens (1986).

Value

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

statistic	the value of the Cramer-von Mises statistic.
p.value	the p-value for the test.
method	the character string "Cramer-von Mises normality test".
data.name	a character string giving the name(s) of the data.

Author(s)

Juergen Gross

References

Stephens, M.A. (1986): Tests based on EDF statistics. In: D'Agostino, R.B. and Stephens, M.A., eds.: Goodness-of-Fit Techniques. Marcel Dekker, New York.

Thode Jr., H.C. (2002): Testing for Normality. Marcel Dekker, New York.

See Also

shapiro.test for performing the Shapiro-Wilk test for normality. ad.test, lillie.test, pearson.test, sf.test for performing further tests for normality. qqnorm for producing a normal quantile-quantile plot.

Examples

```
cvm.test(rnorm(100, mean = 5, sd = 3))
cvm.test(runif(100, min = 2, max = 4))
```

---

lillie.test   *Lilliefors (Kolmogorov-Smirnov) test for normality*

Description

Performs the Lilliefors (Kolmogorov-Smirnov) test for the composite hypothesis of normality, see e.g. Thode (2002, Sec. 5.1.1).

Usage

```
lillie.test(x)
```

Arguments

x	a numeric vector of data values, the number of which must be greater than 4. Missing values are allowed.

Details

The Lilliefors (Kolmogorov-Smirnov) test is an EDF omnibus test for the composite hypothesis of normality. The test statistic is the maximal absolute difference between empirical and hypothetical cumulative distribution function. It may be computed as $D = \max\{D^+, D^-\}$ with

$$D^+ = \max_{i=1,...,n} \{i/n - p_{(i)}\}, D^- = \max_{i=1,...,n} \{p_{(i)} - (i-1)/n\},$$

where $p_{(i)} = \Phi([x_{(i)} - \overline{x}]/s)$. Here, Φ is the cumulative distribution function of the standard normal distribution, and $\overline{x}$ and s are mean and standard deviation of the data values. The p-value is computed from the Dallal-Wilkinson (1986) formula, which is claimed to be only reliable when the p-value is smaller than 0.1. If the Dallal-Wilkinson p-value turns out to be greater than 0.1, then the p-value is computed from the distribution of the modified statistic $Z = D(\sqrt{n} - 0.01 + 0.85/\sqrt{n})$, see Stephens (1974), the actual p-value formula being obtained by a simulation and approximation process.

Value

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

`statistic`	the value of the Lilliefors (Kolomogorv-Smirnov) statistic.
`p.value`	the p-value for the test.
`method`	the character string "Lilliefors (Kolmogorov-Smirnov) normality test".
`data.name`	a character string giving the name(s) of the data.

Note

The Lilliefors (Kolomorov-Smirnov) test is the most famous EDF omnibus test for normality. Compared to the Anderson-Darling test and the Cramer-von Mises test it is known to perform worse. Although the test statistic obtained from `lillie.test(x)` is the same as that obtained from `ks.test(x, "pnorm", mean(x), sd(x))`, it is not correct to use the p-value from the latter for the composite hypothesis of normality (mean and variance unknown), since the distribution of the test statistic is different when the parameters are estimated.

The function call `lillie.test(x)` essentially produces the same result as the S-PLUS function call `ks.gof(x)` with the distinction that the p-value is not set to 0.5 when the Dallal-Wilkinson approximation yields a p-value greater than 0.1. (Actually, the alternative p-value approximation is provided for the complete range of test statistic values, but is only used when the Dallal-Wilkinson approximation fails.)

Author(s)

Juergen Gross

References

Dallal, G.E. and Wilkinson, L. (1986): An analytic approximation to the distribution of Lilliefors' test for normality. The American Statistician, 40, 294–296.

Stephens, M.A. (1974): EDF statistics for goodness of fit and some comparisons. Journal of the American Statistical Association, 69, 730–737.

Thode Jr., H.C. (2002): Testing for Normality. Marcel Dekker, New York.

See Also

shapiro.test for performing the Shapiro-Wilk test for normality. ad.test, cvm.test, pearson.test, sf.test for performing further tests for normality. qqnorm for producing a normal quantile-quantile plot.

Examples

```
lillie.test(rnorm(100, mean = 5, sd = 3))
lillie.test(runif(100, min = 2, max = 4))
```

---

pearson.test	*Pearson chi-square test for normality*

Description

Performs the Pearson chi-square test for the composite hypothesis of normality, see e.g. Thode (2002, Sec. 5.2).

