Package 'Analitica'

June 14, 2025

Title Exploratory Data Analysis and Group Comparison Tools

Version 1.8.1

Description Provides tools for descriptive statistics, graphical exploration,
outlier detection, homoscedasticity tests, and multiple comparison procedures.
Includes manual implementations of Levene's test, Bartlett's test, Fligner-Killeen,
and post hoc comparisons such as Tukey, Scheffe, Games-Howell, Brunner-Munzel, and others.
Useful for teaching, applied analysis, and reproducible research.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

Depends R (>= 4.4)

Imports ggplot2, dplyr, tidyr, ggridges, patchwork, moments, magrittr,
rlang, tidyselect, multcompView,

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

NeedsCompilation no

Author Carlos Jiménez-Gallardo [aut, cre]

Maintainer Carlos Jiménez-Gallardo <carlos.jimenez@ufrontera.cl>

Repository CRAN

Date/Publication 2025-06-14 09:00:10 UTC

Contents

Analitica   *Analitica: Tools for Exploratory Data Analysis and Group Comparisons*

Description

The **Analitica** package provides tools for exploratory statistical analysis, data visualization, and comparison of groups using both parametric and non-parametric methods. It supports univariate and grouped descriptive summaries, outlier detection, homoscedasticity testing, and multiple post hoc procedures.

Details

Designed for applied analysis workflows, this package includes intuitive plotting functions and manual implementations of key statistical tests often needed in educational or research contexts.

Main Features

- `descripYG`: Descriptive statistics with visualizations (histograms, boxplots, density ridges).
- `Levene.Test`: Manual implementation of Levene's test for homogeneity of variances.
- `BartlettTest`: Manual implementation of Bartlett's test.
- `FKTest`: Manual implementation of the Fligner-Killeen test.
- `grubbs_outliers`: Outlier detection based on Grubbs' test.
- `GHTest`, `DuncanTest`, `SNKTest`, etc.: Post hoc comparison procedures.

Author(s)

Carlos Jiménez-Gallardo

BartlettTest   *Bartlett's Test for Homogeneity of Variances (Manual Implementation)*

Description

Conducts Bartlett's test to evaluate whether multiple groups have equal variances, based on a formula interface and raw data vectors, without requiring a fitted model. This implementation provides flexibility for exploratory variance testing in custom workflows.

Usage

```
BartlettTest(formula, data, alpha = 0.05)
```

Arguments

formula	A formula of the form y ~ group, where y is a numeric response and group is a factor indicating group membership.
data	A data frame containing the variables specified in the formula.
alpha	Significance level for the test (default is 0.05).

Details

Bartlett's test is appropriate when group distributions are approximately normal. It tests the null hypothesis that all groups have equal variances (homoscedasticity).

Advantages: - Straightforward to compute. - High sensitivity to variance differences under normality.

Disadvantages: - Highly sensitive to non-normal distributions. - Less robust than alternatives like Levene's test for skewed or heavy-tailed data.

Value

An object of class ″homocedasticidad″, containing:

- Statistic: Bartlett's chi-squared test statistic.
- df: Degrees of freedom associated with the test.
- p_value: The p-value for the test statistic.
- Decision: A character string indicating the conclusion ("Heterocedastic" or "Homocedastic").
- Method: A character string indicating the method used ("Bartlett").

References

Bartlett, M. S. (1937). "Properties of sufficiency and statistical tests." *Proceedings of the Royal Society of London*, Series A, 160(901), 268–282.

Examples

```
data(d_e, package = "Analitica")
res <- BartlettTest(Sueldo_actual ~ labor, data = d_e)
summary(res)

summary(BartlettTest(Sueldo_actual ~ as.factor(labor), data = d_e))
```

---

bar_error   *Bar Plot with Error Bars (Standard Deviation or Standard Error)*

Description

Creates a bar plot of group means with error bars representing either the standard deviation (SD) or the standard error (SE).

Usage

```
bar_error(
  dataSet,
  vD,
  vI,
  variation = ″sd″,
  title = ″Bar plot with error bars″,
  label_y = ″Y Axis″,
  label_x = ″X Axis″
)
```

Arguments

dataSet	A data.frame or tibble containing the data.
vD	A string indicating the name of the numeric dependent variable.
vI	A string indicating the name of the categorical independent variable (grouping variable).
variation	Type of variation to display: "sd" for standard deviation or "se" for standard error. Default is "sd".
title	Title of the plot. Default is "Bar plot with error bars".
label_y	Label for the Y-axis. Default is "Y Axis".
label_x	Label for the X-axis. Default is "X Axis".

Value

A ggplot object representing the plot.

Examples

```
data(d_e, package = "Analitica")
bar_error(d_e, vD = Sueldo_actual, vI = labor, variation = "sd")
```

BMpTest	*Brunner-Munzel Test (Permutation Version) for Two Independent Groups*

Description

Performs the Brunner-Munzel test using a permutation approach, suitable for comparing two independent samples when the assumption of equal variances may not hold.

