

# Package ‘DistributionTest’

February 26, 2020

**Type** Package

**Title** Powerful Goodness-of-Fit Tests Based on the Likelihood Ratio

**Version** 1.1

**Date** 2020-02-26

**Author** Ning Cui [aut, cre], Maoyuan Zhou [ctb]

**Maintainer** Ning Cui <2433971953@qq.com>

**Description** Provides new types of omnibus tests which are generally much more powerful than traditional tests (including the Kolmogorov-Smirnov, Cramer-von Mises and Anderson-Darling tests), see Zhang (2002) <doi:10.1111/1467-9868.00337>.

**License** GPL (>= 3)

**Imports** stats, MASS

**Repository** CRAN

**Encoding** UTF-8

**LazyData** true

**NeedsCompilation** no

**Date/Publication** 2020-02-26 05:20:03 UTC

## R topics documented:

za.test . . . . .	2
zc.test . . . . .	3
zk.test . . . . .	5
<b>Index</b>	<b>8</b>

---

za.test	<i>perform a one- or two-sample analogue ZA of Anderson-Darling statistic</i>
---------	---

---

### Description

The new statistics ZA appear similar to the Anderson-Darling, but it's generally much more powerful, see Jin Zhang(2002).

### Usage

```
za.test(x, y, para = NULL, N = 1000)
```

### Arguments

x	a numeric vector of data values, the number of which must be greater than 7. Missing values are allowed.
y	When tested as a single sample, a numeric vector of data values, the number of which must be greater than 7. Missing values are allowed. When tested as two-sample, a character string indicating what type of test was performed. Distributions "unif", "exponential", "normal", "lognormal", "gamma", "t" and "weibull" are recognised. Here, the names of "exponential", "normal" and "lognormal" are simplified as "exp", "norm" and "lognorm" respectively.
para	A named list giving the parameters of the distribution specified and this can be omitted.
N	The number of replications in Monte Carlo simulation. The default value is 1000

### Details

The ZA test is an EDF omnibus test for the composite hypothesis of distribution. The test statistic is

$$Za = - \sum_{i=1}^n \left[ \frac{\ln[F_0(X_{(i)})]}{n - i + 0.5} + \frac{\ln[1 - F_0(X_{(i)})]}{i - 0.5} \right],$$

where  $F_0(x)$  is a hypothesized distribution function to be tested. Here,  $F_0(X_{(i)}) = \Phi(x)$ ,  $\Phi$  is the cumulative distribution function of the specified distribution. The p-value is computed by Monte Carlo simulation.

### Value

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

statistic	the value of the ZA statistic.
p.value	the p-value for the test.
method	the character string "ZA test for given distribution".
data.name	a character string giving the name(s) of the data.

**Note**

The Za test is the recommended EDF test by Jin Zhang.

**Author(s)**

Ning Cui

**References**

Jin Zhang: Goodness-of-Fit Tests Based on the Likelihood Ratio. *Journal of the Royal Statistical Society*, 64, 281-294.

Jin Zhang, Yuehua Wu: Likelihood-ratio tests for normality. *Computational Statistics & Data Analysis*, 49, 709-721.

Jin Zhang: Powerful Two-Sample Tests Based on the Likelihood Ratio. *Technometrics*, 48:1, 95-103.

**See Also**

[ks.test](#) for performing a one- or two-sample Kolmogorov-Smirnov test. [zc.test](#), [zk.test](#) for performing a powerful goodness-of-fit test based on the likelihood ratio.

**Examples**

```
x<-rbeta(50,shape1 = 0.6,shape2 = 0.8)
y<-rnorm(50)
za.test(x,y)
za.test(x,"unif")
za.test(x,"norm")
za.test(x,"unif",para = list(min=1,max=2))
za.test(x,"exp",para = list(rate=1))
za.test(x,"norm",para = list(mean=1,sd=2))
za.test(x,"lognorm",para = list(mean=1,sd=2))
za.test(x,"weibull",para = list(shape=1,scale=2))
za.test(x,"gamma",para = list(shape=2,scale=1))
za.test(x,"t",para = list(df=3))
```

---

zc.test	<i>perform a one- or two-sample analogue ZC of the Cramer-von Mises statistic</i>
---------	---

---

**Description**

The new statistics ZC appear similar to the Cramer-von Mises statistic, but it's generally much more powerful, see Jin Zhang(2002).

**Usage**

```
zc.test(x, y, para = NULL, N = 1000)
```

**Arguments**

x	a numeric vector of data values, the number of which must be greater than 7. Missing values are allowed.
y	When tested as a single sample, a numeric vector of data values, the number of which must be greater than 7. Missing values are allowed. When tested as two-sample, a character string indicating what type of test was performed. Distributions "unif", "exponential", "normal", "lognormal", "gamma", "t" and "weibull" are recognised. Here, the names of "exponential", "normal" and "lognormal" are simplified as "exp", "norm" and "lognorm" respectively.
para	A named list giving the parameters of the distribution specified and this can be omitted.
N	The number of replications in Monte Carlo simulation. The default value is 1000

**Details**

The ZC test is an EDF omnibus test for the composite hypothesis of distribution. The test statistic is

$$Zc = \sum_{i=1}^n \left[ \ln \frac{F_0(X_{(i)})^{-1} - 1}{(n - 0.5)/(i - 0.75) - 1} \right],$$

where  $F_0(x)$  is a hypothesized distribution function to be tested. Here,  $F_0(X_{(i)}) = \Phi(x)$ ,  $\Phi$  is the cumulative distribution function of the specified distribution. The p-value is computed by Monte Carlo simulation.

