

Package ‘estats’

November 4, 2025

Type Package

Title Fast and Light-Weight Energy Statistics

Version 1.0

Date 2025-10-27

Author Michail Tsagris [aut, cre],
Manos Papadakis [aut]

Maintainer Michail Tsagris <mtsagris@uoc.gr>

Depends R (>= 4.0)

Imports dcov, pdcor, Rfast, Rfast2

Description Fast and memory-less computation of the energy statistics related quantities for vectors and matrices. References include: Szekely G. J. and Rizzo M. L. (2014), <[doi:10.1214/14-AOS1255](https://doi.org/10.1214/14-AOS1255)>. Szekely G. J. and Rizzo M. L. (2023), <ISBN:9781482242744>. Tsagris M. and Papadakis M. (2025). <[doi:10.48550/arXiv.2501.02849](https://doi.org/10.48550/arXiv.2501.02849)>.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2025-11-04 19:10:02 UTC

Contents

estats-package	2
Approximate distance variance	2
Distance correlation matrix	3
Distance variance, covariance and correlation	4
Energy based normality test	6
Energy distance between matrices	7
Energy test of equal univariate distributions	8
Hypothesis test for the distance correlation with high dimensional matrices	9
Hypothesis testing for many partial distance correlations	10
Hypothesis testing for the partial distance correlation	12
Many pPartial distance correlations	13
Partial distance correlation	14
Permutation based and asymptotic (approximate) distance covariance hypothesis test . .	15

Index**18**

<code>estats-package</code>	<i>Fast and Light-Weight Energy Statistics</i>
-----------------------------	--

Description

Description: Fast and memory-less computation of the energy statistics related quantities for vectors and matrices.

Details

Package:	<code>estats</code>
Type:	Package
Version:	1.0
Date:	2025-10-27
License:	GPL-2

Maintainers

Michail Tsagris <mtsagris@uoc.gr>.

Author(s)

Michail Tsagris <mtsagris@uoc.gr> and Manos Papadakis <papadakm95@gmail.com>.

Approximate distance variance

Approximate distance variance

Description

Approximate distance variance.

Usage

```
adcov(x, y, bc = FALSE, K = 100)
```

Arguments

- x A numerical matrix.
- y A numerical matrix.
- bc If you want the bias-corrected distance correlation set this equal to TRUE.
- K The number of projections to perform.

Details

The approximate distance covariance of Huand and Huo (2022) is computed.

Value

The approximate distance covariance.

Author(s)

Michail Tsagris and Manos Papadakis.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Szekely G.J., Rizzo M.L. and Bakirov N.K.(2007). Measuring and Testing Independence by Correlation of Distances. *Annals of Statistics*, 35(6):2769–2794.

Szekely G. J. and Rizzo M. L. (2023). The Energy of Data and Distance Correlation. Chapman and Hall/CRC.

Huang C. and Huo X. (2022). A statistically and numerically efficient independence test based on random projections and distance covariance. *Frontiers in Applied Mathematics and Statistics*, 7: 779841.

Tsagris M. and Papadakis M. (2025). Fast and light-weight energy statistics using the R package Rfast. <https://arxiv.org/abs/2501.02849>

See Also

[adcov](#), [adcov.test](#)

Examples

```
x <- as.matrix(iris[1:50, 1:4])
y <- as.matrix(iris[51:100, 1:4])
res <- dvar(x[, 1])
dcor(x, y)
```

Distance correlation matrix

Distance correlation matrix

Description

Distance correlation matrix.

Usage

```
dcorr(x, bc = FALSE)
```

Arguments

- x A numerical matrix.
- bc If you want the bias-corrected distance correlation set this equal to TRUE.

Details

The squared distance correlation matrix is computed.

Value

A matrix with the pairwise squared distance correlations between all variables in x.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

G.J. Szekely, M.L. Rizzo and N. K. Bakirov (2007). Measuring and Testing Independence by Correlation of Distances. Annals of Statistics, 35(6):2769-2794.

See Also

[dcor](#)

Examples

```
x <- as.matrix(iris[1:50, 1:4])
res <- dcorm(x)
```

Distance variance, covariance and correlation

Distance variance, covariance and correlation

Description

Distance variance, covariance and correlation.

Usage

```
dvar(x, bc = FALSE)
dcov(x, y, bc = FALSE)
dcor(x, y, bc = FALSE)
```

Arguments

- x A numerical matrix or a vector.
- y A numerical matrix or a vector.
- bc If you want the bias-corrected distance correlation set this equal to TRUE.

