Package ‘TestDimorph’

February 23, 2021

Type Package

Title Analysis Of The Interpopulation Difference In Degree of Sexual Dimorphism Using Summary Statistics

Version 0.4.1

Maintainer Bassam A. Abulnoor `<bassam.abulnoor@gmail.com>`

Description Offers a solution for the unavailability of raw data in most anthropological studies by facilitating the calculations of several sexual dimorphism related analyses using the published summary statistics of metric data (mean, standard deviation and sex specific sample size) as illustrated by the works of Greene, D. L. (1989) <doi:10.1002/ajpa.1330790113> and Konigsberg, L. W. (1991) <doi:10.1002/ajpa.1330840110>.

License GPL-3

Depends R (>= 2.10)

Suggests testthat (>= 2.1.0)

Encoding UTF-8

Language en-US

LazyData true

RoxygenNote 7.1.1

Imports caret, corrplot, cutpointr, dplyr, ggplot2, Morpho, multcompView, stats, tidyr,tmvtnorm ,truncnorm ,utils

NeedsCompilation no

Author Bassam A. Abulnoor [aut, cre] (<https://orcid.org/0000-0003-4351-2754>),
MennattAllah H. Attia [aut] (<https://orcid.org/0000-0002-2304-532X>),
Lyle W. Konigsberg [aut] (<https://orcid.org/0000-0003-4052-1575>)

Repository CRAN

Date/Publication 2021-02-23 14:20:03 UTC

R topics documented:

---

accu_model *Evaluation Of Sex prediction Accuracy*

Description

Testing, cross validation and visualization of the accuracy of different sex prediction models using the confusionMatrix and roc curves.

Usage

```
accu_model(
  f,
  x,
  y = NULL,
  method = "lda",
  res_method = "repeatedcv",
  p = 0.75,
  nf = 10,
  nr = 3,
  plot = FALSE,
  Sex = 1,
  Pop = NULL,
  byPop = FALSE,
  ref. = "F",
```

```
  post. = "M",
  ...
)
```

Arguments

`f`	Formula in the form ‘groups ~ x1 + x2 + ...‘. The grouping factor is placed to the left hand side while the numerical measurements are placed to the right hand side
`x`	Data frame to be fitted to the model
`y`	New data frame to be tested, if ‘NULL‘ ‘x‘ is split to test and training datasets, Default: NULL
`method`	A string specifying which classification or regression model to use. For list of supported methods see models.
`res_method`	The resampling method: "boot", "boot632", "optimism_boot", "boot_all", "cv", "repeatedcv", "LOOCV", "LGOCV" (for repeated training/test splits), "none" (only fits one model to the entire training set), timeslice, "adaptive_cv", "adaptive_boot" or "adaptive_LGOCV", Default: 'repeatedcv'
`p`	Percentage of ‘x‘ for testing the model in case ‘y‘ is NULL, Default: 0.75
`nf`	number of folds or of resampling iterations, Default: 10
`nr`	Number of repeats for repeated k fold cross validation, Default: 3
`plot`	Logical; if TRUE returns an roc curve for model accuracy, Default: FALSE
`Sex`	Number of the column containing sex 'M' for male and 'F' for female, Default: 1
`Pop`	Number of the column containing populations' names, Default: NULL
`byPop`	Logical; if TRUE returns the accuracy in different populations of the new data frame, Default: FALSE.
`ref.`	reference category in the grouping factor, Default: 'F'
`post.`	positive category in the grouping factor, Default: 'M'
`...`	additional arguments that can passed to modeling, confusionMatrix function and roc curve generated by plot_roc.

Details

Data frames to be entered as input need to be arranged in a similar manner to [Howells] dataset. The "cut point" is found such that it maximizes the sum of "sensitivity" [TP/(TP+FN)] plus "specificity" [TN/(TN+FP)] where TP is the number of males identified as males, TN is the number of females identified as females, FN is the number of males identified as females, and FP is the number of females identified as males. For methods that employ prior probabilities, they are calculated based on sampling frequencies.