Usage

```
BMpTest(
  grupo1,
  grupo2,
  alpha = 0.05,
  alternative = c("two.sided", "less", "greater"),
  nperm = 10000,
  seed = NULL
)
```

Arguments

grupo1	A numeric vector representing the first group.
grupo2	A numeric vector representing the second group.
alpha	Significance level (default is 0.05).
alternative	Character string specifying the alternative hypothesis: one of "two.sided" (default), "greater", or "less".
nperm	Number of permutations to perform (default = 10000).
seed	Optional random seed for reproducibility (default is NULL).

Details

This version computes an empirical p-value based on resampling, without relying on the t-distribution approximation.

Value

An object of class "comparacion" and "brunnermunzel_perm", containing:

- Resultados: A data frame with comparison name, mean difference, empirical p-value, and significance.

- Promedios: A named numeric vector of group means.

- Orden_Medias: Group names ordered by their mean.

- Metodo: Description of the method used.

References

Brunner, E., & Munzel, U. (2000). "The nonparametric Behrens-Fisher problem: Asymptotic theory and a small-sample approximation." *Biometrical Journal*, 42(1), 17–25.

Examples

```
data(d_e, package = "Analitica")
g1 <- d_e$Sueldo_actual[d_e$labor == 1]
g2 <- d_e$Sueldo_actual[d_e$labor == 2]
resultado <- BMpTest(g1, g2)
summary(resultado)
```

BMTest	*Brunner-Munzel Test for Two Independent Samples*

Description

Performs the Brunner-Munzel nonparametric test for two independent groups, which estimates the probability that a randomly selected value from one group is less than a randomly selected value from the other group.

Usage

```
BMTest(
  grupo1,
  grupo2,
  alpha = 0.05,
  alternative = c("two.sided", "less", "greater")
)
```

Arguments

grupo1	Numeric vector of values from group 1.
grupo2	Numeric vector of values from group 2.
alpha	Significance level (default = 0.05).
alternative	Character string specifying the alternative hypothesis. One of "two.sided" (default), "greater", or "less".

Details

This test is suitable when group variances are unequal and/or sample sizes differ. It does not assume equal variances and is often used as a more robust alternative to the Wilcoxon test.

Advantages: - Handles unequal variances and non-normality. - Recommended when variance homogeneity is questionable.

Disadvantages: - Less well-known and supported. - Requires large sample sizes for accurate inference.

Value

An object of class "comparacion" and "brunnermunzel", containing:

- Resultados: A data frame with test statistics, p-value, and estimated effect size.
- Promedios: A named numeric vector of group means.
- Orden_Medias: Group names ordered by their mean values (descending).
- Metodo: A character string describing the test and hypothesis.
- p_hat: Estimated probability that a value from grupo1 is less than a value from grupo2 (plus 0.5 * ties).

References

Brunner, E., & Munzel, U. (2000). "The nonparametric Behrens-Fisher problem: Asymptotic theory and a small-sample approximation." *Biometrical Journal*, 42(1), 17–25. <https://doi.org/10.1002/(SICI)1521-4036(200001)42:1

Examples

```
data(d_e, package = "Analitica")
g1 <- d_e$Sueldo_actual[d_e$labor == 1]
g2 <- d_e$Sueldo_actual[d_e$labor == 2]
resultado <- BMTest(g1, g2, alternative = "greater")
summary(resultado)
```

BonferroniNPTest   *Bonferroni-Corrected Mann-Whitney Tests (Non-Parametric)*

Description

Performs all pairwise comparisons using the Wilcoxon rank-sum test (Mann-Whitney) with Bonferroni correction for multiple testing.

Usage

```
BonferroniNPTest(formula, data, alpha = 0.05)
```

Arguments

formula	A formula of the form y ~ group.
data	A data frame containing the variables.
alpha	Significance level (default is 0.05).

Details

Suitable for non-parametric data where ANOVA assumptions are violated.

Advantages: - Simple and intuitive non-parametric alternative to ANOVA post hoc tests. - Strong control of Type I error via Bonferroni correction. - Works with unequal group sizes.

Disadvantages: - Conservative with many groups. - Only valid for pairwise comparisons; does not support complex contrasts.

Value

An object of class `"bonferroni_np"` and `"comparaciones"`, containing:

- `Resultados`: Data frame with comparisons, W-statistics, raw and adjusted p-values, and significance levels.
- `Promedios`: Mean ranks of each group.
- `Orden_Medias`: Group names ordered from highest to lowest rank.
- `Metodo`: Name of the method used ("Bonferroni (non-parametric)").

References

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6), 80–83. doi:10.2307/3001968

Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. *Technometrics*, 6(3), 241–252. doi:10.1080/00401706.1964.10490181

Shaffer, J. P. (1995). Multiple Hypothesis Testing. *Annual Review of Psychology*, 46(1), 561–584. doi:10.1146/annurev.ps.46.020195.003021

Examples

```
data(iris)
BonferroniNPTest(Sepal.Length ~ Species, data = iris)
```

BonferroniTest	*Bonferroni-Corrected Pairwise t-Tests*

Description

Performs pairwise t-tests with Bonferroni adjustment for multiple comparisons. This method controls the family-wise error rate by dividing the alpha level by the number of comparisons.

Usage

```
BonferroniTest(modelo, alpha = 0.05)
```

Arguments

modelo	An object of class aov or lm.
alpha	Significance level (default is 0.05).