Details

The distance variance of a matrix/vector, the distance covariance or distance correlation of two matrices is calculated. For the dcov() and dcor(), if x and y are matrices, they must have the same dimensions. We have optimized the code, using the formulas provided in Szekely and Rizzo (2023), but only for the case that both matrices are of the same dimensionality.

Value

The distance covariance or the distance variance.

For the distance correlation a vector with the distance covariance, the distance variance of x, the distance variance of Y and the distance correlation.

Author(s)

Michail Tsagris and Manos Papadakis.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Manos Papadakis <papadakm95@gmail.com>.

References

- Szekely G.J., Rizzo M.L. and Bakirov N.K.(2007). Measuring and Testing Independence by Correlation of Distances. *Annals of Statistics*, 35(6):2769–2794.
- Szekely G. J. and Rizzo M. L. (2023). The Energy of Data and Distance Correlation. Chapman and Hall/CRC.
- Tsagris M. and Papadakis M. (2025). Fast and light-weight energy statistics using the R package Rfast. <https://arxiv.org/abs/2501.02849>

See Also

[pdcor](#) , [dcorm](#)

Examples

```
x <- as.matrix(iris[1:50, 1:4])
y <- as.matrix(iris[51:100, 1:4])
res <- dvar(x[, 1])
dcor(x, y)
```

Energy based normality test
Energy based normality test

Description

Energy based normality test.

Usage

```
normal.etest(x, R = 999)
```

Arguments

- | | |
|---|--|
| x | A numerical vector. |
| R | The number of Monte Carlo samples to generate. |

Details

The energy based normality test is performed where the p-value is computed via parametric bootstrap. The function is faster than the original implementation in the R package "energy".

Value

A vector with two values, the test statistic value and the Monte Carlo (parametric bootstrap) based p-value.

Author(s)

Michail Tsagris
R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Szekely G. J. and Rizzo M.L. (2005) A New Test for Multivariate Normality. Journal of Multivariate Analysis, 93(1): 58–80.

See Also

[eqdist.etest](#)

Examples

```
x <- rnorm(100)
normal.etest(x, R = 299)
```

Energy distance between matrices
Energy distance between matrices

Description

Energy distance between matrices.

Usage

```
edist(x, y = NULL)
```

Arguments

- x A matrix with numbers or a list with matrices.
- y A second matrix with data. The number of columns of x and y must match. The number of rows can be different.

Details

This calculates the energy distance between two matrices. It will work even for tens of thousands of rows, it will just take some time. See the references for more information. If you have many matrices and want to calculate the distance matrix, then put them in a list and use the function.

Value

If "x" is matrix, a numerical value, the energy distance. If "x" is list, a matrix with all pairwise distances of the matrices.

Author(s)

Manos Papadakis

R implementation and documentation: Manos Papadakis <papadakm95@gmail.com>.

References

- Szekely G. J. and Rizzo M. L. (2004) Testing for Equal Distributions in High Dimension, InterStat, November (5).
- Szekely G. J. (2000) Technical Report 03-05, E-statistics: Energy of Statistical Samples, Department of Mathematics and Statistics, Bowling Green State University.
- Sejdinovic D., Sriperumbudur B., Gretton A. and Fukumizu, K. (2013). Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *The Annals of Statistics*, 41(5): 2263–2291.
- Szekely G. J. and Rizzo M. L. (2023). *The Energy of Data and Distance Correlation*. Chapman and Hall/CRC.
- Tsagris M. and Papadakis M. (2025). Fast and light-weight energy statistics using the R package Rfast. <https://arxiv.org/abs/2501.02849>

See Also[dcov](#)**Examples**

```
x <- as.matrix(iris[1:50, 1:4])
y <- as.matrix(iris[51:100, 1:4])
res<-edist(x, y)
z <- as.matrix(iris[101:150, 1:4])
a <- list()
a[[ 1 ]] <- x
a[[ 2 ]] <- y
a[[ 3 ]] <- z
res<-edist(a)

x<-y<-z<-a<-NULL
```

Energy test of equal univariate distributions*Energy test of equal univariate distributions***Description**

Energy test of equal univariate distributions.

Usage

```
eqdist.etest(y, x, R = 999)
```

Arguments

- y A numerical vector or a numerical matrix.
- x A numerical vector or a numerical matrix.
- R The number of permutations to perform.