Value

Visual and numerical accuracy parameters for the tested model

Examples

```
## Not run:
library(TestDimorph)
accu_model(
  Sex ~ GOL + NOL + BNL,
  x = Howells, y = Howells, plot = FALSE
)
# Using a single dataset
accu_model(
  Sex ~ GOL + NOL + BNL,
  x = Howells,
  method = "lda",
  plot = FALSE
)

## End(Not run)
```

`aov_ss` *Sex Specific One way ANOVA From Summary statistics*

Description

Calculates sex specific one way ANOVA from summary statistics.

Usage

```
aov_ss(
  x,
  Pop = 1,
  pairwise = TRUE,
  letters = FALSE,
  es_anova = "none",
  digits = 4,
  CI = 0.95
)
```

Arguments

`x`	A data frame containing summary statistics.
`Pop`	Number of the column containing populations’ names, Default: 1
`pairwise`	Logical; if TRUE runs multiple pairwise comparisons on different populations using Tukey’s post hoc test, Default: TRUE
`letters`	Logical; if TRUE returns letters for pairwise comparisons where significantly different populations are given different letters, Default: FALSE’
`es_anova`	Type of effect size either "f" for f squared,"eta" for eta squared or "none", Default:"none".
`digits`	Number of significant digits, Default: 4
`CI`	confidence interval coverage takes value from 0 to 1, Default: 0.95.

Details

Data is entered as a data frame of summary statistics where the column containing population names is chosen by position (first by default), other columns of summary data should have specific names (case sensitive) similar to baboon.parms_df

Value

Sex specific ANOVA tables and pairwise comparisons in tidy format.

Examples

```
# Comparisons of femur head diameter in four populations
library(TestDimorph)
df <- data.frame(
  Pop = c("Turkish", "Bulgarian", "Greek", "Portuguese "),
  m = c(150.00, 82.00, 36.00, 34.00),
  f = c(150.00, 58.00, 34.00, 24.00),
  M.mu = c(49.39, 48.33, 46.99, 45.20),
  F.mu = c(42.91, 42.89, 42.44, 40.90),
  M.sdev = c(3.01, 2.53, 2.47, 2.00),
  F.sdev = c(2.90, 2.84, 2.26, 2.90)
)
aov_ss(x = df)
```

`Australia` *Australia*

Description

Raw data from Joseph Birdsell's 1938 survey of Australian Aborigines The data is from two regions (B1 and B19), see Gilligan and Bulbeck (2007) for a map of the regions.

Usage

```
Australia
```

Format

A data frame with 94 rows and 9 variables:

Pop (Region) ("B1" = Southwest Australia, "B19" = Northeast Australia), see Gilligan and Bulbeck (2007)

Sex Sex coded as "F" or "M"

Weight.kg body weight in kilograms

Stature.mm Standing height in millimeters

Hum.Lgth Humeral length in millimeters

Rad.Lgth Radius length in millimeters

Fem.Lgth Femoral length in millimeters

Tib.Lgth Tibial length in millimeters

Bi.illiac Bi-illiac breadth in millimeters

References

Gilligan, I., & Bulbeck, D. (2007). Environment and morphology in Australian Aborigines: A re<U+2010>analysis of the Birdsell database. American Journal of Physical Anthropology, 134(1), 75-91.

`baboon.parms_df` *data frame format for the baboon.parms_df for multivariate analysis*

Description

A dataset containing summary statistics for low density lipoprotein (LDL) and apolipoprotein B (apo B) levels in 604 baboons measured on two different diets: a basal diet basal' and a high cholesterol, saturated fat diet (HCSF). The baboons were classified into one of three subspecies (Papio hamadryas anubis, P.h. cynocephalus, or anubis). Each animal was measured on each of the two diets.

Usage

```
baboon.parms_df
```

Format

A data frame with 12 rows and 8 variables

Trait Apolipoprotein B and LDL on two diets

Sub Type of species

M.mu Means of lipoproteins in different species for males

F.mu Means of lipoproteins in different species for females

m Male sample sizes

f Female sample sizes

M.sdev Standard deviations for males

F.sdev Standard deviations for females

Note

The baboon data collection were supported by NIH grant HL28972 and NIH contract HV53030 to the Southwest Foundation for Biomedical Research (Now: Texas Biomedical Research Institute), and funds from the Southwest Foundation for Biomedical Research

References

Konigsberg LW (1991). An historical note on the t-test for differences in sexual dimorphism between populations. American journal of physical anthropology, 84(1), 93–96.