Details

Advantages: - Very simple and easy to implement. - Strong control of Type I error. - Applicable to any set of independent comparisons.

Disadvantages: - Highly conservative, especially with many groups. - Can lead to low statistical power (increased Type II error). - Does not adjust test statistics, only p-values.

Value

An object of class ″bonferroni″ and ″comparaciones″, containing:

- Resultados: Data frame with comparisons, mean differences, t-values, unadjusted and adjusted p-values, and significance.
- Promedios: Named numeric vector of group means.
- Orden_Medias: Group names ordered from highest to lowest mean.
- Metodo: Name of the method used ("Bonferroni-adjusted t-test").

References

Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. *Technometrics*, 6(3), 241–252. doi:10.1080/00401706.1964.10490181

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6), 80–83. doi:10.2307/3001968

Examples

```
data(d_e, package = ″Analitica″)
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- BonferroniTest(mod)
summary(resultado)
```

BrownForsytheTest   *Brown-Forsythe Test for Homogeneity of Variances (Manual Implementation)*

Description

Performs the Brown-Forsythe test using absolute deviations from the median of each group, followed by a one-way ANOVA on those deviations.

Usage

```
BrownForsytheTest(formula, data, alpha = 0.05)
```

Arguments

formula	A formula of the form y ~ group, where y is numeric and group is a factor.
data	A data frame containing the variables.
alpha	Significance level (default is 0.05).

Details

This test is a robust alternative to Bartlett's test, especially useful when the assumption of normality is violated or when outliers are present.

Advantages: - More robust than Bartlett's test under non-normal distributions. - Less sensitive to outliers due to the use of the median.

Disadvantages: - Lower power than Bartlett's test when normality strictly holds. - Assumes that absolute deviations follow similar distributions across groups.

Value

An object of class ″homocedasticidad″, with:

- Statistic: F-statistic.
- df1: Numerator degrees of freedom.
- df2: Denominator degrees of freedom.
- p_value: P-value.
- Decision: "Heterocedastic" or "Homocedastic".
- Method: "Brown-Forsythe".

References

Brown, M. B., & Forsythe, A. B. (1974). "Robust Tests for the Equality of Variances". *Journal of the American Statistical Association*, 69(346), 364–367.

Examples

```
data(d_e, package = "Analitica")
res <- BrownForsytheTest(Sueldo_actual ~ labor, data = d_e)
summary(res)
```

ConoverTest   *Conover-Iman Test for Multiple Comparisons (Non-Parametric)*

Description

Performs non-parametric pairwise comparisons based on rank-transformed data using the Conover-Iman procedure. This method is typically applied as a post hoc test following a significant Kruskal-Wallis test to identify specific group differences.

Usage

```
ConoverTest(formula, data, alpha = 0.05, method.p = "holm")
```

Arguments

formula	A formula of the form y ~ group, where y is a numeric variable and group is a factor indicating group membership.
data	A data frame containing the variables specified in the formula.
alpha	Significance level for hypothesis testing (default is 0.05).
method.p	Method used to adjust p-values for multiple comparisons (default is "holm").

Details

The Conover-Iman test uses rank-based t-statistics, offering improved statistical power over Dunn's test while maintaining flexibility in sample size.

Advantages: - More powerful than Dunn's test, especially with moderate group differences. - Robust to non-normal data and suitable for ordinal or skewed distributions. - Allows for unequal sample sizes across groups.

Disadvantages: - Sensitive to heteroscedasticity (non-constant variances). - Requires appropriate p-value adjustment to control the family-wise error rate.

Value

An object of class "conover" and "comparaciones", containing:

- Resultados: A data frame with pairwise comparisons, t-statistics, raw and adjusted p-values, and significance markers.
- Promedios: A named numeric vector with mean ranks for each group.
- Orden_Medias: A character vector with group names sorted from highest to lowest rank.
- Metodo: A string describing the method used ("Conover (no parametrico)").

References

Conover, W. J. & Iman, R. L. (1979). "Multiple comparisons using rank sums." *Technometrics*, 21(4), 489–495.

Examples

```
data(d_e, package = "Analitica")
ConoverTest(Sueldo_actual ~ labor, data = d_e)
```

---

descripYG   *Descriptive Analysis With Optional Grouping*

Description

Performs a descriptive analysis on a numeric dependent variable, either globally or grouped by an independent variable. Displays summary statistics such as mean, standard deviation, skewness, and kurtosis, and generates associated plots (histogram, boxplot, or density ridges).

Usage

```
descripYG(dataset, vd, vi = NULL)
```

Arguments

dataset   A data.frame or tibble containing the variables.

vd   A numeric variable to analyze (dependent variable).

vi   An optional grouping variable (independent variable, categorical or numeric).

Value

A data.frame with descriptive statistics. Also prints plots to the graphics device.

Examples

```
data(d_e, package = "Analitica")
descripYG(d_e, vd = Sueldo_actual)
descripYG(d_e, vd = Sueldo_actual, vi = labor)
descripYG(d_e,Sueldo_actual,labor)
```

DSCFTest   *Dwass-Steel-Critchlow-Fligner (DSCF) Test (Non-Parametric)*

Description

Robust non-parametric method for multiple comparisons after Kruskal-Wallis. Uses rank-based pairwise tests with a pooled variance estimate.