Details

The test performs the energy test of equal univariate distributions and the p-value is computed via permutations. Both the univariate and multivariate cases are memory-saving, the univariate case is pretty fast, but the multivariate case is not fast enough.

Value

The permutation based p-value.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Szekely G. J. and Rizzo M. L. (2004) Testing for Equal Distributions in High Dimension, InterStat, November (5).

Szekely G. J. (2000) Technical Report 03-05, E-statistics: Energy of Statistical Samples, Department of Mathematics and Statistics, Bowling Green State University.

Sejdinovic D., Sriperumbudur B., Gretton A. and Fukumizu, K. (2013). Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *The Annals of Statistics*, 41(5): 2263–2291.

Szekely G. J. and Rizzo M. L. (2023). The Energy of Data and Distance Correlation. Chapman and Hall/CRC.

Tsagris M. and Papadakis M. (2025). Fast and light-weight energy statistics using the R package Rfast. https://www.researchgate.net/publication/387583091_Fast_and_light-weight_energy_statistics_using_the_R_package

See Also

[normal.etest](#), [dcorm](#)

Examples

```
y <- rnorm(30)
x <- rnorm(40)
eqdist.etest(y, x, R = 99)
```

Hypothesis test for the distance correlation with high dimensional
matrices

*Hypothesis test for the distance correlation with high dimensional ma-
trices*

Description

Hypothesis test for the distance correlation with high dimensional matrices.

Usage

```
dcor.ttest(x, y, logged = FALSE)
```

Arguments

x	A numerical matrix.
y	A numerical matrix (of the same dimensions).
logged	Do you want the logarithm of the p-value to be returned? If yes, set this to TRUE.

Details

The bias corrected distance correlation is used. The hypothesis test is whether the two matrices are independent or not. Note, that this test is size correct as both the sample size and the dimensionality goes to infinity. It will not have the correct type I error for univariate data or for matrices with just a couple of variables.

Value

A vector with 4 elements, the bias corrected distance correlation, the degrees of freedom, the test statistic and its associated p-value.

Author(s)

Manos Papadakis

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Manos Papadakis <papadakm95@gmail.com>.

References

G.J. Szekely, M.L. Rizzo and N. K. Bakirov (2007). Measuring and Testing Independence by Correlation of Distances. *Annals of Statistics*, 35(6): 2769–2794.

Szekely G. J. and Rizzo M. L. (2023). The Energy of Data and Distance Correlation. Chapman and Hall/CRC.

See Also

[dcov](#), [edist](#)

Examples

```
x <- as.matrix(iris[1:50, 1:4])
y <- as.matrix(iris[51:100, 1:4])
dcor.ttest(x, y)
```

Hypothesis testing for many partial distance correlations

Hypothesis testing for many partial distance correlations

Description

Hypothesis testing for many partial distance correlations.

Usage

```
mpdcor.test(y, x, z, R = 500)
```

Arguments

y	A numerical vector.
x	A numerical matrix.
z	A numerical vector.
R	The number of permutations to implement. If R = 1, the the asymptotic p-value is returned only.

Details

Hypothesis testing between y and each column of x, conditional on z is performed.

Value

A matrix with three columns: the unbiased partial distance correlation, the permutation based p-value and the asymptotic p-value as proposed by Shen, Panda and Vogelstein (2022).

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

- Szekely G. J. and Rizzo M. L. (2014). Partial Distance Correlation with Methods for Dissimilarities. *The Annals of Statistics*, 42(6): 2382–2412.
- Shen C., Panda S. and Vogelstein J. T. (2022). The Chi-Square Test of Distance Correlation. *Journal of Computational and Graphical Statistics*, 31(1): 254–262.
- Szekely G. J. and Rizzo M. L. (2023). The Energy of Data and Distance Correlation. Chapman and Hall/CRC.
- Tsagris M. and Papadakis M. (2025). Fast and light-weight energy statistics using the R package Rfast. <https://arxiv.org/abs/2501.02849>
- Kontemeniotis N., Vargiakakis R. and Tsagris M. (2025). On independence testing using the (partial) distance correlation. <https://arxiv.org/abs/2506.15659v1>

See Also

[mpdcor](#) , [pdcor.test](#)

Examples

```
y <- iris[, 1]
x <- matrix( rnorm(150 * 10), ncol = 10 )
z <- iris[, 2]
mpdcor.test(y, x, z)
```

Hypothesis testing for the partial distance correlation
Hypothesis testing for the partial distance correlation

Description

Hypothesis testing for the partial distance correlation.