`baboon.parms_list` *List format for the baboon.parms_df for multivariate analysis*

Description

List format for the baboon.parms_df for multivariate analysis

Usage

`baboon.parms_list`

Format

A list of 7 matrices.

R.res	pooled within group correlation matrix
M.mu	Means of lipoproteins in different species for males
F.mu	Means of lipoproteins in different species for females
m	Male sample sizes
f	Female sample sizes
M.sdev	Standard deviations for males
F.sdev	Standard deviations for females

`baboon.parms_R` *Pooled within group correlation matrix for baboon data*

Description

Pooled within group correlation matrix for baboon data

Usage

`baboon.parms_R`

Format

A 4*4 numerical matrix

`Cremains_measurements` *Measurements from calcined postcranial materials.*

Description

Part of Table 3 from Cavazzuti et al. (2019).

Usage

```
Cremains_measurements
```

Format

A data frame with 22 rows and 8 variables:

Trait Measured feature

M.mu Means of males

F.mu Means of females

m Male sample sizes

f Female sample sizes

M.sdev Standard deviations for males

F.sdev Standard deviations for females

D published value for Chakraborty and Majumder's (1982) measure of sexual dimorphism.

References

Cavazzuti, Claudio, et al. (2019) "Towards a new osteometric method for sexing ancient cremated human remains. Analysis of Late Bronze Age and Iron Age samples from Italy with gendered grave goods." PloS one 14.1: e0209423.

Chakraborty, R., & Majumder, P. P. (1982). On Bennett's measure of sex dimorphism. American journal of physical anthropology, 59(3), 295-298.

`D_index` *Dissimilarity index*

Description

Visual and statistical computation of the area of non-overlap in the trait distribution between two sex groups.

Usage

```
D_index(
  x,
  plot = FALSE,
  fill = "female",
  Trait = 1,
  B = NULL,
  CI = 0.95,
  rand = TRUE,
  digits = 4
)
```

Arguments

`x`	A data frame containing summary statistics.
`plot`	logical; if true a plot of densities for both sexes is returned, Default: FALSE
`fill`	Specify which sex's density to be filled with color in the plot; either "male" in blue color, "female" in pink color or "both", Default: 'female'
`Trait`	Number of the column containing names of measured parameters, Default: 1
`B`	number of bootstrap samples for generating confidence intervals. Higher number means greater accuracy but slower execution. If NULL bootstrap confidence intervals are not produced, Default:NULL
`CI`	confidence interval coverage takes value from 0 to 1, Default: 0.95.
`rand`	logical; if TRUE, uses random seed. If FALSE, then set.seed(42) for repeatability, Default: TRUE
`digits`	Number of significant digits, Default: 4

Details

Chakraborty and Majumder's (1982) D index. The calculations are done using Inman and Bradley's (1989) equations, and the relationship that D = 1 - OVL where OVL is the overlap coefficient described in Inman and Bradley. A parametric bootstrap was used assuming normal distributions. The method is known as the "bias-corrected percentile method" (Efron, 1981) or the "bias-corrected percentile interval" (Tibshirani, 1984)

Value

a table and a graphical representation of the selected traits and their corresponding dissimilarity indices, confidence intervals and significance tests.

References

Chakraborty, Ranajit, and Partha P. Majumder.(1982) "On Bennett's measure of sex dimorphism." American Journal of Physical Anthropology 59.3 : 295-298.

Inman, Henry F., and Edwin L. Bradley Jr.(1989) "The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities." Communications in Statistics-Theory and Methods 18.10:3851-3874.

Efron, B. (1981). Nonparametric standard errors and confidence intervals. Canadian Journal of Statistics, 9(2), 139-158.

