

# Package ‘PEtests’

May 7, 2026

**Type** Package

**Title** Power-Enhanced (PE) Tests for High-Dimensional Data

**Version** 0.1.0

## Description

Two-sample power-enhanced mean tests, covariance tests, and simultaneous tests on mean vectors and covariance matrices for high-dimensional data. Methods of these PE tests are presented in Yu, Li, and Xue (2022) <[doi:10.1080/01621459.2022.2126781](https://doi.org/10.1080/01621459.2022.2126781)>; Yu, Li, Xue, and Li (2022) <[doi:10.1080/01621459.2022.2061354](https://doi.org/10.1080/01621459.2022.2061354)>.

**Author** Xiufan Yu [aut, cre],  
Danning Li [aut],  
Lingzhou Xue [aut],  
Runze Li [aut]

**Maintainer** Xiufan Yu <[xiufan.yu@nd.edu](mailto:xiufan.yu@nd.edu)>

**Imports** stats

**Encoding** UTF-8

**RoxygenNote** 7.2.3.9000

**License** GPL (>= 3)

**Suggests** testthat (>= 3.0.0)

**Config/testthat/edition** 3

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2023-05-22 08:40:02 UTC

## Contents

PEtests-package . . . . .	2
covtest . . . . .	3
covtest.clx . . . . .	4
covtest.lc . . . . .	5
covtest.pe.cauchy . . . . .	6
covtest.pe.comp . . . . .	7

covtest.pe.fisher . . . . .	9
meantest . . . . .	10
meantest.clx . . . . .	11
meantest.cq . . . . .	12
meantest.pe.cauchy . . . . .	13
meantest.pe.comp . . . . .	14
meantest.pe.fisher . . . . .	15
simultest . . . . .	16
simultest.cauchy . . . . .	18
simultest.chisq . . . . .	19
simultest.fisher . . . . .	20
simultest.pe.cauchy . . . . .	21
simultest.pe.chisq . . . . .	22
simultest.pe.fisher . . . . .	23

<b>Index</b>	<b>25</b>
--------------	-----------

---

PEtests-package	<i>Power-Enhanced (PE) Tests for High-Dimensional Data</i>
-----------------	--

---

## Description

The package implements several two-sample power-enhanced mean tests, covariance tests, and simultaneous tests on mean vectors and covariance matrices for high-dimensional data.

## Details

There are three main functions:

`covtest`  
`meantest`  
`simultest`

## References

- Chen, S. X. and Qin, Y. L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Annals of Statistics*, 38(2):808–835. doi:[10.1214/09AOS716](https://doi.org/10.1214/09AOS716)
- Cai, T. T., Liu, W., and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *Journal of the American Statistical Association*, 108(501):265–277. doi:[10.1080/01621459.2012.758041](https://doi.org/10.1080/01621459.2012.758041)
- Cai, T. T., Liu, W., and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 76(2):349–372. doi:[10.1111/rssb.12034](https://doi.org/10.1111/rssb.12034)
- Li, J. and Chen, S. X. (2012). Two sample tests for high-dimensional covariance matrices. *The Annals of Statistics*, 40(2):908–940. doi:[10.1214/12AOS993](https://doi.org/10.1214/12AOS993)
- Yu, X., Li, D., and Xue, L. (2022). Fisher’s combined probability test for high-dimensional covariance matrices. *Journal of the American Statistical Association*, (in press):1–14. doi:[10.1080/01621459.2022.2126781](https://doi.org/10.1080/01621459.2022.2126781)

Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14. doi:10.1080/01621459.2022.2061354

## Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
covtest(X, Y)
meantest(X, Y)
simultest(X, Y)
```

---

covtest

*Two-sample covariance tests for high-dimensional data*

---

## Description

This function implements five two-sample covariance tests on high-dimensional covariance matrices. Let  $\mathbf{X} \in \mathbb{R}^p$  and  $\mathbf{Y} \in \mathbb{R}^p$  be two  $p$ -dimensional populations with mean vectors  $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$  and covariance matrices  $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$ , respectively. The problem of interest is to test the equality of the two covariance matrices:

$$H_{0c} : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2.$$

Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . We denote  $\text{dataX} = (\mathbf{X}_1, \dots, \mathbf{X}_{n_1})^\top \in \mathbb{R}^{n_1 \times p}$  and  $\text{dataY} = (\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2})^\top \in \mathbb{R}^{n_2 \times p}$ .

## Usage

```
covtest(dataX, dataY, method='pe.comp', delta=NULL)
```

## Arguments

dataX	an $n_1$ by $p$ data matrix
dataY	an $n_2$ by $p$ data matrix
method	the method type (default = 'pe.comp'); chosen from <ul style="list-style-type: none"> <li>• 'clx': the <math>l_\infty</math>-norm-based covariance test, proposed in Cai et al. (2013); see <a href="#">covtest.clx</a> for details.</li> <li>• 'lc': the <math>l_2</math>-norm-based covariance test, proposed in Li and Chen (2012); see <a href="#">covtest.lc</a> for details.</li> <li>• 'pe.cauchy': the PE covariance test via Cauchy combination; see <a href="#">covtest.pe.cauchy</a> for details.</li> <li>• 'pe.comp': the PE covariance test via the construction of PE components; see <a href="#">covtest.pe.comp</a> for details.</li> </ul>

- 'pe.fisher': the PE covariance test via Fisher's combination; see [covtest.pe.fisher](#) for details.
- delta This is needed only in method='pe.comp'; see [covtest.pe.comp](#) for details. The default is NULL.

### Value

method the method type  
 stat the value of test statistic  
 pval the p-value for the test.

### References

- Cai, T. T., Liu, W., and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *Journal of the American Statistical Association*, 108(501):265–277.
- Li, J. and Chen, S. X. (2012). Two sample tests for high-dimensional covariance matrices. *The Annals of Statistics*, 40(2):908–940.
- Yu, X., Li, D., and Xue, L. (2022). Fisher's combined probability test for high-dimensional covariance matrices. *Journal of the American Statistical Association*, (in press):1–14.
- Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

### Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
covtest(X,Y)
```