Usage

```
pearson.test(x, n.classes = ceiling(2 * (n^(2/5))), adjust = TRUE)
```

Arguments

x	a numeric vector of data values. Missing values are allowed.
n.classes	The number of classes. The default is due to Moore (1986).
adjust	logical; if TRUE (default), the p-value is computed from a chi-square distribution with n.classes-3 degrees of freedom, otherwise from a chi-square distribution with n.classes-1 degrees of freedom.

Details

The Pearson test statistic is $P = \sum (C_i - E_i)^2/E_i$, where C_i is the number of counted and E_i is the number of expected observations (under the hypothesis) in class i. The classes are build is such a way that they are equiprobable under the hypothesis of normality. The p-value is computed from a chi-square distribution with n.classes-3 degrees of freedom if adjust is TRUE and from a chi-square distribution with n.classes-1 degrees of freedom otherwise. In both cases this is not (!) the correct p-value, lying somewhere between the two, see also Moore (1986).

Value

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

statistic	the value of the Pearson chi-square statistic.
p.value	the p-value for the test.
method	the character string "Pearson chi-square normality test".

data.name	a character string giving the name(s) of the data.
n.classes	the number of classes used for the test.
df	the degress of freedom of the chi-square distribution used to compute the p-value.

Note

The Pearson chi-square test is usually not recommended for testing the composite hypothesis of normality due to its inferior power properties compared to other tests. It is common practice to compute the p-value from the chi-square distribution with n.classes - 3 degrees of freedom, in order to adjust for the additional estimation of two parameters. (For the simple hypothesis of normality (mean and variance known) the test statistic is asymptotically chi-square distributed with n.classes - 1 degrees of freedom.) This is, however, not correct as long as the parameters are estimated by mean(x) and var(x) (or sd(x)), as it is usually done, see Moore (1986) for details. Since the true p-value is somewhere between the two, it is suggested to run pearson.test twice, with adjust = TRUE (default) and with adjust = FALSE. It is also suggested to slightly change the default number of classes, in order to see the effect on the p-value. Eventually, it is suggested not to rely upon the result of the test.

The function call pearson.test(x) essentially produces the same result as the S-PLUS function call chisq.gof((x-mean(x))/sqrt(var(x)), n.param.est=2).

Author(s)

Juergen Gross

References

Moore, D.S. (1986): Tests of the chi-squared type. In: D'Agostino, R.B. and Stephens, M.A., eds.: Goodness-of-Fit Techniques. Marcel Dekker, New York.

Thode Jr., H.C. (2002): Testing for Normality. Marcel Dekker, New York.

See Also

shapiro.test for performing the Shapiro-Wilk test for normality. ad.test, cvm.test, lillie.test, sf.test for performing further tests for normality. qqnorm for producing a normal quantile-quantile plot.

Examples

```
pearson.test(rnorm(100, mean = 5, sd = 3))
pearson.test(runif(100, min = 2, max = 4))
```

---

sf.test   *Shapiro-Francia test for normality*

Description

Performs the Shapiro-Francia test for the composite hypothesis of normality, see e.g. Thode (2002, Sec. 2.3.2).

Usage

```
sf.test(x)
```

Arguments

x   a numeric vector of data values, the number of which must be between 5 and 5000. Missing values are allowed.

Details

The test statistic of the Shapiro-Francia test is simply the squared correlation between the ordered sample values and the (approximated) expected ordered quantiles from the standard normal distribution. The p-value is computed from the formula given by Royston (1993).

Value

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

statistic   the value of the Shapiro-Francia statistic.

p.value   the p-value for the test.

method   the character string "Shapiro-Francia normality test".

data.name   a character string giving the name(s) of the data.

Note

The Shapiro-Francia test is known to perform well, see also the comments by Royston (1993). The expected ordered quantiles from the standard normal distribution are approximated by qnorm(ppoints(x, a = 3/8)), being slightly different from the approximation qnorm(ppoints(x, a = 1/2)) used for the normal quantile-quantile plot by qqnorm for sample sizes greater than 10.

Author(s)

Juergen Gross

References

Royston, P. (1993): A pocket-calculator algorithm for the Shapiro-Francia test for non-normality: an application to medicine. Statistics in Medicine, 12, 181–184.

Thode Jr., H.C. (2002): Testing for Normality. Marcel Dekker, New York.

See Also

shapiro.test for performing the Shapiro-Wilk test for normality. ad.test, cvm.test, lillie.test, pearson.test for performing further tests for normality. qqnorm for producing a normal quantile-quantile plot.

Examples

```
sf.test(rnorm(100, mean = 5, sd = 3))
sf.test(runif(100, min = 2, max = 4))
```

Index