**Value**

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

statistic	the value of the ZC statistic.
p.value	the p-value for the test.
method	the character string "ZC test for given distribution".
data.name	a character string giving the name(s) of the data.

**Note**

The ZC test is the recommended EDF test by Jin Zhang.

**Author(s)**

Ning Cui

**References**

- Jin Zhang: Goodness-of-Fit Tests Based on the Likelihood Ratio. *Journal of the Royal Statistical Society*, 64, 281-294.
- Jin Zhang, Yuehua Wu: Likelihood-ratio tests for normality. *Computational Statistics & Data Analysis*, 49, 709-721.
- Jin Zhang: Powerful Two-Sample Tests Based on the Likelihood Ratio. *Technometrics*, 48:1, 95-103.

**See Also**

[ks.test](#) for performing a one- or two-sample Kolmogorov-Smirnov test. [za.test](#), [zk.test](#) for performing a powerful goodness-of-fit test based on the likelihood ratio.

**Examples**

```
x<-rbeta(50,shape1 = 0.6,shape2 = 0.8)
y<-rnorm(50)
zc.test(x,y)
zc.test(x,"unif")
zc.test(x,"norm")
zc.test(x,"unif",para = list(min=1,max=2))
zc.test(x,"exp",para = list(rate=1))
zc.test(x,"norm",para = list(mean=1,sd=2))
zc.test(x,"lognorm",para = list(mean=1,sd=2))
zc.test(x,"weibull",para = list(shape=1,scale=2))
zc.test(x,"gamma",para = list(shape=2,scale=1))
zc.test(x,"t",para = list(df=3))
```

---

zk.test

*The analogue ZK of the Kolmogorov-Smirnov statistic*


---

**Description**

The new statistics ZK appear similar to the Kolmogorov-Smirnov statistic, but it's generally much more powerful, see Jin Zhang(2002).

**Usage**

```
zk.test(x, y, para = NULL, N = 1000)
```

**Arguments**

- |      |   |
|------|---|
| x    | a numeric vector of data values, the number of which must be greater than 7. Missing values are allowed.  |
| y    | When tested as a single sample, a numeric vector of data values, the number of which must be greater than 7. Missing values are allowed.<br>When tested as two-sample, a character string indicating what type of test was performed. Distributions "unif", "exponential", "normal", "lognormal", "gamma", "t" and "weibull" are recognised. Here, the names of "exponential", "normal" and "lognormal" are simplified as "exp", "norm" and "lognorm" respectively. |
| para | A named list giving the parameters of the distribution specified and this can be omitted.   |
| N    | The number of replications in Monte Carlo simulation. The default value is 1000   |

### Details

The Zk test is an EDF omnibus test for the composite hypothesis of distribution. The test statistic is

$$Z_k = \max_{1 \leq i \leq n} \left\{ (i - 0.5) \ln \frac{i - 0.5}{n F_0(X_{(i)})} + (n - i + 0.5) \ln \frac{n - i + 0.5}{n[1 - F_0(X_{(i)})]} \right\},$$

where  $F_0(x)$  is a hypothesized distribution function to be tested. Here,  $F_0(X_{(i)}) = \Phi(x)$ ,  $\Phi$  is the cumulative distribution function of the specified distribution. The p-value is computed by Monte Carlo simulation.

### Value

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

statistic	the value of the ZK statistic.
p.value	the p-value for the test.
method	the character string "ZK test for given distribution".
data.name	a character string giving the name(s) of the data.

### Note

The Zk test is the recommended EDF test by Jin Zhang.

### Author(s)

Ning Cui

### References

Jin Zhang: Goodness-of-Fit Tests Based on the Likelihood Ratio. *Journal of the Royal Statistical Society*, 64, 281-294.

Jin Zhang, Yuehua Wu: Likelihood-ratio tests for normality. *Computational Statistics & Data Analysis*, 49, 709-721.

Jin Zhang: Powerful Two-Sample Tests Based on the Likelihood Ratio. *Technometrics*, 48:1, 95-103.

### See Also

[ks.test](#) for performing a one- or two-sample Kolmogorov-Smirnov test. [zc.test](#), [za.test](#) for performing a powerful goodness-of-fit test based on the likelihood ratio.

### Examples

```
x<-rbeta(50,shape1 = 0.6,shape2 = 0.8)
y<-rnorm(50)
zk.test(x,y)
zk.test(x,"unif")
zk.test(x,"norm")
zk.test(x,"unif",para = list(min=1,max=2))
```

```
zk.test(x,"exp",para = list(rate=1))
zk.test(x,"norm",para = list(mean=1,sd=2))
zk.test(x,"lognorm",para = list(mean=1,sd=2))
zk.test(x,"weibull",para = list(shape=1,scale=2))
zk.test(x,"gamma",para = list(shape=2,scale=1))
zk.test(x,"t",para = list(df=3))
```

# Index

\*Topic **h**test

za. test, 2

zc. test, 3

zk. test, 5

ks. test, 3, 5, 6

za. test, 2, 5, 6

zc. test, 3, 3, 6

zk. test, 3, 5, 5