---

covtest.clx	<i>Two-sample high-dimensional covariance test (Cai, Liu and Xia, 2013)</i>
-------------	---

---

### Description

This function implements the two-sample  $l_\infty$ -norm-based high-dimensional covariance test proposed in Cai, Liu and Xia (2013). Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . The test statistic is defined as

$$T_{CLX} = \max_{1 \leq i, j \leq p} \frac{(\hat{\sigma}_{ij1} - \hat{\sigma}_{ij2})^2}{\hat{\theta}_{ij1}/n_1 + \hat{\theta}_{ij2}/n_2},$$

where  $\hat{\sigma}_{ij1}$  and  $\hat{\sigma}_{ij2}$  are the sample covariances, and  $\hat{\theta}_{ij1}/n_1 + \hat{\theta}_{ij2}/n_2$  estimates the variance of  $\hat{\sigma}_{ij1} - \hat{\sigma}_{ij2}$ . The explicit formulas of  $\hat{\sigma}_{ij1}$ ,  $\hat{\sigma}_{ij2}$ ,  $\hat{\theta}_{ij1}$  and  $\hat{\theta}_{ij2}$  can be found in Section 2 of Cai, Liu

and Xia (2013). With some regularity conditions, under the null hypothesis  $H_{0c} : \Sigma_1 = \Sigma_2$ , the test statistic  $T_{CLX} - 4 \log p + \log \log p$  converges in distribution to a Gumbel distribution  $G_{cov}(x) = \exp(-\frac{1}{\sqrt{8\pi}} \exp(-\frac{x}{2}))$  as  $n_1, n_2, p \rightarrow \infty$ . The asymptotic  $p$ -value is obtained by

$$p_{CLX} = 1 - G_{cov}(T_{CLX} - 4 \log p + \log \log p).$$

### Usage

```
covtest.clx(dataX,dataY)
```

### Arguments

dataX            an  $n_1$  by  $p$  data matrix  
 dataY            an  $n_2$  by  $p$  data matrix

### Value

stat the value of test statistic  
 pval the  $p$ -value for the test.

### References

Cai, T. T., Liu, W., and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *Journal of the American Statistical Association*, 108(501):265–277.

### Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
covtest.clx(X,Y)
```

---

covtest.lc

*Two-sample high-dimensional covariance test (Li and Chen, 2012)*

---

### Description

This function implements the two-sample  $l_2$ -norm-based high-dimensional covariance test proposed by Li and Chen (2012). Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . The test statistic  $T_{LC}$  is defined as

$$T_{LC} = A_{n_1} + B_{n_2} - 2C_{n_1, n_2},$$

where  $A_{n_1}$ ,  $B_{n_2}$ , and  $C_{n_1, n_2}$  are unbiased estimators for  $\text{tr}(\Sigma_1^2)$ ,  $\text{tr}(\Sigma_2^2)$ , and  $\text{tr}(\Sigma_1 \Sigma_2)$ , respectively. Under the null hypothesis  $H_{0c} : \Sigma_1 = \Sigma_2$ , the leading variance of  $T_{LC}$  is  $\sigma_{T_{LC}}^2 = 4(\frac{1}{n_1} + \frac{1}{n_2})^2 \text{tr}^2(\Sigma^2)$ , which can be consistently estimated by  $\hat{\sigma}_{T_{LC}}^2$ . The explicit formulas of  $A_{n_1}$ ,

$B_{n_2}$ ,  $C_{n_1, n_2}$  and  $\hat{\sigma}_{T_{LC}}^2$  can be found in Equations (2.1), (2.2) and Theorem 1 of Li and Chen (2012). With some regularity conditions, under the null hypothesis  $H_{0c} : \Sigma_1 = \Sigma_2$ , the test statistic  $T_{LC}$  converges in distribution to a standard normal distribution as  $n_1, n_2, p \rightarrow \infty$ . The asymptotic  $p$ -value is obtained by

$$p_{LC} = 1 - \Phi(T_{LC}/\hat{\sigma}_{T_{LC}}),$$

where  $\Phi(\cdot)$  is the cdf of the standard normal distribution.

### Usage

```
covtest.lc(dataX, dataY)
```

### Arguments

dataX            an  $n_1$  by  $p$  data matrix  
 dataY            an  $n_2$  by  $p$  data matrix

### Value

stat the value of test statistic  
 pval the  $p$ -value for the test.

### References

Li, J. and Chen, S. X. (2012). Two sample tests for high-dimensional covariance matrices. *The Annals of Statistics*, 40(2):908–940.

### Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
covtest.lc(X, Y)
```

---

covtest.pe.cauchy	<i>Two-sample PE covariance test for high-dimensional data via Cauchy combination</i>
-------------------	---

---

### Description

This function implements the two-sample PE covariance test via Cauchy combination. Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . Let  $p_{LC}$  and  $p_{CLX}$  denote the  $p$ -values associated with the  $l_2$ -norm-based covariance test (see `covtest.lc` for details) and the  $l_\infty$ -norm-based covariance test (see `covtest.clx` for details), respectively. The PE covariance test via Cauchy combination is defined as

$$T_{Cauchy} = \frac{1}{2} \tan((0.5 - p_{LC})\pi) + \frac{1}{2} \tan((0.5 - p_{CLX})\pi).$$

It has been proved that with some regularity conditions, under the null hypothesis  $H_{0c} : \Sigma_1 = \Sigma_2$ , the two tests are asymptotically independent as  $n_1, n_2, p \rightarrow \infty$ , and therefore  $T_{Cauchy}$  asymptotically converges in distribution to a standard Cauchy distribution. The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - F_{Cauchy}(T_{Cauchy}),$$

where  $F_{Cauchy}(\cdot)$  is the cdf of the standard Cauchy distribution.

