Package: smsets (via r-universe)

October 21, 2026

Date 2026-01-22

Title Simple Multivariate Statistical Estimation and Tests

Version 2.0.0

Description A collection of simple parameter estimation and tests for the comparison of multivariate means and variation, to accompany Chapters 4 and 5 of the book Multivariate Statistical Methods. A Primer (5th edition), by Manly BFJ, Navarro Alberto JA & Gerow K (2024) <doi:10.1201/9781003453482>.

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Depends R (>= 3.6)

LazyData true

Imports stringr, data.table, Hotelling, biotools, effectsize, crayon

Suggests car, knitr, rmarkdown, testthat (>= 3.0.0),

Config/testthat/edition 3

VignetteBuilder knitr

URL https://github.com/ganava4/smsets

BugReports https://github.com/ganava4/smsets/issues

Config/pak/sysreqs libicu-dev

Repository https://ganava4.r-universe.dev

Date/Publication 2026-01-24 17:18:52 UTC

RemoteUrl https://github.com/ganava4/smsets

RemoteRef HEAD

RemoteSha 09f32814f9c8c73cab943688b129fd417958e11a

Contents

BoxM.F . . . 2
Hotelling.mat . . . 4
LeveneT2 . . . 5
Levenetests2s.mv . . . 7
Levenetestsms.mv . . . 9
OnewayMANOVA . . . 11
Penrose.dist . . . 12
print.BoxM.F . . . 14
print.Hotelling.mat . . . 15
print.LeveneT2 . . . 16
print.Levenetests2s.mv . . . 17
print.Levenetestsms.mv . . . 18
print.OnewayMANOVA . . . 20
print.Penrose.dist . . . 21
print.ttests2s.mv . . . 22
print.VanValen . . . 23
skulls . . . 25
sparrows . . . 26
ttests2s.mv . . . 27
VanValen . . . 29

Index **31**

`BoxM.F` *F approximation of Box's M test*

Description

An R function which implements an F approximation for testing the homogeneity of covariance matrices by Box's M. This is an alternative approach to the chi square approximation which requires group sample-sizes to be at least 20.

Usage

```
BoxM.F(x, group)
```

Arguments

`x`	A data frame with $p + 1$ columns (one factor and p response variables).
`group`	The classification factor defining m samples or groups. It must be one of the columns in `x`.

Details

For m samples, the M statistic is given by the equation

$$M = \frac{\prod_{j=1}^{m} |\mathbf{C}_j|^{(n_j-1)/2}}{|\mathbf{C}|^{(n-m)/2}}$$

where

n_j is the sample size of the j-th sample,

$|\mathbf{C}_j|$ is the determinant of the covariance matrix for the jth sample,

$|\mathbf{C}|$ is the determinant of the pooled covariance matrix,

n is the total number of observations.

Large values of M provide evidence that the samples are not from populations with the same covariance matrix. In addition to the observed *M*-value itself, the F-approximation involves the sample sizes and the number of variables analyzed. See the reference for details. Box's test is sensitive to deviations from normality in the distribution of the variables.

Value

Returns an object of class `"BoxM.F"`, a list containing the following components:

`name`	A character string describing the function.
`Cov.Mat`	A list containing the *m* sample covariance matrices
`Cov.pooled`	The pooled covariance matrix
`BoxM.stat`	The approximate F-statistic
`F.BoxM`	The calculated F-statistic
`df.v1`	Numerator degrees of freedom for the F statistic
`df.v2`	Denominator degrees of freedom for the F statistic
`Pvalue`	P-value for the F statistic
`group`	a character string specifying the name of the classification factor defining groups.
`levels.group`	a vector of length *m*, showing the levels in factor `group`.
`data.name`	a character string giving the name of the data.
`variables`	a character string vector containing the variable names.
`data`	the data frame analyzed.

Author(s)

Jorge Navarro Alberto, <ganava4@gmail.com>

References

Manly, B.F.J., Navarro Alberto, J.A. and Gerow, K. (2024) *Multivariate Statistical Methods. A Primer*. 5th Edn. Chapman and Hall/CRC.

Examples

```
data(skulls)
resBoxM.F <- BoxM.F(skulls, Period)
# Brief output
resBoxM.F
```

`Hotelling.mat` *Hotelling's T^2 test with extra information*

Description

An R function which implements Hotelling's T^2 test assuming equal covariance matrices, with extra information.

Usage

```
Hotelling.mat(x, group, level1)
```

Arguments

`x`	a data frame with one two-level factor and *p* response variables.
`group`	two-level factor defining groups. It must be one of the columns in `x`.
`level1`	a character string identifying Sample 1. The string must be one of the factor levels in `group`.

Details

This function is a simplified version of the function `hotelling.test` implemented in the `Hotelling` package for the comparison of mean values of two multivariate samples, under the assumption that covariance matrices are equal. The `summary` methods in `Hotelling.mat` gives more detailed information of the calculations behind the T^2 test.

Value

Returns an object of class `"Hotelling.mat"`, a list containing the following components:

`name`	A character string describing the function.
`T2.list`	A list containing two data frames with the mean vector for the two samples, two covariance matrices, one ma
`group`	a character string specifying the name of the two-level factor defining groups.
`levels.group`	a vector of length two, showing the two levels in factor `group`.
`data.name`	a character string giving the name of the data.
`data`	the data frame analyzed.

Author(s)

Jorge Navarro Alberto, <ganava4@gmail.com>

References

Manly, B.F.J., Navarro Alberto, J.A. and Gerow, K. (2024) *Multivariate Statistical Methods. A Primer*. 5th Edn. Chapman and Hall/CRC.

Examples

```
data(sparrows)
results.T2 <- Hotelling.mat(sparrows, group = Survivorship, level1 = "S")
# Brief output
results.T2
```

`LeveneT2` *Levene's test for two multivariate samples based on Hotelling's T^2 test with extra information*

Description

An R function for the comparison of multivariate variation in two samples, which implements Levene's test based on Hotelling's T^2.