Tibshirani, R. J. (1984). Bootstrap confidence intervals. Technical Report No. 3, Laboratory for Computational Statistics, Department of Statistics, Stanford University.

Examples

```
library(TestDimorph)
data("Cremains_measurements")
# plot and test of significance
D_index(Cremains_measurements[1, ], plot = TRUE)
## Not run:
# confidence interval with bootstrapping
D_index(Cremains_measurements[1, ], rand = FALSE, B = 1000)

## End(Not run)
```

---

extract_sum *Summary Statistics Extraction*

Description

Extract summary data needed for other functions from raw data.

Usage

```
extract_sum(x, Sex = 1, Pop = 2, firstX = 3, test = 1, run = TRUE, ...)
```

Arguments

`x`	Data frame of raw data.
`Sex`	Number of the column containing sex ’M’ for male and ’F’ for female, Default: 1
`Pop`	Number of the column containing populations’ names, Default: 2
`firstX`	Number of column containing measured parameters (First of multiple in case of multivariate analysis), Default: 3
`test`	‘1‘ for Greene t test t_greene, ‘2‘ for univariate, ‘3‘ for sex specific ANOVA aov_ss, ‘4‘ for multivariate, and ‘5‘ for van_vark, Default: 1
`run`	Logical; if TRUE runs the corresponding test after data extraction, Default:TRUE
`...`	Additional arguments that could be passed to the test of choice

Details

Raw data is entered in a wide format data frame similar to Howells data set. The first two columns contain sex ‘Sex‘ (‘M‘ for male and ‘F‘ for female) (Default: ‘1‘) and populations’ names ‘Pop‘ (Default:‘2‘). Starting from ‘firstX‘ column (Default: ‘3‘), measured parameters are entered each in a separate column.

Value

Input for other functions.

Examples

```
# for multivariate test
library(TestDimorph)
extract_sum(Howells, test = 4)
# for univariate test on a specific parameter
library(TestDimorph)
extract_sum(Howells, test = 2, firstX = 4)
```

`Hedges_g` *Hedges' g*

Description

quantifies the size of difference between sexes in measured traits.

Usage

```
Hedges_g(x, Trait = 1, CI = 0.95, B = NULL, rand = TRUE, digits = 4)
```

Arguments

`x`	A data frame containing summary statistics.
`Trait`	Number of the column containing names of measured parameters, Default: 1
`CI`	confidence interval coverage takes value from 0 to 1, Default: 0.95.
`B`	number of bootstrap samples for generating confidence intervals. Higher number means greater accuracy but slower execution. If NULL bootstrap confidence intervals are not produced, Default:NULL
`rand`	logical; if TRUE, uses random seed. If FALSE, then set.seed(42) for repeatability, Default: TRUE
`digits`	Number of significant digits, Default: 4

Details

Calculates Hedges' (1981) g and its confidence intervals using the pooled standard deviation and correcting for bias. See Goulet-Pelletier and Cousineau (2018) for details of the calculations and D_index for description of the bootstrap.

Value

a table of Hedge's g values with confidence interval for different traits.

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.

Goulet-Pelletier, J.-C., & Cousineau, D. (2018). A review of effect sizes and their confidence intervals, part I: The Cohen's d family. The Quantitative Methods for Psychology, 14(4), 242-265.

Examples

```
library(TestDimorph)
data("Cremains_measurements")
# Confidence intervals with non-central t distribution
Hedges_g(Cremains_measurements[1, ])
## Not run:
# confidence interval with bootstrapping
Hedges_g(Cremains_measurements[1, ], rand = FALSE, B = 1000)

## End(Not run)
```

`Howells` *The Howells' craniometric data*

Description

A subset of a dataset that consists of 82 craniometric measurements taken from approximately two thousands and half human crania from 28 geographically diverse populations. The full data set can be found in https://rdrr.io/github/geanes/bioanth/man/howell.html

Usage

```
Howells
```

Format

A data frame with 441 rows and 10 variables:

Sex 'M' for male and 'F' for female

Pop Populations' names

GOL Glabello occipital length

NOL Nasio occipital length

BNL Bastion nasion length

BBH Basion bregma height

XCB Maximum cranial breadth

XFB Maximum frontal breadth

ZYB Bizygomatic breadth

AUB Biauricular breadth

References

Howells WW. (1989). Skull Shapes and the Map. Craniometric Analyses in the Dispersion of Modern Homo. Papers of the Peabody Museum of Archaeology and Ethnology, vol. 79, pp. 189. Cambridge, Mass.: Peabody Museum.