### Usage

```
covtest.pe.cauchy(dataX, dataY)
```

### Arguments

dataX            an  $n_1$  by  $p$  data matrix  
dataY            an  $n_2$  by  $p$  data matrix

### Value

stat the value of test statistic  
pval the  $p$ -value for the test.

### References

Yu, X., Li, D., and Xue, L. (2022). Fisher's combined probability test for high-dimensional covariance matrices. *Journal of the American Statistical Association*, (in press):1–14.

### Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
covtest.pe.cauchy(X, Y)
```

---

covtest.pe.comp	<i>Two-sample PE covariance test for high-dimensional data via PE component</i>
-----------------	---

---

### Description

This function implements the two-sample PE covariance test via the construction of the PE component. Let  $T_{LC}/\hat{\sigma}_{T_{LC}}$  denote the  $l_2$ -norm-based covariance test statistic (see [covtest.lc](#) for details). The PE component is constructed by

$$J_c = \sqrt{p} \sum_{i=1}^p \sum_{j=1}^p T_{ij} \hat{\xi}_{ij}^{-1/2} \mathcal{I}\{\sqrt{2} T_{ij} \hat{\xi}_{ij}^{-1/2} + 1 > \delta_{cov}\},$$

where  $\delta_{cov}$  is a threshold for the screening procedure, recommended to take the value of  $\delta_{cov} = 4 \log(\log(n_1 + n_2)) \log p$ . The explicit forms of  $T_{ij}$  and  $\hat{\xi}_{ij}$  can be found in Section 3.2 of Yu et al. (2022). The PE covariance test statistic is defined as

$$T_{PE} = T_{LC} / \hat{\sigma}_{T_{LC}} + J_c.$$

With some regularity conditions, under the null hypothesis  $H_{0c} : \Sigma_1 = \Sigma_2$ , the test statistic  $T_{PE}$  converges in distribution to a standard normal distribution as  $n_1, n_2, p \rightarrow \infty$ . The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - \Phi(T_{PE}),$$

where  $\Phi(\cdot)$  is the cdf of the standard normal distribution.

### Usage

```
covtest.pe.comp(dataX, dataY, delta=NULL)
```

### Arguments

dataX	an $n_1$ by $p$ data matrix
dataY	an $n_2$ by $p$ data matrix
delta	a scalar; the thresholding value used in the construction of the PE component. If not specified, the function uses a default value $\delta_{cov} = 4 \log(\log(n_1 + n_2)) \log p$ .

### Value

stat the value of test statistic

pval the p-value for the test.

### References

Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

### Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
covtest.pe.comp(X, Y)
```

---

covtest.pe.fisher	<i>Two-sample PE covariance test for high-dimensional data via Fisher's combination</i>
-------------------	---

---

## Description

This function implements the two-sample PE covariance test via Fisher's combination. Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . Let  $p_{LC}$  and  $p_{CLX}$  denote the  $p$ -values associated with the  $l_2$ -norm-based covariance test (see [covtest.lc](#) for details) and the  $l_\infty$ -norm-based covariance test (see [covtest.clx](#) for details), respectively. The PE covariance test via Fisher's combination is defined as

$$T_{Fisher} = -2 \log(p_{LC}) - 2 \log(p_{CLX}).$$

It has been proved that with some regularity conditions, under the null hypothesis  $H_{0c} : \Sigma_1 = \Sigma_2$ , the two tests are asymptotically independent as  $n_1, n_2, p \rightarrow \infty$ , and therefore  $T_{Fisher}$  asymptotically converges in distribution to a  $\chi_4^2$  distribution. The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - F_{\chi_4^2}(T_{Fisher}),$$

where  $F_{\chi_4^2}(\cdot)$  is the cdf of the  $\chi_4^2$  distribution.

## Usage

```
covtest.pe.fisher(dataX, dataY)
```

## Arguments

dataX	an $n_1$ by $p$ data matrix
dataY	an $n_2$ by $p$ data matrix

## Value

stat the value of test statistic  
 pval the  $p$ -value for the test.

## References

Yu, X., Li, D., and Xue, L. (2022). Fisher's combined probability test for high-dimensional covariance matrices. *Journal of the American Statistical Association*, (in press):1–14.

## Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
covtest.pe.fisher(X, Y)
```

meantest

*Two-sample mean tests for high-dimensional data***Description**

This function implements five two-sample mean tests on high-dimensional mean vectors. Let  $\mathbf{X} \in \mathbb{R}^p$  and  $\mathbf{Y} \in \mathbb{R}^p$  be two  $p$ -dimensional populations with mean vectors  $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$  and covariance matrices  $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$ , respectively. The problem of interest is to test the equality of the two mean vectors of the two populations:

$$H_{0m} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2.$$

Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . We denote  $\text{dataX} = (\mathbf{X}_1, \dots, \mathbf{X}_{n_1})^\top \in \mathbb{R}^{n_1 \times p}$  and  $\text{dataY} = (\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2})^\top \in \mathbb{R}^{n_2 \times p}$ .

**Usage**

```
meantest(dataX, dataY, method='pe.comp', delta=NULL)
```

**Arguments**

dataX	an $n_1$ by $p$ data matrix
dataY	an $n_2$ by $p$ data matrix
method	the method type (default = 'pe.comp'); chosen from <ul style="list-style-type: none"> <li>'clx': the <math>l_\infty</math>-norm-based mean test, proposed in Cai et al. (2014); see <a href="#">meantest.clx</a> for details.</li> <li>'cq': the <math>l_2</math>-norm-based mean test, proposed in Chen and Qin (2010); see <a href="#">meantest.cq</a> for details.</li> <li>'pe.cauchy': the PE mean test via Cauchy combination; see <a href="#">meantest.pe.cauchy</a> for details.</li> <li>'pe.comp': the PE mean test via the construction of PE components; see <a href="#">meantest.pe.comp</a> for details.</li> <li>'pe.fisher': the PE mean test via Fisher's combination; see <a href="#">meantest.pe.fisher</a> for details.</li> </ul>
delta	This is needed only in method='pe.comp'; see <a href="#">meantest.pe.comp</a> for details. The default is NULL.

**Value**

method the method type  
stat the value of test statistic  
pval the p-value for the test.

## References

- Chen, S. X. and Qin, Y. L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Annals of Statistics*, 38(2):808–835.
- Cai, T. T., Liu, W., and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 76(2):349–372.
- Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

## Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
meantest(X,Y)
```

---

meantest.clx

*Two-sample high-dimensional mean test (Cai, Liu and Xia, 2014)*


---

## Description

This function implements the two-sample  $l_\infty$ -norm-based high-dimensional mean test proposed in Cai, Liu and Xia (2014). Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . The test statistic is defined as

$$M_{CLX} = \frac{n_1 n_2}{n_1 + n_2} \max_{1 \leq j \leq p} \frac{(\bar{X}_j - \bar{Y}_j)^2}{\frac{1}{n_1 + n_2} [\sum_{u=1}^{n_1} (X_{uj} - \bar{X}_j)^2 + \sum_{v=1}^{n_2} (Y_{vj} - \bar{Y}_j)^2]}$$

With some regularity conditions, under the null hypothesis  $H_{0c} : \Sigma_1 = \Sigma_2$ , the test statistic  $M_{CLX} - 2 \log p + \log \log p$  converges in distribution to a Gumbel distribution  $G_{mean}(x) = \exp(-\frac{1}{\sqrt{\pi}} \exp(-\frac{x}{2}))$  as  $n_1, n_2, p \rightarrow \infty$ . The asymptotic  $p$ -value is obtained by

$$p_{CLX} = 1 - G_{mean}(M_{CLX} - 2 \log p + \log \log p).$$

## Usage

```
meantest.clx(dataX, dataY)
```

## Arguments

dataX            an  $n_1$  by  $p$  data matrix  
dataY            an  $n_2$  by  $p$  data matrix

## Value

stat the value of test statistic  
pval the  $p$ -value for the test.

## References

Cai, T. T., Liu, W., and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 76(2):349–372.

## Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
meantest.clx(X,Y)
```

---

meantest.cq

*Two-sample high-dimensional mean test (Chen and Qin, 2010)*


---

## Description

This function implements the two-sample  $l_2$ -norm-based high-dimensional mean test proposed by Chen and Qin (2010). Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . The test statistic  $M_{CQ}$  is defined as

$$M_{CQ} = \frac{1}{n_1(n_1 - 1)} \sum_{u \neq v}^{n_1} \mathbf{X}'_u \mathbf{X}_v + \frac{1}{n_2(n_2 - 1)} \sum_{u \neq v}^{n_2} \mathbf{Y}'_u \mathbf{Y}_v - \frac{2}{n_1 n_2} \sum_u^{n_1} \sum_v^{n_2} \mathbf{X}'_u \mathbf{Y}_v.$$

Under the null hypothesis  $H_{0m} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ , the leading variance of  $M_{CQ}$  is  $\sigma_{M_{CQ}}^2 = \frac{2}{n_1(n_1-1)} \text{tr}(\widehat{\boldsymbol{\Sigma}}_1^2) + \frac{2}{n_2(n_2-1)} \text{tr}(\widehat{\boldsymbol{\Sigma}}_2^2) + \frac{4}{n_1 n_2} \text{tr}(\widehat{\boldsymbol{\Sigma}}_1 \widehat{\boldsymbol{\Sigma}}_2)$ , which can be consistently estimated by  $\widehat{\sigma}_{M_{CQ}}^2 = \frac{2}{n_1(n_1-1)} \text{tr}(\widehat{\boldsymbol{\Sigma}}_1^2) + \frac{2}{n_2(n_2-1)} \text{tr}(\widehat{\boldsymbol{\Sigma}}_2^2) + \frac{4}{n_1 n_2} \text{tr}(\widehat{\boldsymbol{\Sigma}}_1 \widehat{\boldsymbol{\Sigma}}_2)$ . The explicit formulas of  $\text{tr}(\widehat{\boldsymbol{\Sigma}}_1^2)$ ,  $\text{tr}(\widehat{\boldsymbol{\Sigma}}_2^2)$ , and  $\text{tr}(\widehat{\boldsymbol{\Sigma}}_1 \widehat{\boldsymbol{\Sigma}}_2)$  can be found in Section 3 of Chen and Qin (2010). With some regularity conditions, under the null hypothesis  $H_{0m} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ , the test statistic  $M_{CQ}$  converges in distribution to a standard normal distribution as  $n_1, n_2, p \rightarrow \infty$ . The asymptotic  $p$ -value is obtained by

$$p_{CQ} = 1 - \Phi(M_{CQ}/\widehat{\sigma}_{M_{CQ}}),$$

where  $\Phi(\cdot)$  is the cdf of the standard normal distribution.

## Usage