Usage

```
DSCFTest(formula, data, alpha = 0.05, method.p = "holm")
```

Arguments

`formula`	A formula of the form `y ~ group`.
`data`	A data frame containing the variables.
`alpha`	Significance level (default is 0.05).
`method.p`	Method for p-value adjustment (default is "holm").

Details

Advantages: - Strong control of Type I error with unequal sample sizes. - More powerful than Dunn in many conditions.

Disadvantages: - Computationally more complex. - Less commonly available in standard software.

Value

An object of class `"dscf"` and `"comparaciones"`, including:

- `Resultados`: Data frame with comparisons, z-statistics, p-values, adjusted p-values, and significance levels.
- `Promedios`: Mean ranks of each group.
- `Orden_Medias`: Group names ordered from highest to lowest mean rank.
- `Metodo`: "DSCF (no paramétrico)".

References

Dwass, M. (1960). Some k-sample rank-order tests. In I. Olkin et al. (Eds.), Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling (pp. 198–202). Stanford University Press.

Examples

```
data(d_e, package = "Analitica")
DSCFTest(Sueldo_actual ~ labor, data = d_e)
```

DuncanTest	*Duncan Multiple Range Test (DMRT)*

Description

Performs the Duncan test for pairwise comparisons after an ANOVA. This method is more liberal than Tukey's HSD, using a stepwise approach with critical values from the studentized range distribution.

Usage

```
DuncanTest(modelo, alpha = 0.05)
```

Arguments

modelo	An object of class aov or lm.
alpha	Significance level (default is 0.05).

Details

Advantages: - High power for detecting differences. - Simple to interpret and implement.

Disadvantages: - Inflates Type I error rate. - Not recommended for confirmatory research.

Value

An object of class "duncan" and "comparaciones", containing:

- Resultados: A data frame with pairwise comparisons, mean differences, critical values, p-values, and significance indicators.
- Promedios: A named numeric vector of group means.
- Orden_Medias: A character vector with group names ordered from highest to lowest mean.
- Metodo: A character string indicating the comparison method ("Duncan").

References

Duncan, D. B. (1955). "Multiple range and multiple F tests." Biometrics, 11(1), 1-42.

Examples

```
data(d_e, package = "Analitica")
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- DuncanTest(mod)
summary(resultado)
plot(resultado)
```

DunnTest	*Dunn's Test for Multiple Comparisons (Non-Parametric)*

Description

Performs Dunn's test for pairwise comparisons following a Kruskal-Wallis test. Suitable for non-parametric data (ordinal or non-normal), using rank sums. Includes Holm correction by default for multiple comparisons.

Usage

```
DunnTest(formula, data, alpha = 0.05, method.p = "holm")
```

Arguments

`formula`	A formula of the form y ~ group.
`data`	A data frame containing the variables.
`alpha`	Significance level (default is 0.05).
`method.p`	Method for p-value adjustment (default is "holm").

Details

Advantages: - Simple and widely used non-parametric alternative to Tukey's test. - Handles unequal sample sizes. - Compatible with various p-value corrections (e.g., Holm, Bonferroni).

Disadvantages: - Less powerful than DSCF or Conover when sample sizes vary widely. - Requires ranking all data and can be conservative depending on adjustment.

Value

An object of class `"dunn"` and `"comparaciones"`, including:

- `Resultados`: Data frame with group comparisons, z-values, raw and adjusted p-values, and significance.
- `Promedios`: Mean ranks of each group.
- `Orden_Medias`: Group names ordered from highest to lowest rank.
- `Metodo`: "Dunn (no paramétrico)".

References

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, 6(3), 241–252. doi:10.1080/00401706.1964.10490181

See Also

`kruskal.test`, `dunn.test`

Examples

```
data(d_e, package = "Analitica")
DunnTest(Sueldo_actual ~ labor, data = d_e)
```

---

d_e	*Datos de clientes ficticios*

Description

Conjunto de dato, para ser utilizados como ejemplo. Las variables son:

Usage

```
data(d_e)
```

Format

Un data.frame con N filas y M columnas. Las variables típicas pueden incluir:

id ID del empleado

Sexo Sexo del empleado

FechaNac Fecha Nacimiento

educacion cantidad de años de estudio

labor area de trabajo dentro de la emrpesa

Sueldo_actual sueldo a la fecha

Sueldo_inicial sueldo al ingresar a la empresa

antiguedad meses trabajando en la empresa

experiencia meses de experiencia

ingreso Ingreso mensual estimado

minoria la pertenencia a una minoria

FKTest	*Fligner-Killeen Test for Homogeneity of Variances (Manual Implementation)*

Description

Performs a non-parametric Fligner-Killeen test for equality of variances across two or more groups, using raw vectors via a formula interface.

Usage

```
FKTest(formula, data, alpha = 0.05)
```

Arguments

`formula`	A formula of the form y ~ group, where y is numeric and `group` is a grouping variable (factor or coercible to factor).
`data`	A data frame containing the variables in the formula.
`alpha`	Significance level (default is 0.05).