Howells WW. (1995). Who's Who in Skulls. Ethnic Identification of Crania from Measurements. Papers of the Peabody Museum of Archaeology and Ethnology, vol. 82, pp. 108. Cambridge, Mass.: Peabody Museum.

Howells, W. W. (1973). Cranial Variation in Man: A Study by Multivariate Analysis of Patterns of Difference Among Recent Human Populations (Vol. 67). Cambridge, MA: Peabody Museum of Archaeology and Ethnology.

`Howells_R` *Pooled within group correlation matrix for Howells' data*

Description

Pooled within group correlation matrix for Howells' data

Usage

```
Howells_R
```

Format

A 8*8 numerical matrix

`Howells_summary` *Summary of the Howells' craniometric data*

Description

Summary statistics of the the Howells' data subset.

Usage

```
Howells_summary
```

Format

A data frame with 32 rows and 8 variables:

Trait	Measured feature
Pop	Population name
M.mu	Means of males
F.mu	Means of females
m	Male sample sizes
f	Female sample sizes
M.sdev	Standard deviations for males
F.sdev	Standard deviations for females

References

Howells

`Howells_summary_list` *List format of Howells_summary for multivariate analysis*

Description

List format of Howells_summary for multivariate analysis

Usage

```
Howells_summary_list
```

Format

A list of 7 matrices with structure similar to baboon.parms_list

---

`Howells_V` *Pooled within-group variance-covariance matrix for Howells' data*

Description

Pooled within-group variance-covariance matrix for Howells' data

Usage

```
Howells_V
```

Format

A 8*8 numerical matrix

`MI_index` *Mixture Index ("MI")*

Description

Ipina and Durand's (2010) mixture intersection (MI) measure of sexual dimorphism. This measure is an overlap coefficient where the sum of the frequency of males and the frequency of females equals 1.0. Ipina and Durand (2010) also define a normal intersection (NI) measure which is the overlap coefficient of two normal distributions, equivalent to Inman and Bradley's (1989) overlap coefficient

Usage

```
MI_index(
  x,
  plot = FALSE,
  Trait = 1,
  B = NULL,
  CI = 0.95,
  p.f = 0,
  index_type = "MI",
  rand = TRUE,
  digits = 4
)
```

Arguments

`x`	A data frame containing summary statistics.
`plot`	logical; if true a plot of densities for both sexes is returned, Default: FALSE
`Trait`	Number of the column containing names of measured parameters, Default: 1
`B`	number of bootstrap samples for generating confidence intervals. Higher number means greater accuracy but slower execution. If NULL bootstrap confidence intervals are not produced, Default:NULL
`CI`	confidence interval coverage takes value from 0 to 1, Default: 0.95.
`p.f`	proportion of sample that is female (if p.f>0 then p.m=1-p.f, where p.m is the proportion of males and bootstrap won't be available) , Default: 0
`index_type`	type of coefficient (if "MI" it fits the mixture index. If = "NI" it fits the overlap coefficient for two normal distributions, which is equal to 1 – D_index, Default: 'MI'
`rand`	logical; if TRUE, uses random seed. If FALSE, then set.seed(42) for repeatability, Default: TRUE
`digits`	Number of significant digits, Default: 4

Details

see D_index for bootstrap method.

Value

returns a table of Ipina and Durand's (2010) mixture index ("MI") for different traits with graphical representation.

References

Inman, H. F., & Bradley Jr, E. L. (1989). The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. Communications in Statistics-Theory and Methods, 18(10), 3851-3874.

Ipina, S. L., & Durand, A. I. (2010). Assessment of sexual dimorphism: a critical discussion in a (paleo-) anthropological context. Human Biology, 82(2), 199-220.