Usage

```
LeveneT2(x, group, level1, var.equal = TRUE)
```

Arguments

`x`	A data frame with one two-level factor and *p* response variables.
`group`	Two-level factor defining groups. It must be one of the columns in `x`.
`level1`	A character string identifying Sample 1. The string must be one of the factor levels in `group`.
`var.equal`	A logical variable indicating whether to treat the within-sample covariance matrices of absolute deviations around medians for samples 1 and 2 as equal or not. The default is `TRUE`. If the within-sample covariance matrices of absolute deviations around medians are not assumed equal (`FALSE`), Hotelling's `T^2` test is performed using the Nel and van der Merwe's (1986) solution to the multivariate Behrens-Fisher problem as implemented in **Hotelling** package (Curran and Hersh, 2021).

Details

`LeveneT2` makes use of Hotelling's T^2 to test the variation in two multivariate samples. This test is an alternative procedure that should be more robust than Box's test which is known to be rather sensitive to the assumption that the samples are from multivariate normal distributions.

In `LeveneT2` the data values are transformed into absolute deviations from their respective sample medians

$$ADM_{ijk} = |x_{ijk} - M_{jk}|$$

where

x_{ijk} is the value of variable X_k for the ith individual in sample j, and

M_{jk} is the median of X_k in sample j.

The unequal variation question between samples $j = 1$ and $j = 2$ becomes a T^2-test for the difference of the mean ADM vectors.

Value

Returns an object of class "LeveneT2", a list containing the following components:

name	A character string describing the function.
medians	A list containing two vectors. The first vector medians1 contains the medians for all variables in sample 1 a
bygroup.data	A list with two data frames matlevel1 and matlevel2 containing the original variables for samples 1 and
absdev.median	A list with two data frames abs.dev.median1 and abs.dev.median2 containing the absolute deviations fr
LeveneT2.test	A list of class hotelling.test containing the list stats and the scalar pval, produced by function hotel
var.equal	a logical variable indicating whether the two variances were treated as being equal TRUE or not FALSE.
group	a character string specifying the name of the two-level factor defining groups.
levels.group	a vector of length two, showing the two levels in factor group.
data.name	a character string giving the name of the data.
variables	a character string vector containing the variable names.
data	the data frame analyzed.

The extractor function print.LeveneT2 returns an annotated output of the test.

Author(s)

Jorge Navarro Alberto, <ganava4@gmail.com>

References

Curran, J. and Hersh, T. (2021). *Hotelling: Hotelling's T^2 Test and Variants*. R package version 1.0-8, https://CRAN.R-project.org/package=Hotelling.

Manly, B.F.J., Navarro Alberto, J.A. and Gerow, K. (2024) *Multivariate Statistical Methods. A Primer*. 5th Edn. Chapman and Hall/CRC.

Nel, D.G. and van de Merwe, C.A. (1986). A solution to the multivariate Behrens-Fisher problem. *Comm. Statist. Theor. Meth.*, A15, 12, 3719-3736.

Examples

```
data(sparrows)
LeveneT2.sparrows <- LeveneT2(sparrows, group = Survivorship, level1 = "S",
                              var.equal = TRUE)
# Brief output
LeveneT2.sparrows
```

`Levenetests2s.mv` *Multiple two-sample Levene tests for the comparison of variation in multivariate data*

Description

Performs multiple two-sample Levene tests, based on two-sample t-tests applied to absolute differences around medians for more than one response vector, with corrected significance levels using any of the adjustment methods for multiple comparisons offered by `p.adjust`. This function includes the argument `alternative =` useful to specify the type of alternative, either one-sided (lower-/ upper-tail) or two-sided. Effects sizes are also computed with respect to the two-sample t-tests.

Usage

```
Levenetests2s.mv(
  x,
  group,
  level1,
  alternative = "two.sided",
  var.equal = FALSE,
  P.adjust = "none",
  unit = "units"
)
```

Arguments

`x`	a data frame with one two-level factor and *p* response variables.
`group`	two-level factor defining groups. It must be one of the columns in `x`.
`level1`	a character string identifying Sample 1. The string must be one of the factor levels in `group`.
`alternative`	a character string specifying the alternative hypothesis, must be one of `"two.sided"` (default), `"greater"` or `"less"`. You can specify just the initial letter.
`var.equal`	a logical variable indicating whether to treat the two variances as being equal. If `TRUE` then the pooled variance is used to estimate the variance otherwise the Welch (or Satterthwaite) approximation to the degrees of freedom is used.
`P.adjust`	p-value correction method, a character string. Can be abbreviated. See 'Details'.
`unit`	Physical units of the response variable useful to fully characterize raw effect sizes

Details

This function focuses on the univariate Levene test for the comparison of mean values for two samples, when more than one variable is involved in the data analysis, so that type one error rates ("false significances") in the series of Levene tests are adjusted according to the number of response

variables analyzed. The pairwise comparisons between the two levels in group with corrections for multiple testing are made over more than one response vector.

The methods implemented in P.adjust are the same as those contained in the p.adjust.methods: "bonferroni", "holm", "hochberg", "hommel", "BH", (Benjamini-Hochberg) or its alias "fdr" (False Discovery Rate), and "BY" (Benjamini & Yekutieli). The default pass-through option ("none") is also included.

Value

Returns an object of class "Levenetests2s.mv", a list containing the following components:

name	A character string describing the function.
medians	A list containing two vectors of length *p*, being *p* the number of response variables. medians1 and medians
absdev.median	A list containing two data frames, abs.dev.median1 and abs.dev.median2, corresponding to the absolute
means.absdev	A list containing two vectors of length *p*, (means.absdev1 and means.absdev1), corresponding to the mea
vars.absdev	A list containing two vectors of length *p*, (vars.absdev1 and vars.absdev1), corresponding to the varianc
t.list	A list containing *p* vectors of length 5, each vector containing the t-statistic, the degrees of freedom, the adj
alternative	A character string specifying the alternative hypothesis chosen.
var.equal	A logical variable indicating whether the two variances were treated as being equal TRUE or not FALSE.
P.adjust	A character string indicating the correction method chosen
group	A character string specifying the name of the two-level factor defining groups.
levels.group	a vector of length two showing the two levels in factor group.
data.name	a character string giving the name of the data.
data	the data frame analyzed.

The extractor function print.Levenetests2s.mv returns an annotated output of the Levene tests (or, equivalently, the two-sample t-tests applied to the absolute differences around medians).