```
meantest.cq(dataX, dataY)
```

## Arguments

dataX            an  $n_1$  by  $p$  data matrix  
dataY            an  $n_2$  by  $p$  data matrix

## Value

stat the value of test statistic  
pval the  $p$ -value for the test.

## References

Chen, S. X. and Qin, Y. L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Annals of Statistics*, 38(2):808–835.

## Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
meantest.cq(X,Y)
```

---

meantest.pe.cauchy	<i>Two-sample PE mean test for high-dimensional data via Cauchy combination</i>
--------------------	---

---

## Description

This function implements the two-sample PE covariance test via Cauchy combination. Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . Let  $p_{CQ}$  and  $p_{CLX}$  denote the  $p$ -values associated with the  $l_2$ -norm-based covariance test (see `meantest.cq` for details) and the  $l_\infty$ -norm-based covariance test (see `meantest.clx` for details), respectively. The PE covariance test via Cauchy combination is defined as

$$M_{Cauchy} = \frac{1}{2} \tan((0.5 - p_{CQ})\pi) + \frac{1}{2} \tan((0.5 - p_{CLX})\pi).$$

It has been proved that with some regularity conditions, under the null hypothesis  $H_{0m} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ , the two tests are asymptotically independent as  $n_1, n_2, p \rightarrow \infty$ , and therefore  $M_{Cauchy}$  asymptotically converges in distribution to a standard Cauchy distribution. The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - F_{Cauchy}(M_{Cauchy}),$$

where  $F_{Cauchy}(\cdot)$  is the cdf of the standard Cauchy distribution.

## Usage

```
meantest.pe.cauchy(dataX, dataY)
```

## Arguments

dataX	an $n_1$ by $p$ data matrix
dataY	an $n_2$ by $p$ data matrix

## Value

stat the value of test statistic  
pval the  $p$ -value for the test.

## References

- Chen, S. X. and Qin, Y. L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Annals of Statistics*, 38(2):808–835.
- Cai, T. T., Liu, W., and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 76(2):349–372.

## Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
meantest.pe.cauchy(X,Y)
```

---

meantest.pe.comp	<i>Two-sample PE mean test for high-dimensional data via PE component</i>
------------------	---

---

## Description

This function implements the two-sample PE mean via the construction of the PE component. Let  $M_{CQ}/\hat{\sigma}_{M_{CQ}}$  denote the  $l_2$ -norm-based mean test statistic (see [meantest.cq](#) for details). The PE component is constructed by

$$J_m = \sqrt{p} \sum_{i=1}^p M_i \hat{\nu}_i^{-1/2} \mathcal{I}\{\sqrt{2} M_i \hat{\nu}_i^{-1/2} + 1 > \delta_{mean}\},$$

where  $\delta_{mean}$  is a threshold for the screening procedure, recommended to take the value of  $\delta_{mean} = 2 \log(\log(n_1 + n_2)) \log p$ . The explicit forms of  $M_i$  and  $\hat{\nu}_j$  can be found in Section 3.1 of Yu et al. (2022). The PE covariance test statistic is defined as

$$M_{PE} = M_{CQ}/\hat{\sigma}_{M_{CQ}} + J_m.$$

With some regularity conditions, under the null hypothesis  $H_{0m} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ , the test statistic  $M_{PE}$  converges in distribution to a standard normal distribution as  $n_1, n_2, p \rightarrow \infty$ . The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - \Phi(M_{PE}),$$

where  $\Phi(\cdot)$  is the cdf of the standard normal distribution.

## Usage

```
meantest.pe.comp(dataX, dataY, delta=NULL)
```

## Arguments

dataX	an $n_1$ by $p$ data matrix
dataY	an $n_2$ by $p$ data matrix
delta	a scalar; the thresholding value used in the construction of the PE component. If not specified, the function uses a default value $\delta_{mean} = 2 \log(\log(n_1 + n_2)) \log p$ .

**Value**

stat the value of test statistic  
 pval the p-value for the test.

**References**

Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

**Examples**

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
meantest.pe.comp(X,Y)
```

---

meantest.pe.fisher	<i>Two-sample PE mean test for high-dimensional data via Fisher's combination</i>
--------------------	---

---

**Description**

This function implements the two-sample PE covariance test via Fisher's combination. Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . Let  $p_{CQ}$  and  $p_{CLX}$  denote the  $p$ -values associated with the  $l_2$ -norm-based covariance test (see [meantest.cq](#) for details) and the  $l_\infty$ -norm-based covariance test (see [meantest.clx](#) for details), respectively. The PE covariance test via Fisher's combination is defined as

$$M_{Fisher} = -2 \log(p_{CQ}) - 2 \log(p_{CLX}).$$

It has been proved that with some regularity conditions, under the null hypothesis  $H_{0m} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ , the two tests are asymptotically independent as  $n_1, n_2, p \rightarrow \infty$ , and therefore  $M_{Fisher}$  asymptotically converges in distribution to a  $\chi_4^2$  distribution. The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - F_{\chi_4^2}(M_{Fisher}),$$

where  $F_{\chi_4^2}(\cdot)$  is the cdf of the  $\chi_4^2$  distribution.