Details

This test is particularly useful when the assumption of normality is violated, as it is robust to outliers and distributional deviations. It serves as a reliable alternative to Bartlett's test when data do not follow a normal distribution.

Advantages: - Non-parametric: No assumption of normality. - Robust to outliers. - Suitable for heterogeneous sample sizes.

Disadvantages: - Less powerful than parametric tests under normality. - May be computationally intensive with large datasets.

Value

An object of class `"homocedasticidad"`, containing:

Statistic The Fligner-Killeen chi-squared statistic.

df Degrees of freedom.

p_value The p-value for the test.

Decision `"Homoscedastic"` or `"Heteroscedastic"` depending on the test result.

Method A string indicating the method used ("Fligner-Killeen").

References

Fligner, M. A., & Killeen, T. J. (1976). "Distribution-free two-sample tests for scale." *Journal of the American Statistical Association*, 71(353), 210–213. <https://doi.org/10.1080/01621459.1976.10480351>

Examples

```
data(d_e, package = "Analitica")
res <- FKTest(Sueldo_actual ~ labor, data = d_e)
summary(res)
```

GabrielTest

Gabriel's Post Hoc Test for Multiple Comparisons

Description

A modification of Tukey's test for use with moderately unequal sample sizes.

Usage

```
GabrielTest(modelo, alpha = 0.05)
```

Arguments

modelo	An object of class aov or lm.
alpha	Significance level (default is 0.05).

Details

Advantages: - More powerful than Tukey for unequal group sizes. - Controls error rates effectively with moderate imbalance.

Disadvantages: - Can be anti-conservative with large differences in group sizes. - Less common in standard statistical software.

Value

An object of class "gabriel" and "comparaciones", containing:

- Resultados: Data frame with comparisons, mean differences, adjusted critical value, p-value, and significance level.
- Promedios: Named numeric vector of group means.
- Orden_Medias: Vector of group names ordered from highest to lowest mean.
- Metodo: Name of the method used ("Gabriel").

References

Hochberg, Y., & Tamhane, A. C. (1987). Multiple Comparison Procedures.

Examples

```
data(d_e, package = "Analitica")
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- GabrielTest(mod)
summary(resultado)
plot(resultado)
```

GHTest	*Games-Howell Post Hoc Test*

Description

Performs the Games-Howell test for pairwise comparisons after ANOVA, without assuming equal variances or sample sizes. It is suitable when Levene or Bartlett test indicates heterogeneity of variances.

Usage

```
GHTest(modelo, alpha = 0.05)
```

Arguments

modelo	An object from aov or lm.
alpha	Significance level (default is 0.05).

Details

Advantages: - Excellent for heteroscedastic data. - Controls Type I error across unequal group sizes.

Disadvantages: - Slightly conservative in small samples. - More complex to compute than Tukey.

Value

An object of class "gameshowell" and "comparaciones", which contains:

- Resultados: A data frame with pairwise comparisons, including mean differences, t-values, degrees of freedom, p-values, and significance labels.
- Promedios: A named numeric vector of group means.
- Orden_Medias: A character vector with group names ordered by their means.
- Metodo: A character string indicating the method used ("Games-Howell").

References

Games, P. A., & Howell, J. F. (1976). "Pairwise Multiple Comparison Procedures with Unequal N's and/or Variances: A Monte Carlo Study". Journal of Educational Statistics, 1(2), 113–125. <https://doi.org/10.1002/j.2162-6057.1976.tb00211.x>

Examples

```
data(d_e, package = "Analitica")
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- GHTest(mod)
summary(resultado)
plot(resultado)
```

---

grubbs_outliers *Outlier Detection Using Grubbs' Test (Iterative)*

Description

Detects one or more outliers in a numeric variable using the iterative Grubbs' test, which assumes the data follow a normal distribution.

Usage

```
grubbs_outliers(dataSet, vD, alpha = 0.05)
```

Arguments

dataSet	A data.frame containing the data.
vD	Unquoted name of the numeric variable to be tested for outliers.
alpha	Significance level for the test (default is 0.05).

Details

The function applies Grubbs' test iteratively, removing the most extreme value and retesting until no further significant outliers are found. The test is valid only under the assumption of normality.

Value

A data.frame identical to the input, with an added logical column outL indicating which observations were identified as outliers (TRUE or FALSE).

References

Grubbs, F. E. (1969). "Procedures for Detecting Outlying Observations in Samples." *Technometrics*, 11(1), 1–21. doi:10.1080/00401706.1969.10490657

Examples

```
data(d_e, package = "Analitica")
d<-grubbs_outliers(d_e, Sueldo_actual)
```

HolmTest   *Holm-Adjusted Pairwise Comparisons*

Description

Performs pairwise t-tests with p-values adjusted using Holm's sequential method.

Usage

```
HolmTest(modelo, alpha = 0.05)
```

Arguments

modelo   An object of class aov or lm.

alpha   Significance level (default is 0.05).

Details

Advantages: - Controls family-wise error rate more efficiently than Bonferroni. - Easy to apply over any set of p-values.

Disadvantages: - Does not adjust test statistics, only p-values. - Slightly more conservative than false discovery rate (FDR) methods.