Author(s)

Jorge Navarro Alberto, <ganava4@gmail.com>

References

Manly, B.F.J., Navarro Alberto, J.A. and Gerow, K. (2024) *Multivariate Statistical Methods. A Primer*. 5th Edn. Chapman and Hall/CRC.

Hedges, L. V. 1981. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics* 6(2): 107–128.

Examples

```
data(sparrows)
res.Levene2s.mv <- Levenetests2s.mv(sparrows, Survivorship, "S",
                                    alternative = "less", var.equal = TRUE,
                                    P.adjust = "bonferroni", unit = "mm")
res.Levene2s.mv
```

`Levenetestsms.mv` *Robust Levene's Tests for the Comparison of Variation of m Samples in Multivariate Data*

Description

Performs Levene's tests for *m* samples on *p* responses, based on (univariate) One-Way ANOVAs and One-Way MANOVAs applied to absolute differences around medians. Significance levels of the univariate tests of variation can be corrected using any of the adjustment methods for multiple comparisons offered by `p.adjust`. Effects sizes are also computed with respect to the One-Way ANOVAs.

Usage

```
Levenetestsms.mv(x, group, var.equal = FALSE, P.adjust = "none")
```

Arguments

`x`	a data frame containing a factor with *m* levels and *p* response variables.
`group`	a factor with *m* levels defining samples. It must be one of the columns in `x`.
`var.equal`	a logical variable indicating whether to treat the *m* variances of absolute deviations around medians (the variances of the measure of variation among samples!) as being equal for the One-Way ANOVAs. If `TRUE` then a simple F test for the equality of means is performed. If `FALSE` the Welch (1951) approximation to the degrees of freedom is used, as implemented in `oneway.test`.
`P.adjust`	p-value correction method of univariate Levene's tests (One-Way ANOVAs), a character string. Can be abbreviated. See 'Details'.

Details

This function focuses on robust Levene's tests, both univariate and multivariate, for the comparison of variation among *m* samples in multivariate data. These tests can be chosen as alternatives to Box's test which is sensitive to deviations from normality. The application of Levene's test one variable at a time from a set of *p* variables can be computed by repeating *p* times **car**'s package function `leveneTest` (Fox and Weisberg 2019), when `center = median`. However, there are *p* univariate Levene's tests possible, each one consisting of one-way ANOVAs applied to the absolute deviations around medians. Therefore, the p-values produced in the ANOVAs can be subject to corrections for multiple testing, depending on the number of response variables analyzed. The methods implemented in `P.adjust` are the same as those contained in the `p.adjust.methods`: `"bonferroni"`, `"holm"`, `"hochberg"`, `"hommel"`, `"BH"`, (Benjamini-Hochberg) or its alias `"fdr"` (False Discovery Rate), and `"BY"` (Benjamini & Yekutieli). The default pass-through option (`"none"`) is also included. Four measures of effect size are also computed with respect to the univariate F tests, which are interpreted as effect sizes of variation among samples. User-friendly summaries of all analyses (including the multivariate Levene's test) can be invoked using the `print` method for this function.

Value

Returns an list of class "Levenetestsms.mv", a list containing the following components:

name	A character string describing the function.
medians	A $m \times p$ matrix; the cell (*m*,*p*) contains the median of the *p*-th response in sample *m*.
absdev_medians	A list containing *m* data frames, one data frame for each level of group, and each data frame having *p* colu
df_absdev	A data frame containing the absolute deviations around medians, seen as a compact version of absdev_me
means_absdev	A $m \times p$ matrix; the cell (*m*,*p*) contains the mean absolute deviation around the median of the *p*-th respon
vars_absdev	A $m \times p$ matrix; the cell (*m*,*p*) contains the variance of absolute deviations around the median of the *p*-th
OneWayANOVAs	A list containing the results of the *p* tests for equal means of absolute deviations around medians in a one-
ANOVATables	A list containing *p* analysis of variance tables produced by anova.lm, each table corresponding to a one-w
var.equal	A logical variable indicating whether the two variances were treated as being equal TRUE or not FALSE.
P.adjust	A character string indicating the correction method chosen.
Eff_sizes	A list of length *p* containing four effect size measures for an F-test in one-way ANOVA, and their respecti
OWM_absdev	A list of class "manova" containing the results of the One-Way MANOVA applied to the absolute deviatio
group	A character string specifying the name of the *m*-level factor defining samples.
levels.group	A vector of length *m* showing the levels in factor group.
variables	A vector of length *p* showing the names of response variables.
data.name	A character string giving the name of the data.
data	The data frame analyzed.

The extractor function print.Levenetestsms.mv returns an annotated output of the univariate Levene tests and, optionally, the multivariate Levene's test.

Author(s)

Jorge Navarro Alberto, <ganava4@gmail.com>

References

Ben-Shachar, M., Lüdecke, D., and Makowski, D. (2020). effectsize: Estimation of Effect Size Indices and Standardized Parameters. *Journal of Open Source Software*, 5(56), 2815. doi: 10.21105/joss.02815

Fox, J., and Weisberg, S. (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. https://www.john-fox.ca/Companion/.

Manly, B.F.J., Navarro Alberto, J.A. and Gerow, K. (2024) *Multivariate Statistical Methods. A Primer*. 5th Edn. Chapman and Hall/CRC.

Welch, B.L. (1951). On the comparison of several mean values: an alternative approach. *Biometrika*, 38, 330-336. doi:10.2307/2332579.

Examples

```
data(skulls)
res.Levenems.mv <- Levenetestsms.mv(skulls, Period, var.equal = TRUE,
                                    P.adjust = "bonferroni")
res.Levenems.mv
```

OnewayMANOVA *Tests in One-way MANOVA with extra information*

Description

An R function to test the difference of mean vectors among the levels of a single factor with respect to *p* response variables. Sum of squares and cross-products matrices involved in the MANOVA can be optionally displayed. Test statistics produced are the same as those implemented in `summary.manova`

Usage

```
OnewayMANOVA(x, group)
```

Arguments

`x`	A data frame with one factor and *p* response variables.
`group`	Factor defining groups. It must be one of the columns in `x`.

