

# Package ‘Hotelling’

September 11, 2018

**Version** 1.0-5

**Date** 2018-09-11

**Title** Hotelling's  $T^2$  Test and Variants

**Description** A set of R functions which implements Hotelling's  $T^2$  test and some variants of it. Functions are also included for Aitchison's additive log ratio and centred log ratio transformations.

**Depends** corpcor

**License** GPL ( $\geq 2$ )

**URL** <http://www.stat.auckland.ac.nz/showperson?firstname=James&surname=Curran>

**RoxygenNote** 6.1.0

**NeedsCompilation** no

**Author** James Curran [aut, cre]

**Maintainer** James Curran <j.curran@auckland.ac.nz>

**Repository** CRAN

**Date/Publication** 2018-09-11 06:00:07 UTC

## R topics documented:

Hotelling-package . . . . .	2
alr . . . . .	2
bottle.df . . . . .	3
clr . . . . .	4
container.df . . . . .	5
hotelling.stat . . . . .	5
hotelling.test . . . . .	6
plot.hotelling.test . . . . .	8
print.hotelling.test . . . . .	9

<b>Index</b>	<b>11</b>
--------------	-----------

---

Hotelling-package	<i>Hotelling A set of R functions and data sets which implements Hotelling's <math>T^2</math> test, and some variants of it. Functions are also included for Aitchison's additive log ratio and centred log ratio transformations.</i>
-------------------	--

---

### Description

Hotelling A set of R functions and data sets which implements Hotelling's  $T^2$  test, and some variants of it. Functions are also included for Aitchison's additive log ratio and centred log ratio transformations.

### Author(s)

James Curran Maintainer: James Curran <j.curran@auckland.ac.nz> ~~ The author and/or maintainer of the package ~~

---

alr	<i>Additive log ratio transformation</i>
-----	--

---

### Description

Aitchison's additive log ratio transformation for compositional data

### Usage

```
alr(form, data, group = NULL)
```

### Arguments

form	a formula which specifies the denominator variable as the response
data	a data frame in which the data is stored
group	if not NULL then a character string specifying the name of the grouping variable

### Details

This function will give a warning if zeros are present because the transformed data will have -Infs.

### Value

a data frame with the ALR transformation applied to data. Each row in the data frame is standardized with respect to a specific variable by dividing by that variable. The logarithms of the resulting ratios are returned. If a grouping variable is specified, then this is preserved.

**Author(s)**

James M. Curran

**References**

Aitchison, J. (1986). "The Statistical Analysis of Compositional Data", Chapman and Hall, reprinted in 2003 with additional material by The Blackburn Press

**See Also**

clr

**Examples**

```
data(bottle.df)

## transform with respect to manganese
alr(Mn~., bottle.df, "Number")

## transform the data with respect to barium, but removing the
## grouping in column 1
alr(Ba~., bottle.df[,-1])
```

---

bottle.df

*Bottle data*

---

**Description**

This data contains the elemental concentration of five different elements (Manganese, Barium, Strontium, Zirconium, and Titanium) in samples of glass taken from six different Heineken beer bottles. 20 measurements were taken from each bottle.

**References**

R. L. Bennett. *Aspects of the analysis and interpretation of glass trace evidence*. Master's thesis, Department of Chemistry, University of Waikato, 2002.

---

clr *Centered log ratio transformation*

---

**Description**

Aitchison's centered log ratio transformation for compositional data

**Usage**

```
clr(data, group = NULL)
```

**Arguments**

data            a data frame in which the data is stored  
group           if not NULL then a character string specifying the name of the grouping variable

**Details**

This function will give a warning if zeros are present because the transformed data will have -Infs.

**Value**

a data frame with the CLR transformation applied to data. Each row in the data frame is standardized by dividing by the geometric mean of that row. The logarithms of the resulting ratios are returned. If a grouping variable is specified, then this is preserved.