Examples

```
library(TestDimorph)
data("Cremains_measurements")
# plot and test of significance
MI_index(Cremains_measurements[1, ], plot = TRUE)
#' #NI index
MI_index(Cremains_measurements[1, ], index_type = "NI")
## Not run:
# confidence interval was bootstrapping
MI_index(Cremains_measurements[1, ], rand = FALSE, B = 1000)

## End(Not run)
```

`models` *List of accu_model supported models*

Description

List of accu_model supported models

Usage

```
models
```

Format

A list of length 20.

`multivariate` *Multivariate Analysis Of Sexual Dimorphism*

Description

Multivariate extension of Greene t test t_greene

Usage

```
multivariate(
  x,
  R.res = NULL,
  Trait = 1,
  Pop = 2,
  type_manova = "II",
  manova_test_statistic = "W",
  interact_manova = TRUE,
  es_manova = "none",
  univariate = FALSE,
  padjust = "none",
  ...,
  lower.tail = FALSE,
  CI = 0.95,
  digits = 4
)
```

Arguments

`x`	Data frame or list containing summary statistics for multiple parameters measured in both sexes in two or more populations.
`R.res`	Pooled within correlational matrix, Default: NULL
`Trait`	Number of the column containing names of measured parameters, Default: 1
`Pop`	Number of the column containing populations' names, Default: 2
`type_manova`	type of MANOVA test "I","II" or "III", Default:"II".
`manova_test_statistic`	type of test statistic used either "W" for "Wilks","P" for "Pillai", "HL" for "Hotelling-Lawley" or "R" for "Roy's largest root", Default: "W".
`interact_manova`	Logical; if TRUE calculates MANOVA for the interaction effects,Default: TRUE.
`es_manova`	effect size either ,"eta" for eta squared, or "none"for not reporting an effect size, Default:"none".
`univariate`	Logical; if TRUE conducts multiple univariate analyses on different parameters separately, Default: FALSE
`padjust`	Method of p.value adjustment for multiple comparisons following p.adjust Default: "none".

`...`	Additional arguments that could be passed to univariate
`lower.tail`	Logical; if TRUE probabilities are ‘P[X <= x]‘, otherwise, ‘P[X > x]‘., Default: FALSE
`CI`	confidence interval coverage takes value from 0 to 1, Default: 0.95.
`digits`	Number of significant digits, Default: 4

Details

Data can be entered either as a data frame of summary statistics as in baboon.parms_df. In that case the pooled within correlational matrix ‘R.res‘ should be entered as a separate argument as in baboon.parms_R. Another acceptable format is a named list of matrices containing different summary statistics as well as the correlational matrix as in baboon.parms_list. By setting the option ‘univariate‘ to ‘TRUE‘, multiple ‘ANOVA‘s can be run on each parameter independently with the required p.value correction using p.adjust.methods.

Value

MANOVA table. When the term is followed by ‘(E)‘ an exact f-value is calculated.

Examples

```
# x is a data frame with separate correlational matrix
library(TestDimorph)
multivariate(baboon.parms_df, R.res = baboon.parms_R)
# x is a list with the correlational matrix included
library(TestDimorph)
multivariate(baboon.parms_list, univariate = TRUE, padjust = "bonferroni")
# reproduces results from Konigsberg (1991)
multivariate(baboon.parms_df, R.res = baboon.parms_R)[3, ]
multivariate(baboon.parms_df, R.res = baboon.parms_R, interact_manova = FALSE)
```

`NHANES_1999` *NHANES 1999*

Description

Raw data from 1999-2000 NHANES (National Health and Nutrition Examination Survey). Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2020, `https://www.cdc.gov/nchs/nhanes/index.htm`

Usage

```
NHANES_1999
```

Format

A data frame with 1430 rows and 5 variables:

Sex (RIAGENDR) Sex coded as "F" or "M"

Pop (RIDRETH1) Self-reported race, coded as "Black" = Non-Hispanic Black, "Mex.Am" = Mexican American, or "White" = Non-Hispanic White