Details

This function is a simplified version of `manova`, focusing in multivariate analysis of variance for one single factor with respect to *p* responses. The `print` method in `OnewayMANOVA` is similar to that in `summary.manova`, producing the same approximate F tests in the one-way MANOVA. A simplified printout of the sums of squares and product matrices involved in the analysis can optionally be chosen.

Value

Returns an object of class `"OnewayMANOVA"`, a list containing the following components:

`name`	A character string describing the function.
`T`	The total sum of squares and cross-product matrix, defined as $\mathbf{T} = \mathbf{B} + \mathbf{W}$, with $\mathbf{B}$ and $\mathbf{W}$ described below
`W`	The within-sample or residual sum of squares and cross-product matrix.
`B`	The between-sample sum of squares and cross-product matrix
`x.mnv`	An object of class "manova" (and some other classes) produced by function `manova`, to be passed as argumer
`group`	A character string specifying the name of the factor defining groups.
`levels.group`	A vector showing the levels in factor `group`.
`data.name`	A character string giving the name of the data.
`variables`	A character string vector containing the variable names.
`data`	The data frame analyzed.

Author(s)

Jorge Navarro Alberto, <ganava4@gmail.com>

References

Manly, B.F.J., Navarro Alberto, J.A. and Gerow, K. (2024) *Multivariate Statistical Methods. A Primer*. 5th Edn. Chapman and Hall/CRC.

Examples

```
data(skulls)
res.MANOVA <- OnewayMANOVA(skulls, group = Period)
# Brief output
res.MANOVA
```

`Penrose.dist` *Penrose's distance calculator*

Description

Computes Penrose's distance between *m* multivariate populations or samples, when information is available on the means and variances.

Usage

```
Penrose.dist(x, group)
```

Arguments

`x`	A data frame with $p + 1$ columns (one factor and *p* response variables).
`group`	The classification factor defining m samples or groups. It must be one of the variables in `x`.

Details

Let the mean of X_k in population *i* be μ_{ki}, $k = 1, ..., p; i = 1, ..., m$ and assume that the variance of variable X_k is V_k. The Penrose (1953) distance P_{ij} between population *i* and population *j* is given by

$$P_{ij} = \sum_{k=1}^{p} \frac{(\mu_{ki} - \mu_{kj})^2}{pV_k}$$

Penrose's distances between multivariate samples are computed using this expression, but μ_{ki}, μ_{kj} and V_k being replaced by their corresponding sample estimates.

A disadvantage of Penrose's measure is that it does not consider the correlations between the p variables.

The function requires package **biotools** (da Silva, 2017, 2021).

Value

Returns an object of class "Penrose.dist", a list containing the following components:

name	A character string describing the function.
means.vec	A numeric matrix with *p* rows and *m* columns giving the mean of each variable per group.
covs.list	A list containing the *m* sample covariance matrices.
Samp.sizes	A table showing the number of observations used in the calculation of the covariance matrix for each group.
PooledCov	The pooled covariance matrix. This matrix can be accessed and used as an input argument for the calculatio
Penrose.mat	The Penrose distances given as a "matrix" object.
Penros.dist	The Penrose distances given as a "dist" object.
group	A character string specifying the name of the classification factor defining groups.
levels.group	a vector of length *m*, showing the levels in factor group.
data.name	a character string giving the name of the data.
variables	a character string vector containing the variable names.
data	the data frame analyzed.

Author(s)

Jorge Navarro Alberto, <ganava4@gmail.com>

References

da Silva, A.R. (2021). *biotools: Tools for Biometry and Applied Statistics in Agricultural Science*. R package version 4.2. https://cran.r-project.org/package=biotools.

da Silva, A.R., Malafaia, G., and Menezes, I.P.P. (2017). biotools: an R function to predict spatial gene diversity via an individual-based approach. *Genetics and Molecular Research* 16. https://doi.org/10.4238/gmr16029655.

Goslee, S.C. and Urban, D.L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software* 22(7):1-19. DOI:10.18637/jss.v022.i07

Manly, B.F.J., Navarro Alberto, J.A. and Gerow, K. (2024) *Multivariate Statistical Methods. A Primer*. 5th Edn. Chapman and Hall/CRC.

Penrose, L.W. (1953). Distance, size and shape. *Annals of Eugenics* 18: 337-43.

Examples

```
data(skulls)
res.Penrose <- Penrose.dist(x = skulls, group = Period)
# Brief output
res.Penrose
```

`print.BoxM.F` *Prints Box'M test based on an F-statistic*

Description

Prints the results produced by BoxM.F function, with the option to display the matrices involved in the calculations

Usage

```
## S3 method for class 'BoxM.F'
print(x, long = FALSE, ...)
```

Arguments

`x`	an object of class BoxM.F
`long`	a logical variable indicating whether a long output is desired (`TRUE`) or not (`FALSE`, the default). The long output shows the covariance matrix for each group and the pooled covariance matrix.
`...`	further arguments passed to or from other methods.

Value

Displays the results of Box's *M* test for homogeneity of covariance matrices, based on the F-approximation computed by the `BoxM.F` function. The argument `x`, invisibly, as for all print methods, is a list of class "`BoxM.F`". This `print` method provides two sorts of output depending on whether the `long` argument is `TRUE` or `FALSE` (the default). The "short" output displays:

- A heading describing the analysis.
- The data frame analyzed.
- The variables used for the test.
- The factor defining the populations or samples and their levels.
- The value of the Box's *M* statistic, the corresponding approximate *F*-statistic, the degrees of freedom for the numerator and the denominator of the *F*-statistic, and the *p*-value.

In addition to the above information, the "long" output lists:

- The covariance matrix for each sample.
- The pooled covariance matrix.

Examples

```
data(skulls)
resBoxM.F <- BoxM.F(skulls, Period)
# Long output
print(resBoxM.F, long = TRUE)
```

`print.Hotelling.mat` *Prints Hotelling's T^2 test*

Description

Prints the results produced by the `Hotelling.mat` function

Usage

```
## S3 method for class 'Hotelling.mat'
print(x, long = FALSE, ...)
```

Arguments

`x`	an object of class `"Hotelling.mat"`
`long`	a logical variable indicating whether a long output is desired (`TRUE`) or not (`FALSE`, the default)
`...`	further arguments passed to or from other methods.