Usage

```
pdcor.test(x, y, z, type = 1, R = 500)
```

Arguments

x	A numerical vector or matrix.
y	A numerical vector or matrix.
z	A numerical vector or matrix.
type	In case that all x, y, and z are vectors the user may select the type = 2 which is even faster, but at the expense of requiring more memory.
R	The number of permutations to implement. If R = 1, the the asymptotic p-value is returned only.

Details

Hypothesis testing using the unbiased partial distance correlation between x and y conditioning on z is computed. **Note:** currently, only two cases are supported, all x, y, and z are vectors or they are all matrices with the same dimensions.

Value

A vector with the unbiased partial distance correlation, the permutation based p-value and the asymptotic p-value as proposed by Shen, Panda and Vogelstein (2022).

Author(s)

Michail Tsagris and Nikolaos Kontemeniotis .

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Nikolaos Kontemeniotis <kontemeniotisn@gmail.com>.

References

- Szekely G. J. and Rizzo M. L. (2014). Partial Distance Correlation with Methods for Dissimilarities. *The Annals of Statistics*, 42(6): 2382–2412.
- Shen C., Panda S. and Vogelstein J. T. (2022). The Chi-Square Test of Distance Correlation. *Journal of Computational and Graphical Statistics*, 31(1): 254–262.

Szekely G. J. and Rizzo M. L. (2023). The Energy of Data and Distance Correlation. Chapman and Hall/CRC.

Tsagris M. and Papadakis M. (2025). Fast and light-weight energy statistics using the R package Rfast. <https://arxiv.org/abs/2501.02849>

Kontemeniotis N., Vargiakakis R. and Tsagris M. (2025). On independence testing using the (partial) distance correlation. <https://arxiv.org/abs/2506.15659v1>

See Also

[pdcor](#)

Examples

```
x <- iris[, 1]
y <- iris[, 2]
z <- iris[, 3]
pdcor.test(x, y, z)
```

Many pPartial distance correlations
Many partial distance correlations

Description

Many partial distance correlations.

Usage

`mpdcor(y, x, z)`

Arguments

<code>y</code>	A numerical vector.
<code>x</code>	A numerical matrix.
<code>z</code>	A numerical vector.

Details

This computes the unbiased pdcor between `y` and each column of `x`, conditional on the vector `z`.

Value

A vector with many unbiased partial distance correlations.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

- Szekely G. J. and Rizzo M. L. (2014). Partial Distance Correlation with Methods for Dissimilarities. *The Annals of Statistics*, 42(6): 2382–2412.
- Szekely G. J. and Rizzo M. L. (2023). The Energy of Data and Distance Correlation. Chapman and Hall/CRC.
- Tsagris M. and Papadakis M. (2025). Fast and light-weight energy statistics using the R package Rfast. <https://arxiv.org/abs/2501.02849>
- Kontemeniotis N., Vargiakakis R. and Tsagris M. (2025). On independence testing using the (partial) distance correlation. <https://arxiv.org/abs/2506.15659v1>

See Also

[pdcor](#) , [mpdcor.test](#)

Examples

```
y <- iris[, 1]
x <- matrix( rnorm(150 * 10), ncol = 10 )
z <- iris[, 2]
mpdcor(y, x, z)
pdcor(y, x[, 1], z)
```

Partial distance correlation

Partial distance correlation

Description

Partial distance correlation.

Usage

```
pdcor(x, y, z)
```

Arguments

- x A numerical vector or matrix.
- y A numerical vector or matrix.
- z A numerical vector or matrix.

Details

The unbiased partial distance correlation between x and y conditioning on z is computed. **Note:** currently, only two cases are supported, all x, y, and z are vectors or they are all matrices with the same dimensions.

Value

The unbiased partial distance correlation.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

- Szekely G. J. and Rizzo M. L. (2014). Partial Distance Correlation with Methods for Dissimilarities. *The Annals of Statistics*, 42(6): 2382–2412.
- Szekely G. J. and Rizzo M. L. (2023). The Energy of Data and Distance Correlation. Chapman and Hall/CRC.
- Tsagris M. and Papadakis M. (2025). Fast and light-weight energy statistics using the R package Rfast. <https://arxiv.org/abs/2501.02849>
- Kontemeniotis N., Vargiakakis R. and Tsagris M. (2025). On independence testing using the (partial) distance correlation. <https://arxiv.org/abs/2506.15659v1>

See Also

[pdcor.test](#), [mpdcor](#)

Examples

```
x <- iris[, 1]
y <- iris[, 2]
z <- iris[, 3]
pdcor(x, y, z)
```

Permutation based and asymptotic (approximate) distance covariance hypothesis test

Permutation based and asymptotic distance (approximate) covariance hypothesis test

Description

Permutation based and asymptotic (approximate) distance covariance hypothesis test.