BMXWT Body weight in kilograms

BMXHT Standing height in centimeters

BMXARML Upper arm length in centimeters

Note

This is not the complete dataset. It is selected so that age in years is greater than or equal to 20 and less than or equal to 40

`raw_gen` *Raw Data Generation By Log-normal Or Truncated Distribution*

Description

Generates raw data from summary statistics using uni/multivariate log/truncated normal distribution

Usage

```
raw_gen(
  x,
  Trait = 1,
  Pop = 2,
  R.res = NULL,
  dist = c("truncated", "log"),
  lower = -Inf,
  upper = Inf,
  format = c("wide", "long"),
  complete_cases = FALSE
)
```

Arguments

`x`	Data frame or list containing summary statistics for multiple parameters measured in both sexes in two or more populations.
`Trait`	Number of the column containing names of measured parameters, Default: 1
`Pop`	Number of the column containing populations' names, Default: 2
`R.res`	Pooled within correlational matrix, Default: NULL
`dist`	univariate distribution used for data generation either ‘log‘ for log normal or ‘truncated‘ for truncated distribution, Default: 'truncated'

`lower`	vector of lower bounds, Default: -Inf
`upper`	vector of upper bounds, Default: Inf
`format`	form of the resultant data frame either 'long' or 'wide', Default: 'wide'
`complete_cases`	Logical; if TRUE rows with missing values will be removed, Default: FALSE

Details

If data generation is desired using multivariate distribution data is entered in the form of a list of summary statistics and pooled within correlational matrix as in baboon.parms_list, or the summary statistics are entered separately in the form of a data frame as in baboon.parms_df with a separate correlational matrix as in baboon.parms_R. If data frame is entered without a correlational matrix, data generation is carried out using univariate distribution. N.B: Transformation of raw summary data to logged data is only possible for univariate distribution and if multivariate log normal distribution is desired logged values should be entered directly with ‘dist‘ set to ‘truncated‘.

Value

a data frame of raw data

Examples

```
# Data generation using univariate distribution
library(TestDimorph)
raw_gen(baboon.parms_df)
# Data generation using multivariate distribution
library(TestDimorph)
raw_gen(baboon.parms_list)
```

`t_greene` *Greene t test of Sexual Dimorphism*

Description

Calculation and visualization of the differences in degree sexual dimorphism between two populations using summary statistics as input.

Usage

```
t_greene(
  x,
  Pop = 1,
  es = "none",
  plot = FALSE,
  ...,
  alternative = c("two.sided", "less", "greater"),
  padjust = "none",
  letters = FALSE,
```

```
  digits = 4,
  CI = 0.95
)
```

Arguments

`x`	A data frame containing summary statistics.
`Pop`	Number of the column containing populations' names, Default: 1
`es`	Type of effect size either "d" for Cohen's d,"g" for Hedge's g or "none" , Default:"none".
`plot`	Logical; if TRUE graphical matrix of p values, Default: FALSE
`...`	additional arguments that can be passed to [corrplot][corrplot::corrplot] function.
`alternative`	a character string specifying the alternative hypothesis, must be one of "two.sided", "greater" or "less".
`padjust`	Method of p.value adjustment for multiple comparisons following p.adjust Default: "none".
`letters`	Logical; if TRUE returns letters for pairwise comparisons where significantly different populations are given different letters, Default: FALSE'
`digits`	Number of significant digits, Default: 4
`CI`	confidence interval coverage takes value from 0 to 1, Default: 0.95.