Value

Displays the results of the comparison of mean values of two multivariate samples, under the assumption that covariance matrices are equal, using Hotelling's T^2 test. The argument `x`, invisibly, as for all print methods, is a list of class `"Hotelling.mat"`. This `print` method provides two sorts of output depending on whether the `long` argument is `TRUE` or `FALSE` (the default). The "short" output displays:

- A description of the analysis.
- The data frame analyzed.
- The labels of the two-level group factor (samples), with an order determined by the user in the `Hotelling.mat` argument `level`.
- The value of Hotelling's T-statistic.
- The value of the F-statistic with its corresponding degrees of freedom for numerator and denominator.
- The P-value.

In addition to this summary, the "long" output shows:

- The mean vectors and covariance matrices for each sample.
- The pooled covariance matrix.
- The inverse of the covariance matrix.

Examples

```
data(sparrows)
results.T2 <- Hotelling.mat(sparrows, group = Survivorship, level1 = "S")
# Long output
print(results.T2, long = TRUE)
```

`print.LeveneT2` *Prints Levene's test based on Hotelling's T^2 test*

Description

Prints the results produced by LeveneT2, consisting of a Levene's test for two multivariate samples based on Hotelling's T^2 test.

Usage

```
## S3 method for class 'LeveneT2'
print(x, long = FALSE, ...)
```

Arguments

x	an object of class "LeveneT2".
long	a logical variable indicating whether a long output is desired (TRUE) or not (FALSE, the default).
...	further arguments passed to or from other methods.

Value

Displays the results of the comparison of multivariate variation in two samples in which data values are transformed into absolute deviations from their respective sample medians, and mean vectors of absolute deviations are compared using Hotelling's T^2 test. The argument x, invisibly, as for all print methods, is a list of class "LeveneT2". This print method provides two sorts of output depending on whether the long argument is TRUE or FALSE (the default). The "short" output displays:

- A description of the analysis.
- The data frame analyzed.
- The names of responses in the data frame.
- The labels of the two-level group factor (samples), with an order determined by the argument level1 in LeveneT2.
- The value of Hotelling's T-statistic.
- The value of the F-statistic with its corresponding degrees of freedom for numerator and denominator. When the within-sample covariance matrices of absolute deviations around medians are not assumed equal (var.equal = FALSE), these degrees of freedom are approximated using the Nel and van der Merwe's (1986) solution to the multivariate Behrens-Fisher problem, as implemented in **Hotelling** package (Curran and Hersh, 2021).

- The P-value.

In addition to the above information, the "long" output lists:

- Sub-data frames containing the original responses and medians, separately for each sample.
- The absolute deviations from sample medians for samples 1 and 2.
- Vectors of mean absolute deviations around medians for samples 1 and 2, used in Hotelling's T test.

References

Curran, J. and Hersh, T. (2021). *Hotelling: Hotelling's T^2 Test and Variants*. R package version 1.0-8, https://CRAN.R-project.org/package=Hotelling.

Nel, D.G. and van de Merwe, C.A. (1986). A solution to the multivariate Behrens-Fisher problem. *Comm. Statist. Theor. Meth.*, A15, 12, 3719-3736.

Examples

```
data(sparrows)
LeveneT2.sparrows <- LeveneT2(sparrows, group = Survivorship, level1 = "S",
                              var.equal = TRUE)
# Long output
print(LeveneT2.sparrows, long = TRUE)
```

---

print.Levenetests2s.mv

Prints multiple two-sample Levene tests for the comparison of variation in multivariate data

---

Description

Prints the results produced by Levenetests2s.mv, consisting of two-sample Levene's tests computed from two-sample t-tests applied to absolute differences around medians for more than one response vector.

Usage

```
## S3 method for class 'Levenetests2s.mv'
print(x, ...)
```

Arguments

x	an object of class "Levenetests2s.mv"
...	further arguments passed to or from other methods.

Details

Summarize

Value

An annotated output of two-sample Levene's tests computed from two-sample t-tests applied to absolute differences around medians for more than one response vector, with (optionally) corrected significance levels. The argument `x`, invisibly, as for all print methods, is a list of class "`Levenetests2s.mv`". This `print` method provides a user-friendly display of particular elements in `x`:

- A description of the analysis.
- The data frame analyzed.
- The labels of the two-level group factor (samples), with an order determined by the user in the `Levenetests2s.mv` argument `level1`.
- The *t*-based Levene's test results for each response variable; these include:
  - The variable name.
  - Sample medians classified by group levels.
  - Means and variances of sample absolute deviations from the median classified by group levels.
  - The value of the *t*-statistic, the degrees of freedom and the *p*-value.
  - Effect sizes: raw and Hedge's (1981). The units of raw effect sizes are shown according to the argument `unit =` in `Levenetests2s.mv`.
- The type of alternative hypothesis for all tests.
- The method of significance level adjustment for multiple comparisons used.

References

Hedges, L. V. 1981. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics* 6(2): 107–128.

Examples

```
data(sparrows)
res.Levene2s.mv <- Levenetests2s.mv(sparrows, Survivorship, "S",
                                    alternative = "less", var.equal = TRUE,
                                    P.adjust = "bonferroni", unit = "mm")
print(res.Levene2s.mv)
```

`print.Levenetestsms.mv`

Prints multiple Levene's tests of multivariate data.

Description

Prints the results produced by `Levenetestsms.mv`, consisting of *m*-sample Levene's tests computed from one-way ANOVAs applied to absolute differences around medians for *p* responses. It optionally produces a single multivariate Levene's test which is basically a one-way MANOVA of absolute differences around medians.

Usage

```
## S3 method for class 'Levenetestsms.mv'
print(
  x,
  format = "oneway.test",
  EffectSize = TRUE,
  multivariate = FALSE,
  mv_statistic = c("Pillai", "Wilks", "Hotelling-Lawley", "Roy"),
  long = FALSE,
  ...
)
```

Arguments

x	an object of class "Levenetestsms.mv".
format	a character string specifying the one-way ANOVA results are formatted, must be one of "oneway.test" (default), "anova.lm" or "both".
EffectSize	a logical variable. If TRUE (default), the measures of effect size η^2, ω^2, ϵ^2, and Cohen's f are returned.
multivariate	a logical variable. If TRUE, the multivariate Levene's test is returned. The test statistic chosen can be set with the optional argument mv_statistic.
mv_statistic	a character string. The name of the test statistic to be used in the multivariate Levene's test (any of the four tests implemented in summary.manova). Pillai's test is the default. Partial matching is used so the name can be abbreviated.
long	A logical variable indicating whether a long output is desired (TRUE) or not (FALSE, the default). The additional objects returned are three matrices: 1) matrix of medians for the original data classified by sample and response variable, 2) matrix of mean absolute deviations around means, and 3) matrix of variances of absolute deviations around means.
...	further arguments passed to or from other methods.