Value

An object of class "holm" and "comparaciones", containing:

- Resultados: Data frame of comparisons, mean differences, t-values, unadjusted and adjusted p-values, and significance codes.
- Promedios: Named numeric vector of group means.
- Orden_Medias: Character vector with group names ordered from highest to lowest mean.
- Metodo: Name of the method used ("Holm-adjusted t-test").

References

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.

Examples

```
data(d_e, package = "Analitica")
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- HolmTest(mod)
summary(resultado)
plot(resultado)
```

Levene.Test   *Levene's Test for Homogeneity of Variances (Manual Implementation)*

Description

Performs Levene's test for equality of variances across groups using a formula interface. This test evaluates the null hypothesis that the variances are equal across groups, and is commonly used as a preliminary test before ANOVA or other parametric analyses.

Usage

```
Levene.Test(formula, data, alpha = 0.05, center = "median")
```

Arguments

formula	A formula of the form y ~ group, where y is numeric and group is a factor.
data	A data frame containing the variables in the formula.
alpha	Significance level (default is 0.05).
center	Character string: use "median" (default) or "mean" as the center for deviations.

Details

Levene's test is based on an analysis of variance (ANOVA) applied to the absolute deviations from each group's center (either the mean or, more robustly, the median). It is less sensitive to departures from normality than Bartlett's test.

Advantages: - Robust to non-normality, especially when using the median. - Suitable for equal or unequal sample sizes across groups. - Widely used in practice for checking homoscedasticity.

Disadvantages: - Less powerful than parametric alternatives under strict normality.

Value

An object of class "homocedasticidad", containing:

Statistic F statistic of the Levene test.

df Degrees of freedom (between and within groups).

p_value The p-value for the test.

Decision "Homoscedastic" or "Heteroscedastic" depending on the test result.

Method A string indicating the method used ("Levene").

References

Levene, H. (1960). "Robust Tests for Equality of Variances." In Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling (pp. 278–292). Stanford University Press.

Examples

```
data(d_e, package = "Analitica")
res <- Levene.Test(Sueldo_actual ~ labor, data = d_e)
summary(res)
```

LSDTest   *Least Significant Difference (LSD) Test*

Description

Performs unadjusted pairwise t-tests following a significant ANOVA.

Usage

```
LSDTest(modelo, alpha = 0.05)
```

Arguments

modelo	An object of class aov or `lm`.
alpha	Significance level (default is 0.05).

Details

Advantages: - Very powerful when assumptions are met. - Simple and easy to interpret.

Disadvantages: - High risk of Type I error without correction. - Not recommended if many comparisons are made.

Value

An object of class "comparaciones" with LSD results.

References

Fisher, R. A. (1935). The Design of Experiments. Oliver & Boyd.

Examples

```
data(d_e, package = "Analitica")
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- LSDTest(mod)
summary(resultado)
plot(resultado)
```

MWTest *Mann-Whitney U Test (Wilcoxon Rank-Sum, Manual Implementation)*

Description

Performs the Mann-Whitney U test (Wilcoxon rank-sum) for two independent groups, using a manual implementation. Suitable when the assumptions of parametric tests (normality, homogeneity of variances) are not met.

Usage

```
MWTest(
  grupo1,
  grupo2,
  alpha = 0.05,
  alternative = c("two.sided", "less", "greater"),
  continuity = TRUE
)
```

Arguments

grupo1	Numeric vector for the first group.
grupo2	Numeric vector for the second group.
alpha	Significance level (default = 0.05).
alternative	Character string specifying the alternative hypothesis. Options are `"two.sided"` (default), `"less"`, or `"greater"`.
continuity	Logical indicating whether to apply continuity correction (default = TRUE).

Details

Advantages: - Does not assume normality. - More powerful than t-test for skewed distributions.

Disadvantages: - Only compares two groups at a time. - Sensitive to unequal variances or shapes.

This implementation allows one- or two-sided alternatives and optionally applies a continuity correction.

Value

An object of class `"comparacion"` and `"mannwhitney"`, containing:

- `Resultados`: A data frame with the comparison name, difference in means, p-value, and significance.
- `Promedios`: A named numeric vector of group means.
- `Orden_Medias`: A character vector of group names ordered from highest to lowest mean.
- `Metodo`: A string describing the test and hypothesis direction.

References

Mann, H. B., & Whitney, D. R. (1947). "On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other." *Annals of Mathematical Statistics*, 18(1), 50–60.

Examples

```
data(d_e, package = "Analitica")
g1 <- d_e$Sueldo_actual[d_e$labor == 1]
g2 <- d_e$Sueldo_actual[d_e$labor == 2]
resultado <- MWTest(g1, g2, alternative = "greater")
summary(resultado)
```

NemenyiTest   *Nemenyi Test for Multiple Comparisons (Non-Parametric)*

Description

Performs the Nemenyi test after a significant Kruskal-Wallis or Friedman test. Based on the studentized range distribution applied to mean ranks.

Usage

```
NemenyiTest(formula, data, alpha = 0.05)
```

Arguments

formula	A formula of the form y ~ group.
data	A data frame containing the variables.
alpha	Significance level (default is 0.05).