**Usage**

```
meantest.pe.fisher(dataX, dataY)
```

**Arguments**

dataX            an  $n_1$  by  $p$  data matrix  
 dataY            an  $n_2$  by  $p$  data matrix

**Value**

stat the value of test statistic  
 pval the p-value for the test.

**References**

Chen, S. X. and Qin, Y. L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Annals of Statistics*, 38(2):808–835.

Cai, T. T., Liu, W., and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 76(2):349–372.

**Examples**

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
meantest.pe.fisher(X,Y)
```

---

simultest	<i>Two-sample simultaneous tests on high-dimensional mean vectors and covariance matrices</i>
-----------	---

---

**Description**

This function implements six two-sample simultaneous tests on high-dimensional mean vectors and covariance matrices. Let  $\mathbf{X} \in \mathbb{R}^p$  and  $\mathbf{Y} \in \mathbb{R}^p$  be two  $p$ -dimensional populations with mean vectors  $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$  and covariance matrices  $(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$ , respectively. The problem of interest is the simultaneous inference on the equality of mean vectors and covariance matrices of the two populations:

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ and } \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2.$$

Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . We denote  $\text{dataX} = (\mathbf{X}_1, \dots, \mathbf{X}_{n_1})^\top \in \mathbb{R}^{n_1 \times p}$  and  $\text{dataY} = (\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2})^\top \in \mathbb{R}^{n_2 \times p}$ .

**Usage**

```
simultest(dataX, dataY, method='pe.fisher', delta_mean=NULL, delta_cov=NULL)
```

**Arguments**

dataX	an $n_1$ by $p$ data matrix
dataY	an $n_2$ by $p$ data matrix
method	the method type (default = 'pe.fisher'); chosen from <ul style="list-style-type: none"> <li>• 'cauchy': the simultaneous test via Cauchy combination; see <a href="#">simultest.cauchy</a> for details.</li> </ul>

- 'chisq': the simultaneous test via chi-squared approximation; see [simultest.chisq](#) for details.
  - 'fisher': the simultaneous test via Fisher's combination; see [simultest.fisher](#) for details.
  - 'pe.cauchy': the PE simultaneous test via Cauchy combination; see [simultest.pe.cauchy](#) for details.
  - 'pe.chisq': the PE simultaneous test via chi-squared approximation; see [simultest.pe.chisq](#) for details.
  - 'pe.fisher': the PE simultaneous test via Fisher's combination; see [simultest.pe.fisher](#) for details.
- delta\_mean the thresholding value used in the construction of the PE component for the mean test statistic. It is needed only in PE methods such as method='pe.cauchy', method='pe.chisq', and method='pe.fisher'; see [simultest.pe.cauchy](#), [simultest.pe.chisq](#), and [simultest.pe.fisher](#) for details. The default is NULL.
- delta\_cov the thresholding value used in the construction of the PE component for the covariance test statistic. It is needed only in PE methods such as method='pe.cauchy', method='pe.chisq', and method='pe.fisher'; see [simultest.pe.cauchy](#), [simultest.pe.chisq](#), and [simultest.pe.fisher](#) for details. The default is NULL.

## Value

method the method type  
 stat the value of test statistic  
 pval the p-value for the test.

## References

- Chen, S. X. and Qin, Y. L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Annals of Statistics*, 38(2):808–835.
- Li, J. and Chen, S. X. (2012). Two sample tests for high-dimensional covariance matrices. *The Annals of Statistics*, 40(2):908–940.
- Yu, X., Li, D., and Xue, L. (2022). Fisher's combined probability test for high-dimensional covariance matrices. *Journal of the American Statistical Association*, (in press):1–14.
- Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

## Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
simultest(X,Y)
```

## Description

This function implements the two-sample simultaneous test on high-dimensional mean vectors and covariance matrices using Cauchy combination. Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . Let  $p_{CQ}$  and  $p_{LC}$  denote the  $p$ -values associated with the  $l_2$ -norm-based mean test proposed in Chen and Qin (2010) (see [meantest.cq](#) for details) and the  $l_2$ -norm-based covariance test proposed in Li and Chen (2012) (see [covtest.lc](#) for details), respectively. The simultaneous test statistic via Cauchy combination is defined as

$$C_{n_1, n_2} = \frac{1}{2} \tan((0.5 - p_{CQ})\pi) + \frac{1}{2} \tan((0.5 - p_{LC})\pi).$$

It has been proved that with some regularity conditions, under the null hypothesis  $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$  and  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ , the two tests are asymptotically independent as  $n_1, n_2, p \rightarrow \infty$ , and therefore  $C_{n_1, n_2}$  asymptotically converges in distribution to a standard Cauchy distribution. The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - F_{Cauchy}(C_{n_1, n_2}),$$

where  $F_{Cauchy}(\cdot)$  is the cdf of the standard Cauchy distribution.

## Usage

```
simultest.cauchy(dataX, dataY)
```

## Arguments

`dataX` an  $n_1$  by  $p$  data matrix  
`dataY` an  $n_2$  by  $p$  data matrix

## Value

`stat` the value of test statistic

`pval` the  $p$ -value for the test.

## References

Chen, S. X. and Qin, Y. L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Annals of Statistics*, 38(2):808–835.