Value

An annotated output of *m*-sample Levene's tests computed from one-way ANOVAs applied to absolute differences around medians for *p* responses, with (optionally) corrected significance levels. The argument x, invisibly, as for all print methods, is a list of class "Levenetestsms.mv". This print method provides a user-friendly display of particular elements in x:

- A description of the analysis.
- The data frame analyzed
- The labels of the group factor with *m* levels (samples).
- The univariate Levene's tests based on the results of one-way ANOVAs for each response variable; These results consist of the *F*-tests displayed in up to two different output formats corresponding to the formatted results produced by the conventional functions oneway.test and anova.lm. P-values are recomputed whenever P.adjust method is different from "none".

In addition to the above information, the long = TRUE output lists:

- Sample medians classified by group levels and variables.
- Means and variances of sample absolute deviations from the median classified by group levels and variables.

If `multivariate = TRUE`, the multivariate Levene's test is appended, based on the test statistic specified in `mv_statistic`. If `multivariate = TRUE` and `mv_statistic` is omitted, Pillai's statistic is chosen by default.

Examples

```
data(skulls)
res.Levenems.mv <- Levenetestsms.mv(skulls, Period, var.equal = TRUE,
                                    P.adjust = "bonferroni")
print(res.Levenems.mv, format = "both", multivariate = TRUE, long = TRUE)
```

---

print.OnewayMANOVA *Prints a one-way MANOVA with extra information*

Description

Prints the results produced by the OnewayMANOVA function

Usage

```
## S3 method for class 'OnewayMANOVA'
print(
  x,
  test = c("Pillai", "Wilks", "Hotelling-Lawley", "Roy"),
  long = FALSE,
  ...
)
```

Arguments

x	An object of class OnewayMANOVA.
test	The name of the test statistic to be used (the four tests implemented in summary.manova). Pillai's test is the default. Partial matching is used so the name can be abbreviated.
long	A logical variable indicating whether a long output is desired (TRUE) or not (FALSE, the default)
...	further arguments passed to or from other methods.

Value

Displays the results of a One-way MANOVA, i.e., the test of the difference of mean vectors among the levels of a single factor with respect to p response variables. The argument `x`, invisibly, as for all print methods, is a list of class "`OnewayMANOVA`". This `print` method provides two sorts of output depending on whether the `long` argument is `TRUE` or `FALSE` (the default). The "short" output displays:

- A heading describing the function.
- The data frame analyzed.
- The variables involved in the calculation of distances.
- The factor defining the populations or samples and their levels.
- The One-way MANOVA table specifying the `test` chosen for the F-test approximation, like in `summary.manova`.

In addition to the above information, the "long" output lists:

- The Between-Sample Sum of Squares and Crossed Products matrix, B
- The Within-Sample Total Sum of Squares and Crossed Products matrix, W.
- The Total Sample Sum of Squares and Crossed Products matrix, T.

Examples

```
data(skulls)
res.MANOVA <- OnewayMANOVA(skulls, group = Period)
# Long output, Wilks' test
print(res.MANOVA, test = "Wilks", long = TRUE)
```

---

print.Penrose.dist *Prints Penrose's distance matrix*

Description

Prints the results produced by `Penrose.dist`, the Penrose's distance calculator.

Usage

```
## S3 method for class 'Penrose.dist'
print(x, long = FALSE, ...)
```

Arguments

`x`	an object of class Penrose.dist
`long`	a logical variable indicating whether a long output is desired (`TRUE`) or not (`FALSE`, the default). In addition to Penrose's distances, the long output displays the covariance matrix for each group with their population / sample sizes, the mean vector for each group, and the pooled covariance matrix.
`...`	further arguments passed to or from other methods.

Value

Displays Penrose's distances between *m* multivariate populations or samples. The argument `x`, invisibly, as for all print methods, is a list of class `"Penrose.dist"`. This `print` method provides two sorts of output depending on whether the `long` argument is `TRUE` or `FALSE` (the default). The "short" output displays:

- A heading describing the function.
- The data frame analyzed.
- The variables involved in the calculation of distances.
- The factor defining the populations or samples and their levels.
- The Penrose distance matrix (lower triangular form).

In addition to the above information, the "long" output lists:

- The population or sample sizes.
- The mean vector for each population / sample.
- The covariance matrix for each population / sample
- The pooled covariance matrix.

Examples

```
data(skulls)
res.Penrose <- Penrose.dist(x = skulls, group = Period)
# Long output
print(res.Penrose, long = TRUE)
```

`print.ttests2s.mv` *Prints multiple two-sample t-tests for a multivariate data set*

Description

Prints the results produced by `ttests2s.mv`, consisting of two-sample *t*-tests on more than one response vector with corrected significance levels for multiple comparisons, as offered by `p.adjust`. Effects sizes are also displayed.

Usage

```
## S3 method for class 'ttests2s.mv'
print(x, ...)
```

Arguments

`x`	an object of class `"ttests2s.mv"`
`...`	further arguments passed to or from other methods.

Value

An annotated output of multiple two-sample *t*-tests on more than one response vector with (optionally) corrected significance levels. The argument `x`, invisibly, as for all print methods, is a list of class "`ttests2s.mv`". This `print` method provides a user-friendly display of particular elements in `x`:

- A description of the analysis.
- The data frame analyzed.
- The labels of the two-level group factor (samples), with an order determined by the user in the `ttests2s.mv` argument `level1`.
- The *t*-test results for each response variable; these include:
  - The variable name.
  - Sample means and variances classified by group levels.
  - The value of the *t*-statistic, the degrees of freedom and the *p*-value.
  - Effect sizes: raw and Hedge's (1981). The units of raw effect sizes are shown according to the argument `unit =` in `ttests2s.mv`.
- The type of alternative hypothesis for all tests.
- The method of significance level adjustment for multiple comparisons used.