Details

Advantages: - Easy to implement for equal-sized groups. - Conservative control of family-wise error rate.

Disadvantages: - Only valid with equal group sizes. - No p-values are directly calculated (based on critical differences only).

Value

An object of class "nemenyi" and "comparaciones", including:

- Resultados: Data frame with group comparisons, rank differences, critical value, p-values, and significance codes.
- Promedios: Mean ranks of each group.
- Orden_Medias: Group names ordered from highest to lowest rank.
- Metodo: Name of the method ("Nemenyi (no paramétrico)").

References

Nemenyi, P. (1963). Distribution-free Multiple Comparisons.

Examples

```
set.seed(123)
datos <- data.frame(
 grupo = rep(c("A", "B", "C", "D"), each = 10),
 medida = c(
   rnorm(10, mean = 10),
   rnorm(10, mean = 12),
   rnorm(10, mean = 15),
   rnorm(10, mean = 11)
 )
)
table(datos$grupo)
#> A  B  C  D
#>10 10 10 10
# Aplicar el test de Nemenyi
resultado <- NemenyiTest(medida ~ grupo, data = datos)
# Ver los resultados
summary(resultado)
# O simplemente
resultado$Resultados
# Ver orden de medias (rangos)
resultado$Orden_Medias
```

---

plot.comparaciones	*Generic plot for multiple comparison tests (with multcompView let-ters)*

Description

This function generates a bar plot displaying group means along with significance letters based on multiple comparisons. It uses `multcompView` to assign letters indicating statistically different groups.

Usage

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

Arguments

x   An object of class `comparaciones`.

...   Additional arguments (currently not used).

Value

No return value. Called for side effects: displays a bar plot with significance letters.

See Also

[multcompLetters](), [ggplot]()

Examples

```
# Assuming you have an object of class 'comparaciones' named res
# plot(res)
```

ScheffeTest   *Scheffé Test for Multiple Comparisons*

Description

Performs Scheffé's post hoc test after fitting an ANOVA model. This test compares all possible pairs of group means, using a critical value based on the F-distribution.

Usage

```
ScheffeTest(modelo, alpha = 0.05)
```

Arguments

modelo   An object of class aov or lm representing an ANOVA model.

alpha   Significance level (default is 0.05).

Details

The Scheffé test is a conservative method, making it harder to detect significant differences, but reducing the likelihood of Type I errors (false positives). It is especially appropriate when the comparisons were not pre-planned and the number of contrasts is large.

Assumptions: normally distributed residuals and homogeneity of variances.

Advantages: - Very robust to violations of assumptions. - Suitable for complex comparisons, not just pairwise.

Disadvantages: - Very conservative; reduced power. - Not ideal for detecting small differences.

Value

An object of class `"scheffe"` and `"comparaciones"`, containing:

- `Resultados`: A data frame of pairwise comparisons with difference, critical value, p-value, and significance code.
- `Promedios`: A named numeric vector of group means.
- `Orden_Medias`: A character vector with group names ordered from highest to lowest mean.
- `Metodo`: A character string indicating the test name ("Scheffe").

References

Scheffé, H. (1953). "A method for judging all contrasts in the analysis of variance." *Biometrika*, 40(1/2), 87–104. <https://doi.org/10.1093/biomet/40.1-2.87>

Examples

```
data(d_e, package = "Analitica")
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- ScheffeTest(mod)
summary(resultado)
plot(resultado)
```

SNKTest   *Student-Newman-Keuls (SNK) Test for Multiple Comparisons*

Description

Performs the Student-Newman-Keuls (SNK) post hoc test for pairwise comparisons after fitting an ANOVA model. The test uses a stepwise approach where the critical value depends on the number of means spanned between groups (range r).

Usage

```
SNKTest(modelo, alpha = 0.05)
```

Arguments

modelo	An object of class aov or lm representing an ANOVA model.
alpha	Significance level (default is 0.05).

Details

SNK is more powerful but less conservative than Tukey's HSD, increasing the chance of detecting real differences while slightly raising the Type I error rate.

Assumptions: normality, homogeneity of variances, and independence of observations.

Advantages: - More powerful than Tukey when differences are large. - Intermediate control of Type I error.

Disadvantages: - Error control is not family-wise. - Type I error increases with more comparisons.

Value

An object of class "snk" and "comparaciones", containing:

- Resultados: A data frame with pairwise comparisons, including mean differences, critical values, p-values, and significance codes.
- Promedios: A named numeric vector of group means.
- Orden_Medias: A character vector with group names ordered from highest to lowest mean.
- Metodo: A character string indicating the test used ("SNK").

References

Student, Newman, and Keuls (1952). "Student-Newman-Keuls Procedure". See also: <https://doi.org/10.1002/bimj.2003100

Examples

```
data(d_e, package = "Analitica")
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- SNKTest(mod)
summary(resultado)
plot(resultado)
```

---

summary.comparacion   *Summary Method for Objects of Class 'comparacion'*

Description

Displays a formatted summary of the results from a pairwise comparison test of two independent groups. Compatible with objects returned by functions like BMTest() or MWTest().

Usage

```
## S3 method for class 'comparacion'
summary(object, ...)
```

Arguments

object   An object of class "comparacion".