Li, J. and Chen, S. X. (2012). Two sample tests for high-dimensional covariance matrices. *The Annals of Statistics*, 40(2):908–940.

Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

**Examples**

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
simultest.cauchy(X,Y)
```

---

simultest.chisq      *Two-sample simultaneous test using chi-squared approximation*

---

**Description**

This function implements the two-sample simultaneous test on high-dimensional mean vectors and covariance matrices using chi-squared approximation. Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . Let  $M_{CQ}/\hat{\sigma}_{MCQ}$  denote the  $l_2$ -norm-based mean test statistic proposed in Chen and Qin (2010) (see [meantest.cq](#) for details), and let  $T_{LC}/\hat{\sigma}_{TLC}$  denote the  $l_2$ -norm-based covariance test statistic proposed in Li and Chen (2012) (see [covtest.lc](#) for details). The simultaneous test statistic via chi-squared approximation is defined as

$$S_{n_1, n_2} = M_{CQ}^2/\hat{\sigma}_{MCQ}^2 + T_{LC}^2/\hat{\sigma}_{TLC}^2.$$

It has been proved that with some regularity conditions, under the null hypothesis  $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$  and  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ , the two tests are asymptotically independent as  $n_1, n_2, p \rightarrow \infty$ , and therefore  $S_{n_1, n_2}$  asymptotically converges in distribution to a  $\chi_2^2$  distribution. The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - F_{\chi_2^2}(S_{n_1, n_2}),$$

where  $F_{\chi_2^2}(\cdot)$  is the cdf of the  $\chi_2^2$  distribution.

**Usage**

```
simultest.chisq(dataX, dataY)
```

**Arguments**

```
dataX            n1 by p data matrix
dataY            n2 by p data matrix
```

**Value**

```
stat the value of test statistic
pval the p-value for the test.
```

**References**

Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

**Examples**

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
simultest.chisq(X,Y)
```

---

simultest.fisher

*Two-sample simultaneous test using Fisher's combination*


---

**Description**

This function implements the two-sample simultaneous test on high-dimensional mean vectors and covariance matrices using Fisher's combination. Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . Let  $p_{CQ}$  and  $p_{LC}$  denote the  $p$ -values associated with the  $l_2$ -norm-based mean test proposed in Chen and Qin (2010) (see [meantest.cq](#) for details) and the  $l_2$ -norm-based covariance test proposed in Li and Chen (2012) (see [covtest.lc](#) for details), respectively. The simultaneous test statistic via Fisher's combination is defined as

$$J_{n_1, n_2} = -2 \log(p_{CQ}) - 2 \log(p_{LC}).$$

It has been proved that with some regularity conditions, under the null hypothesis  $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$  and  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ , the two tests are asymptotically independent as  $n_1, n_2, p \rightarrow \infty$ , and therefore  $J_{n_1, n_2}$  asymptotically converges in distribution to a  $\chi_4^2$  distribution. The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - F_{\chi_4^2}(J_{n_1, n_2}),$$

where  $F_{\chi_4^2}(\cdot)$  is the cdf of the  $\chi_4^2$  distribution.

**Usage**

```
simultest.fisher(dataX, dataY)
```

**Arguments**

```
dataX      an  $n_1$  by  $p$  data matrix
dataY      an  $n_2$  by  $p$  data matrix
```

**Value**

```
stat the value of test statistic
pval the p-value for the test.
```

## References

- Chen, S. X. and Qin, Y. L. (2010). A two-sample test for high-dimensional data with applications to gene-set testing. *Annals of Statistics*, 38(2):808–835.
- Li, J. and Chen, S. X. (2012). Two sample tests for high-dimensional covariance matrices. *The Annals of Statistics*, 40(2):908–940.
- Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

## Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
simultest.fisher(X,Y)
```

---

simultest.pe.cauchy      *Two-sample PE simultaneous test using Cauchy combination*

---

## Description

This function implements the two-sample PE simultaneous test on high-dimensional mean vectors and covariance matrices using Cauchy combination. Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . Let  $M_{PE}$  and  $T_{PE}$  denote the PE mean test statistic and PE covariance test statistic, respectively. (see [meantest.pe.comp](http://meantest.pe.comp) and [covtest.pe.comp](http://covtest.pe.comp) for details). Let  $p_m$  and  $p_c$  denote their respective  $p$ -values. The PE simultaneous test statistic via Cauchy combination is defined as

$$C_{PE} = \frac{1}{2} \tan((0.5 - p_m)\pi) + \frac{1}{2} \tan((0.5 - p_c)\pi).$$

It has been proved that with some regularity conditions, under the null hypothesis  $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$  and  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ , the two tests are asymptotically independent as  $n_1, n_2, p \rightarrow \infty$ , and therefore  $C_{PE}$  asymptotically converges in distribution to a standard Cauchy distribution. The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - F_{Cauchy}(C_{PE}),$$

where  $F_{Cauchy}(\cdot)$  is the cdf of the standard Cauchy distribution.

## Usage

```
simultest.pe.cauchy(dataX, dataY, delta_mean=NULL, delta_cov=NULL)
```

**Arguments**

dataX	an $n_1$ by $p$ data matrix
dataY	an $n_2$ by $p$ data matrix
delta_mean	a scalar; the thresholding value used in the construction of the PE component for mean test; see <a href="#">meantest.pe.comp</a> for details.
delta_cov	a scalar; the thresholding value used in the construction of the PE component for covariance test; see <a href="#">covtest.pe.comp</a> for details.

**Value**

stat the value of test statistic  
 pval the p-value for the test.

**References**

Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

**Examples**

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
simultest.pe.cauchy(X,Y)
```