**Author(s)**

James M. Curran

**References**

Aitchison, J. (1986). "The Statistical Analysis of Compositional Data", Chapman and Hall, reprinted in 2003 with additional material by The Blackburn Press

**See Also**

alr

**Examples**

```
data(bottle.df)

## transform preserving grouping
clr(bottle.df, "Number")

## transform the data but remove the
## grouping in column 1
```

```
clr(bottle.df[,-1])
```

---

container.df	<i>Container data</i>
--------------	-----------------------

---

### Description

This data contains the elemental concentration of nine different elements (Titanium, Aluminium, Iron, Manganese, Magnesium, Calcium, Barium, Strontium, and Zirconium) in specimens of glass taken from two different containers. Ten measurements were taken from each container.

### References

Jose R. Almirall. Discrimination of glass samples by solution based ICP-OES PhD thesis, Department of Chemistry, Florida International University, 1998.

---

hotelling.stat	<i>Calculate Hotelling's two sample T-squared test statistic</i>
----------------	--

---

### Description

Calculate Hotelling's T-squared test statistic for the difference in two multivariate means.

### Usage

```
hotelling.stat(x, y, shrinkage = FALSE)
```

### Arguments

x	a nx by p matrix containing the data points from sample 1
y	a ny by p matrix containing the data points from sample 2
shrinkage	set to TRUE if the covariance matrices are to be estimated using Schaefer and Strimmer's James-Stein shrinkage estimator

### Details

Note, the sample size requirements are that  $n_x + n_y - 1 > p$ . The procedure will stop if this is not met and the shrinkage estimator is not being used. The shrinkage estimator has not been rigorously tested for this application (small p, smaller n).

**Value**

A list containing the following components:

statistic	Hotelling's (unscaled) T-squared statistic
m	The scaling factor - this can be used by multiplying it with the test statistic, or dividing the critical F value
df	a vector of length containing the numerator and denominator degrees of freedom
nx	The sample size of sample 1
ny	The sample size of sample 2
p	The number of variables to be used in the comparison

**Author(s)**

James M. Curran

**References**

Hotelling, H. (1931). "The generalization of Student's ratio." *Annals of Mathematical Statistics* 2 (3): 360–378.

Schaefer, J., and K. Strimmer (2005). "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics." *Statist. Appl. Genet. Mol. Biol.* 4: 32.

Opgen-Rhein, R., and K. Strimmer (2007). "Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach." *Statist. Appl. Genet. Mol. Biol.* 6: 9.

**Examples**

```
data(container.df)
split.data = split(container.df[, -1], container.df$gp)
x = split.data[[1]]
y = split.data[[2]]
hotelling.stat(x, y)
hotelling.stat(x, y, TRUE)
```

---

hotelling.test

*Two-sample Hotelling's T-squared test*

---

**Description**

Performs a two-sample Hotelling's T-squared test for the difference in two multivariate means

**Usage**

```
hotelling.test(x, ...)

## Default S3 method:
hotelling.test(x, y, shrinkage = FALSE, perm = FALSE,
  B = 10000, progBar = (perm && TRUE), ...)

## S3 method for class 'formula'
hotelling.test(x, data = NULL, pair = c(1, 2), ...)
```

**Arguments**

x	a matrix containing the data points from sample 1 or a formula specifying the elements to be used as a response and the grouping variable as a predictor
...	any additional arguments. This is useful to pass the optional arguments for the default call from the formula version
y	a matrix containing the data points from sample 2
shrinkage	if TRUE then Shaefer and Strimmer's James-Stein shrinkage estimator is used to calculate the sample covariance matrices
perm	if TRUE then permutation testing is used to estimate the non-parametric P-value for the hypothesis test
B	if perm is TRUE, then B is the number of permutations to perform
progBar	if TRUE and perm is TRUE then a progress bar will be displayed whilst the permutation procedure is carried out
data	a data frame needs to be specified if a formula is to be used to perform the test
pair	a vector of length two which can be used when the grouping factor has more than two levels to select different pairs of groups. For example for a 3-level factor, pairs could be set to c(1, 3) to perform Hotelling's test between groups 1 and 3