...   Additional arguments (currently ignored).

Value

Invisibly returns a one-row data frame with the summary statistics.

```
summary.homocedasticidad
```
Summary Method for Homoscedasticity Test Results

Description

Displays a summary of variance homogeneity tests such as Bartlett, Fligner-Killeen, or Levene, applied to a fitted formula using numeric data and groupings.

Usage

```
## S3 method for class 'homocedasticidad'
summary(object, ...)
```

Arguments

object	An object of class "homocedasticidad".
...	Currently ignored.

Value

Invisibly returns the input object (invisible). Printed output includes: test name, statistic, degrees of freedom, p-value, and decision at the 0.05 level.

```
T2Test
```
Tamhane's T2 Post Hoc Test

Description

Performs the Tamhane T2 test for pairwise comparisons after an ANOVA model, assuming unequal variances and/or unequal sample sizes. This test is appropriate when the assumption of homogeneity of variances is violated, such as when Levene's test or Bartlett's test is significant.

Usage

```
T2Test(modelo, alpha = 0.05)
```

Arguments

modelo	An object of class aov or lm.
alpha	Significance level (default is 0.05).

Details

The test uses a modified t-test with Welch-Satterthwaite degrees of freedom and a conservative approach to control for multiple comparisons.

Advantages: - Controls Type I error under heteroscedasticity. - No assumption of equal sample sizes.

Disadvantages: - Conservative; may reduce power. - Not as powerful as Games-Howell in some contexts.

Value

An object of class `"tamhanet2"` and `"comparaciones"`, containing:

- `Resultados`: A data frame with pairwise comparisons, mean differences, t-values, degrees of freedom, p-values, and significance codes.
- `Promedios`: A named numeric vector of group means.
- `Orden_Medias`: A character vector with group names ordered from highest to lowest mean.
- `Metodo`: A character string indicating the method used ("Tamhane T2").

References

Tamhane, A. C. (1977). "Multiple comparisons in model I one-way ANOVA with unequal variances." *Communications in Statistics - Theory and Methods*, 6(1), 15–32. <https://doi.org/10.1080/03610927708827524>

Examples

```
data(d_e, package = "Analitica")
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- T2Test(mod)
summary(resultado)
plot(resultado)
```

T3Test   *Dunnett's T3 Post Hoc Test*

Description

Performs Dunnett's T3 test for pairwise comparisons after an ANOVA model. This test is recommended when group variances are unequal and sample sizes differ. It is based on the studentized range distribution and provides conservative control over Type I error without assuming homoscedasticity.

Usage

```
T3Test(modelo, alpha = 0.05)
```

Arguments

`modelo`	An object of class `aov` or `lm`.
`alpha`	Significance level (default is 0.05).

Details

Advantages: - More powerful than T2 when group sizes are small. - Adjusted for unequal variances.

Disadvantages: - Complex critical value estimation. - Less frequently used and harder to find in software.

Value

An object of class `"dunnettt3"` and `"comparaciones"`, containing:

- `Resultados`: A data frame with pairwise comparisons, mean differences, q-values, degrees of freedom, p-values, and significance indicators.
- `Promedios`: A named numeric vector of group means.
- `Orden_Medias`: A character vector of group names ordered from highest to lowest mean.
- `Metodo`: A character string with the test name ("Dunnett T3").

References

Dunnett, C. W. (1980). "Pairwise multiple comparisons in the unequal variance case." *Journal of the American Statistical Association*, 75(372), 796–800. <https://doi.org/10.1080/01621459.1980.10477558>

Examples

```
data(d_e, package = "Analitica")
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- T3Test(mod)
summary(resultado)
plot(resultado)
```

TukeyTest   *Tukey HSD Test for Multiple Comparisons*

Description

Performs Tukey's Honest Significant Difference (HSD) test for all pairwise comparisons after fitting an ANOVA model. This post hoc method uses the studentized range distribution and is appropriate when variances are equal across groups and observations are independent.

Usage

```
TukeyTest(modelo, alpha = 0.05)
```

Arguments

`modelo`	An object of class `aov` or `lm` representing an ANOVA model.
`alpha`	Significance level (default is 0.05).

Details

Tukey's test controls the family-wise error rate and is widely used when group comparisons have not been planned in advance.

Advantages: - Strong control of Type I error rate. - Ideal for balanced designs with equal variances.

Disadvantages: - Assumes equal variances and sample sizes. - Less powerful with heteroscedasticity.

Value

An object of class `"tukey"` and `"comparaciones"`, containing:

- `Resultados`: A data frame of pairwise comparisons with mean differences, critical value, p-value, and significance level.
- `Promedios`: A named numeric vector of group means.
- `Orden_Medias`: A character vector with group names ordered from highest to lowest mean.
- `Metodo`: A character string indicating the method used ("Tukey").

References

Tukey, J. W. (1949). "Comparing individual means in the analysis of variance." *Biometrics*, 5(2), 99–114. <https://doi.org/10.2307/3001913>

Examples

```
data(d_e, package = "Analitica")
mod <- aov(Sueldo_actual ~ as.factor(labor), data = d_e)
resultado <- TukeyTest(mod)
summary(resultado)
plot(resultado)
```

Index