---

simultest.pe.chisq      *Two-sample PE simultaneous test using chi-squared approximation*

---

**Description**

This function implements the two-sample PE simultaneous test on high-dimensional mean vectors and covariance matrices using chi-squared approximation. Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . Let  $M_{PE}$  and  $T_{PE}$  denote the PE mean test statistic and PE covariance test statistic, respectively. (see [meantest.pe.comp](#) and [covtest.pe.comp](#) for details). The PE simultaneous test statistic via chi-squared approximation is defined as

$$S_{PE} = M_{PE}^2 + T_{PE}^2.$$

It has been proved that with some regularity conditions, under the null hypothesis  $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$  and  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ , the two tests are asymptotically independent as  $n_1, n_2, p \rightarrow \infty$ , and therefore  $S_{PE}$  asymptotically converges in distribution to a  $\chi_2^2$  distribution. The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - F_{\chi_2^2}(S_{PE}),$$

where  $F_{\chi_2^2}(\cdot)$  is the cdf of the  $\chi_2^2$  distribution.

**Usage**

```
simultest.pe.chisq(dataX,dataY,delta_mean=NULL,delta_cov=NULL)
```

**Arguments**

`dataX` an  $n_1$  by  $p$  data matrix  
`dataY` an  $n_2$  by  $p$  data matrix  
`delta_mean` a scalar; the thresholding value used in the construction of the PE component for mean test; see [meantest.pe.comp](#) for details.  
`delta_cov` a scalar; the thresholding value used in the construction of the PE component for covariance test; see [covtest.pe.comp](#) for details.

**Value**

`stat` the value of test statistic

`pval` the p-value for the test.

**References**

Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

**Examples**

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
simultest.pe.chisq(X,Y)
```

---

`simultest.pe.fisher` *Two-sample PE simultaneous test using Fisher's combination*

---

**Description**

This function implements the two-sample PE simultaneous test on high-dimensional mean vectors and covariance matrices using Fisher's combination. Suppose  $\{\mathbf{X}_1, \dots, \mathbf{X}_{n_1}\}$  are i.i.d. copies of  $\mathbf{X}$ , and  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_{n_2}\}$  are i.i.d. copies of  $\mathbf{Y}$ . Let  $M_{PE}$  and  $T_{PE}$  denote the PE mean test statistic and PE covariance test statistic, respectively. (see [meantest.pe.comp](#) and [covtest.pe.comp](#) for details). Let  $p_m$  and  $p_c$  denote their respective  $p$ -values. The PE simultaneous test statistic via Fisher's combination is defined as

$$J_{PE} = -2 \log(p_m) - 2 \log(p_c).$$

It has been proved that with some regularity conditions, under the null hypothesis  $H_0 : \mu_1 = \mu_2$  and  $\Sigma_1 = \Sigma_2$ , the two tests are asymptotically independent as  $n_1, n_2, p \rightarrow \infty$ , and therefore  $J_{PE}$  asymptotically converges in distribution to a  $\chi_4^2$  distribution. The asymptotic  $p$ -value is obtained by

$$p\text{-value} = 1 - F_{\chi_4^2}(J_{PE}),$$

where  $F_{\chi_4^2}(\cdot)$  is the cdf of the  $\chi_4^2$  distribution.

### Usage

```
simultest.pe.fisher(dataX,dataY,delta_mean=NULL,delta_cov=NULL)
```

### Arguments

dataX	an $n_1$ by $p$ data matrix
dataY	an $n_2$ by $p$ data matrix
delta_mean	a scalar; the thresholding value used in the construction of the PE component for mean test; see <a href="#">meantest.pe.comp</a> for details.
delta_cov	a scalar; the thresholding value used in the construction of the PE component for covariance test; see <a href="#">covtest.pe.comp</a> for details.

### Value

stat the value of test statistic

pval the p-value for the test.

### References

Yu, X., Li, D., Xue, L., and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *Journal of the American Statistical Association*, (in press):1–14.

### Examples

```
n1 = 100; n2 = 100; pp = 500
set.seed(1)
X = matrix(rnorm(n1*pp), nrow=n1, ncol=pp)
Y = matrix(rnorm(n2*pp), nrow=n2, ncol=pp)
simultest.pe.fisher(X,Y)
```

# Index

covtest, [2](#), [3](#)  
covtest.clx, [3](#), [4](#), [6](#), [9](#)  
covtest.lc, [3](#), [5](#), [6](#), [7](#), [9](#), [18–20](#)  
covtest.pe.cauchy, [3](#), [6](#)  
covtest.pe.comp, [3](#), [4](#), [7](#), [21–24](#)  
covtest.pe.fisher, [4](#), [9](#)

meantest, [2](#), [10](#)  
meantest.clx, [10](#), [11](#), [13](#), [15](#)  
meantest.cq, [10](#), [12](#), [13–15](#), [18–20](#)  
meantest.pe.cauchy, [10](#), [13](#)  
meantest.pe.comp, [10](#), [14](#), [21–24](#)  
meantest.pe.fisher, [10](#), [15](#)

PEtests-package, [2](#)

simultest, [2](#), [16](#)  
simultest.cauchy, [16](#), [18](#)  
simultest.chisq, [17](#), [19](#)  
simultest.fisher, [17](#), [20](#)  
simultest.pe.cauchy, [17](#), [21](#)  
simultest.pe.chisq, [17](#), [22](#)  
simultest.pe.fisher, [17](#), [23](#)