**Value**

A list (which is also of class 'hotelling.test') with the following elements:

stats	a list containing all of the output from hotelling.stat
pval	the P-value from the test
results	if perm == TRUE, then all of the permutation test statistics are stored in results

**Methods (by class)**

- default: Two-sample Hotelling's T-squared test
- formula: Two-sample Hotelling's T-squared test

**Author(s)**

James M. Curran

## References

- Hotelling, H. (1931). “The generalization of Student’s ratio.” *Annals of Mathematical Statistics* 2 (3): 360–378.
- Schaefer, J., and K. Strimmer (2005). “A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics.” *Statist. Appl. Genet. Mol. Biol.* 4: 32.
- Opgen-Rhein, R., and K. Strimmer (2007). “Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach.” *Statist. Appl. Genet. Mol. Biol.* 6: 9.
- Campbell, G.P. and J. M. Curran (2009). “The interpretation of elemental composition measurements from forensic glass evidence III.” *Science and Justice*, 49(1),2-7.

## See Also

hotelling.stat

## Examples

```
data(container.df)
fit = hotelling.test(~gp, data = container.df)
fit

subs.df = container.df[1:10,]
subs.df$gp = rep(1:2, c(5,5))
fitPerm = hotelling.test(Al+Fe~gp, data = subs.df, perm = TRUE)
fitPerm
plot(fitPerm)

data(bottle.df)
fit12 = hotelling.test(~Number, data = bottle.df)
fit12

fit23 = hotelling.test(~Number, data = bottle.df, pair = c(2,3))
fit23
```

---

plot.hotelling.test     *Plots the results from a permutation based version of Hotelling’s T-squared test for the difference in two multivariate sample means*

---

## Description

Plots a histogram of the distribution of the permuted test statistics for a permutation version of Hotelling’s T-squared

## Usage

```
## S3 method for class 'hotelling.test'
plot(x, ...)
```



### Arguments

x                    an object of type `hotelling.test`  
...                  any additional arguments to be passed to the `hist` command

### Details

This function only works if you have performed a permutation test. It will return an error message if not. It could be programmed to draw the relevant F distribution in the standard case, but this seems rather pointless.

### Author(s)

James M. Curran

### Examples

```
data(bottle.df)
bottle.df = subset(bottle.df, Number == 1)
bottle.df$Number = rep(1:2,c(10,10))
fit = hotelling.test(~Number, bottle.df, perm = TRUE)
plot(fit)
plot(fit, col = "lightblue")
```

---

`print.hotelling.test`    *Prints the results from a Hotelling's T-squared test for the difference in two multivariate sample means*

---

### Description

Prints the test statistic, degrees of freedom and P-value from Hotelling's T-squared test for the difference in two multivariate sample means

### Usage

```
## S3 method for class 'hotelling.test'
print(x, ...)
```

### Arguments

x                    an object of type `hotelling.test`  
...                  any additional arguments to be passed to the `hist` command

### Author(s)

James M. Curran

**Examples**

```
data(bottle.df)
bottle.df = subset(bottle.df, Number == 1)
bottle.df$Number = rep(1:2,c(10,10))
fit = hotelling.test(~Number, bottle.df, perm = TRUE)
fit
fit = hotelling.test(~Number, bottle.df)
fit

## an explicit call
print(fit)
```

# Index

\*Topic **datasets**

bottle.df, 3  
container.df, 5

\*Topic **htest**

hotelling.stat, 5  
hotelling.test, 6

\*Topic **package**

Hotelling-package, 2

\*Topic **plot**

plot.hotelling.test, 8

\*Topic **print**

print.hotelling.test, 9

\*Topic **transformation**

alr, 2  
clr, 4

alr, 2

bottle.df, 3

clr, 4

container.df, 5

hotel.stat (hotelling.stat), 5

hotel.test (hotelling.test), 6

Hotelling (Hotelling-package), 2

Hotelling-package, 2

hotelling.stat, 5

hotelling.test, 6

plot.hotelling.test, 8

print.hotelling.test, 9