Usage

```
dcov.test(x, y, R = 1)
adcov.test(x, y, R = 499)
```

Arguments

- x A numerical matrix or a vector. For the approximate distance covariance test (adcov.test()) this can only be a matrix.
- y A numerical matrix (of the same dimensions) or a vector. For the approximate distance covariance test (adcov.test()) this can only be a matrix (the number of variables need not be the same).
- R For the dcov.test() if R=1, the asymptotic p-value of Shen, Panda and Vogelstein (2022) is returned. If R > 1, the permutation based p-value is computed. For the adcov.test() this must be a large number because the permutation based p-value is returned.

Details

The bias corrected distance correlation is used. The hypothesis test is whether the two matrices are independent or not. If R=1, the test is based on the distance correlation. If R > 1 the test is based upon the distance covariance. For the approximate distance covariance test of Huang and Huo (2022) that is based upon permutations is performed.

Value

A vector with 2 elements, the bias corrected distance correlation or covariance, and the associated permutation or asymptotic based p-value.

Author(s)

Manos Papadakis

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

- Shen C., Panda S. and Vogelstein J. T. (2022). The Chi-Square Test of Distance Correlation. *Journal of Computational and Graphical Statistics*, 31(1): 254–262.
- G.J. Szekely, M.L. Rizzo and N. K. Bakirov (2007). Measuring and Testing Independence by Correlation of Distances. *Annals of Statistics*, 35(6): 2769–2794.
- Szekely G. J. and Rizzo M. L. (2023). The Energy of Data and Distance Correlation. Chapman and Hall/CRC.
- Huang C. and Huo X. (2022). A statistically and numerically efficient independence test based on random projections and distance covariance. *Frontiers in Applied Mathematics and Statistics*, 7: 779841.

See Also

[dcov](#), [adcov](#)

Examples

```
x <- as.matrix(iris[1:50, 1:4])
y <- as.matrix(iris[51:100, 1:4])
res <- dcov.test(x, y)
```

Index

- * **Energy distances**
 - Energy test of equal univariate distributions, 8
- * **total sum**
 - Energy test of equal univariate distributions, 8
- adcov**, 3, 16
 - adcov (Approximate distance variance), 2
 - adcov.test, 3
 - adcov.test (Permutation based and asymptotic (approximate) distance covariance hypothesis test), 15
 - Approximate distance variance, 2
- dcor**, 4
 - dcor (Distance variance, covariance and correlation), 4
 - dcor.ttest (Hypothesis test for the distance correlation with high dimensional matrices), 9
 - dcorm, 5, 9
 - dcorm (Distance correlation matrix), 3
 - dcov, 8, 10, 16
 - dcov (Distance variance, covariance and correlation), 4
 - dcov.test (Permutation based and asymptotic (approximate) distance covariance hypothesis test), 15
 - Distance correlation matrix, 3
 - Distance variance, covariance and correlation, 4
 - dvar (Distance variance, covariance and correlation), 4
- edist**, 10
 - edist (Energy distance between matrices), 7
- Energy based normality test, 6
 - Energy distance between matrices, 7
 - Energy test of equal univariate distributions, 8
- eqdist.estest**, 6
 - eqdist.estest (Energy test of equal univariate distributions), 8
- estats-package**, 2
 - Hypothesis test for the distance correlation with high dimensional matrices, 9
 - Hypothesis testing for many partial distance correlations, 10
 - Hypothesis testing for the partial distance correlation, 12
- Many pPartial distance correlations, 13
 - mpdcor, 11, 15
 - mpdcor (Many pPartial distance correlations), 13
 - mpdcor.test, 14
 - mpdcor.test (Hypothesis testing for many partial distance correlations), 10
- normal.estest**, 9
 - normal.estest (Energy based normality test), 6
- Partial distance correlation, 14
 - pdcor, 5, 13, 14
 - pdcor (Partial distance correlation), 14
 - pdcor.test, 11, 15
 - pdcor.test (Hypothesis testing for the partial distance correlation), 12
- Permutation based and asymptotic (approximate) distance covariance hypothesis test, 15