References

Hedges, L. V. 1981. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics* 6(2): 107–128.

Examples

```
data(sparrows)
ttests.sparrows <- ttests2s.mv(sparrows, group = Survivorship, level1 = "S",
                               var.equal = TRUE, P.adjust = "holm",
                               unit = "mm")
print(ttests.sparrows)
```

---

print.VanValen *Prints van Valen's test*

Description

Displays the results of van Valen's test produced by the `VanValen` function and, optionally, the matrices involved in the calculations.

Usage

```
## S3 method for class 'VanValen'
print(x, long = FALSE, ...)
```

Arguments

x	an object of class VanValen.
long	a logical variable indicating whether a long output is desired (TRUE) or not (FALSE, the default)
...	further arguments passed to or from other methods.

Value

Displays the results of van Valen's test produced by the VanValen function. The argument x, invisibly, as for all print methods, is a list of class "VanValen". This print method provides two sorts of output depending on whether the long argument is TRUE or FALSE (the default). The "short" output displays:

- A two-line heading describing the analysis.
- The data frame analyzed.
- The variables used for the comparison of samples.
- The labels of the two-level group factor (samples), with an order determined by the user in the argument level1 of VanValen.
- The value of the *t*-statistic, the degrees of freedom and the *p*-value.
- The type of alternative hypothesis for the *t*-test.

In addition to the above information, the "long" output lists:

- Sub-data frames containing the standardized data, separately for each sample.
- The sample medians for the standardized data, samples 1 and 2.
- Sub-data frames containing the deviations from sample medians for the standardized values, separately for each sample.
- Sub-data frames containing the pooled distances (*d*'s), separately for each sample. These two samples of *d*-values are compared by a *t*-test.
- The means and variances for each sample of *d*-values.

Examples

```
data(sparrows)
res.VanValen <- VanValen(sparrows, "Survivorship", "S",
                         alternative = "less", var.equal = TRUE)
# Long output
print(res.VanValen, long = TRUE)
```

skulls *Egyptian male skulls*

Description

Measurements made on male skulls from the area of Thebes in Egypt. There are samples of 30 skulls from each of five periods: the Early Predynastic period (circa 4000 BC), the Late Predynastic period (circa 3300 BC), the 12th and 13th Dynasties (circa 1850 BC), the Ptolemaic period (circa 200 BC), and the Roman period (circa AD 150). Four measurements (mm) are available on each skull.

Usage

```
data(skulls)
```

Format

A data frame with 150 rows and 5 variables:

`Period` A factor with five levels

`Maximum_breadth` a numeric vector

`Basibregmatic_height` a numeric vector

`Basialveolar_length` a numeric vector

`Nasal_height` a numeric vector

References

Thomson, A. and Randall-Maciver, P. (1905). *Ancient Races of the Thebaid*, Oxford University Press, Oxford, London.

Manly, B.F.J., Navarro Alberto, J.A. and Gerow, K. (2024) *Multivariate Statistical Methods. A Primer*. 5th Edition. Boca Raton, CRC Press.

Examples

```
data(skulls)
str(skulls)
```

sparrows *Body measurements of female sparrows*

Description

Data extracted from the classical report by Hermon Bumpus (1898) who measured morphological variables in sparrows, after a severe storm. This data subset consists of five body measurements of 49 female sparrows, classified according to their survival status (21 survived, 28 did not survive).

Usage

```
data(sparrows)
```

Format

A data frame with 49 rows and 6 variables:

`Survivorship` A factor with two levels ("S" = Survived, "NS" = Did not survive)

`Total_length` Total length (mm), a numeric vector

`Alar_extent` Alar extent (mm), a numeric vector

`L_beak_head` Length of beak and head (mm), a numeric vector

`L_humerus` Length of humerus (mm), a numeric vector

`L_keel_sternum` Length of keel of sternum (mm), a numeric vector

References

Bumpus, H.C. (1898). The elimination of the unfit as illustrated by the introduced sparrow, *Passer domesticus*. *Biological Lectures*, 11th Lecture. Marine Biology Laboratory, Woods Hole, MA, 209–26.

Manly, B.F.J., Navarro Alberto, J.A. and Gerow, K. (2024) *Multivariate Statistical Methods. A Primer*. 5th Edition. Boca Raton, CRC Press.

Examples

```
data(sparrows)
str(sparrows)
```

`ttests2s.mv` *Multiple two-sample t-tests for multivariate data*

Description

Performs multiple two-sample t-tests on more than one response vector with corrected significance levels using any of the adjustment methods for multiple comparisons offered by `p.adjust`. Effects sizes are also computed.

Usage

```
ttests2s.mv(
  x,
  group,
  level1,
  alternative = "two.sided",
  var.equal = FALSE,
  P.adjust = "none",
  unit = "units"
)
```

Arguments

`x`	A data frame with one two-level factor and *p* response variables.
`group`	Two-level factor defining groups. It must be one of the columns in `x`.
`level1`	A character string identifying Sample 1. The string must be one of the factor levels in `group`.
`alternative`	a character string specifying the alternative hypothesis, must be one of `"two.sided"` (default), `"greater"` or `"less"`. You can specify just the initial letter.
`var.equal`	a logical variable indicating whether to treat the two variances as being equal. If `TRUE` then the pooled variance is used to estimate the variance otherwise the Welch (or Satterthwaite) approximation to the degrees of freedom is used.
`P.adjust`	p-value correction method, a character string. Can be abbreviated.
`unit`	A character string in cases in which all response variables are measured using the same physical units. Useful to fully characterize raw effect sizes. The default value is the character string `"units"`.

Details

This function extends the univariate `t.test` for the comparison of mean values for two samples, when more than one variable is involved in the data analysis, so that type one error rates ("false significances") in a series of univariate t-tests are adjusted according to the number of response variables analyzed. The pairwise comparisons between the two levels in `group` with corrections for multiple testing are made over more than one response vector thus, the function is a variation of `pairwise.t.test`.

The methods implemented are the same as those contained in the p.adjust.methods for p.adjust: "bonferroni", "holm", "hochberg", "hommel", "BH" (Benjamini-Hochberg) or its alias "fdr" (False Discovery Rate), and "BY" (Benjamini & Yekutieli). The default pass-through option ("none") is also included.

Value

Returns an object of class "ttests2s.mv", a list containing the following components:

name	A character string describing the function
t.list	A list containing *p* vectors of length 5, each vector having the computed t-statistic, the degrees of freedom fo
alternative	A character string specifying the alternative hypothesis chosen.
var.equal	A logical variable indicating whether the two variances were treated as being equal TRUE or not FALSE.
P.adjust	A character string indicating the correction method chosen
raw.ES	The raw effect size (scalar) expressed in the pre-specified units
unit	A character string indicating the units chosen
Hedges.d	The post hoc effect size Hedges' estimator (scalar)
group	A character string specifying the name of the two-level factor defining groups.
levels.group	A vector of length two showing the two levels in factor group.
data.name	A character string giving the name of the data.
data	the data frame analyzed.

The extractor function print.ttests2s.mv returns an annotated output of each t-test and effect size estimation.

Author(s)

Jorge Navarro Alberto, <ganava4@gmail.com>

References

Hedges, L. V. 1981. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics* 6(2): 107–128.

Examples

```
data(sparrows)
ttests.sparrows <- ttests2s.mv(sparrows, group = Survivorship, level1 = "S",
                               var.equal = TRUE, P.adjust = "bonferroni",
                               unit = "mm")
ttests.sparrows
```

VanValen *van Valen's test*

Description

Computes van Valen's test for the comparison of the variation in two multivariate samples. The comparison is made in terms of distances between all standardized variables from their corresponding standardized medians, thus producing two sets of pooled distances, one per sample, whose means are then compared by a two-sample t-test.

Usage

```
VanValen(x, group, level1, alternative = "two.sided", var.equal = FALSE)
```

Arguments

x	a data frame with one two-level factor and *p* response variables.
group	two-level factor defining groups. It must be one of the columns in x.
level1	a character string identifying Sample 1. The string must be one of the factor levels in group.
alternative	a character string specifying the alternative hypothesis in the t-test for the comparison of mean pooled distances. Must be one of "two.sided" (default), "greater" or "less". You can specify just the initial letter.
var.equal	a logical variable indicating whether to treat the two variances of pooled distances as being equal. If TRUE then the pooled variance is used to estimate the variance; otherwise the Welch (or Satterthwaite) approximation to the degrees of freedom is used.

Details

To ensure that all variables are given equal weight, each variable is first standardized in van Valen's test, so that the mean is zero and variance is one for all samples combined before the calculation of the pooled distances. These are given by

$$d_{ij} = \sqrt{\sum_{k=1}^{p} (x_{ijk} - M_{jk})^2}$$

where

x_{ijk} is the value of the standardized variable X_k for the ith individual in sample j, and

M_{jk} is the median of the same standardized variable in the jth sample.

The sample means of the d_{ij} values are compared with a t-test. If one sample is more variable than another, then the mean d_{ij} values will tend to be higher in that sample. The expression for d_{ij} in van Valen's is based on an implicit assumption that if the two samples being tested differ, then one sample will be more variable than the other for all variables. A significant result cannot be expected

in a case where, for example, X_1 and X_2 are more variable in sample 1, but X_3 and X_4 are more variable in sample 2. The effect of the differing variances would then tend to cancel out in the calculation of d_{ij}. Thus, Van Valen's test is not appropriate for situations where changes in the level of variation are not expected to be consistent for all variables.

Value

Returns an object of class "VanValen", a list containing the following components:

name	A character string describing the function.
std.data	A list with two data frames matlevel1 and matlevel2 containing the values of the standardized variables fo
medians.std	A list containing two vectors. The first vector medians.std1 contains the medians for all standardized varia
dev.median	A list with two data frames dev.median1 and dev.median2 containing the deviations from sample medians
d.list	A list with two data frames d.level1 and d.level2 containing the pooled distances of standardized variabl
means.d	A named numeric vector carrying the mean pooled distances for samples 1 and 2, respectively
vars.d	A named numeric vector carrying the variance of pooled distances for samples 1 and 2, respectively
t.vec	A named numeric vector containing the t-statistic, the degrees of freedom and the p-value for the test, respec
alternative	a character string specifying the alternative hypothesis chosen.
var.equal	A logical variable indicating whether the two variances were treated as being equal TRUE or not FALSE.
group	A character string specifying the name of the two-level factor defining groups.
levels.group	A vector of length two, showing the two levels in factor group.
data.name	A character string giving the name of the data.
variables	A character string vector containing the variable names.
data	The data frame analyzed.

Author(s)

Jorge Navarro Alberto, <ganava4@gmail.com>

References

Manly, B.F.J., Navarro Alberto, J.A. and Gerow, K. (2024) *Multivariate Statistical Methods. A Primer*. 5th Edn. CRC Press.

van Valen, L. (1978) The statistics of variation. *Evolutionary Theory* 4: 33-43. (Erratum *Evolutionary Theory* 4: 202.)

Examples

```
data(sparrows)
res.VanValen <- VanValen(sparrows, "Survivorship", "S",
                         alternative = "less", var.equal = TRUE)
# Brief output
res.VanValen
```

Index

∗ **datasets**
skulls, 25
sparrows, 26

anova.lm, *10*, *19*

BoxM.F, 2

Hotelling.mat, 4, *15*
hotelling.test, *4*, *6*

LeveneT2, 5, *16*
leveneTest, *9*
Levenetests2s.mv, 7, *17*
Levenetestsms.mv, 9, *18*

oneway.test, *9, 10*, *19*
OnewayMANOVA, 11, *20*

p.adjust, *7*, *9*
pairwise.t.test, *27*
Penrose.dist, 12
print.BoxM.F, 14
print.Hotelling.mat, 15
print.LeveneT2, *6*, 16
print.Levenetests2s.mv, *8*, 17
print.Levenetestsms.mv, *10*, 18
print.OnewayMANOVA, 20
print.Penrose.dist, 21
print.ttests2s.mv, 22, *28*
print.VanValen, 23

skulls, 25
sparrows, 26
summary.manova, *19*

ttests2s.mv, *22*, 27

VanValen, 29