Details

The input is a data frame of summary statistics where the column containing population names is chosen by position (first by default), other columns of summary data should have specific names (case sensitive) similar to baboon.parms_df

Value

data frame of t.test results

See Also

[multcompView::multcompLetters()] [corrplot::corrplot()]

Examples

```
# Comparisons of femur head diameter in four populations
library(TestDimorph)
df <- data.frame(
  Pop = c("Turkish", "Bulgarian", "Greek", "Portuguese "),
  m = c(150.00, 82.00, 36.00, 34.00),
  f = c(150.00, 58.00, 34.00, 24.00),
  M.mu = c(49.39, 48.33, 46.99, 45.20),
  F.mu = c(42.91, 42.89, 42.44, 40.90),
  M.sdev = c(3.01, 2.53, 2.47, 2.00),
  F.sdev = c(2.90, 2.84, 2.26, 2.90)
```

```
)
t_greene(
  df,
  plot = TRUE,
  method = "ellipse",
  padjust = "none",
  type = "lower",
  col = c(
    "#AEB6E5",
    "#B1A0DB",
    "#B788CD",
    "#BC6EB9",
    "#BC569E",
    "#B6407D",
    "#A93154"
  ),
  tl.cex = 0.8,
  tl.col = "black",
  insig =
    "label_sig",
  tl.srt = 0.1,
  pch.cex = 2.5,
  tl.pos = "ld",
  win.asp = 1,
  number.cex = 0.5,
  na.label = "NA"
)
```

`univariate` *Univariate Analysis Of Sexual Dimorphism*

Description

Calculation and visualization of the differences in degree sexual dimorphism between multiple populations using a modified one way ANOVA and summary statistics as input

Usage

```
univariate(
  x,
  Pop = 1,
  type_anova = "II",
  interact_anova = TRUE,
  es_anova = "none",
  pairwise = FALSE,
  padjust = "none",
  ...,
  lower.tail = FALSE,
  CI = 0.95,
```

```
  N = NULL,
  digits = 4
)
```

Arguments

`x`	A data frame containing summary statistics.
`Pop`	Number of the column containing populations' names, Default: 1
`type_anova`	type of ANOVA test "I","II" or "III", Default:"II".
`interact_anova`	Logical; if TRUE calculates interaction effect, Default: TRUE.
`es_anova`	Type of effect size either "f" for f squared,"eta" for eta squared or "none", Default:"none".
`pairwise`	Logical; if TRUE runs multiple pairwise comparisons on different populations using t_greene Default: FALSE
`padjust`	Method of p.value adjustment for multiple comparisons following p.adjust Default: "none".
`...`	Additional arguments that could be passed to the t_greene function
`lower.tail`	Logical; if TRUE probabilities are ‘P[X <= x]‘, otherwise, ‘P[X > x]‘., Default: FALSE
`CI`	confidence interval coverage takes value from 0 to 1, Default: 0.95.
`N`	Number of pairwise comparisons for [p.adjust.methods], if left ‘NULL‘ it will follow the formula ‘n(n 1)/2‘ where ‘n‘ is the number of populations , Default: NULL
`digits`	Number of significant digits, Default: 4

Details

Data is entered as a data frame of summary statistics where the column containing population names is chosen by position (first by default), other columns of summary data should have specific names (case sensitive) similar to baboon.parms_df

Value

ANOVA tale.

Examples

```
# Comparisons of femur head diameter in four populations
library(TestDimorph)
df <-
  data.frame(
    Pop = c("Turkish", "Bulgarian", "Greek", "Portuguese "),
    m = c(150.00, 82.00, 36.00, 34.00),
    M.mu = c(49.39, 48.33, 46.99, 45.20),
    M.sdev = c(3.01, 2.53, 2.47, 2.00),
    f = c(150.00, 58.00, 34.00, 24.00),
    F.mu = c(42.91, 42.89, 42.44, 40.90),
```

```
    F.sdev = c(2.90, 2.84, 2.26, 2.90)
  )
univariate(df, pairwise = TRUE, padjust = "bonferroni")
```

`van_vark` *Graphical and statistical representation of dimorphism differences.*

Description

Provides testing for differences in patterning of sexual dimorphism between populations, as well as for evolutionary trends that may characterize other species. The test is based on the computation of the first q canonical variates (q=2 by default) or multiple discriminant functions to develop various tests of sexual dimorphism in any two populations A and B.

Usage

```
van_vark(
  x,
  W,
  q = 2,
  Trait = 1,
  Pop = 2,
  plot = TRUE,
  lower.tail = FALSE,
  digits = 4
)
```

Arguments

`x`	Data frame of means and sample sizes for different populations.
`W`	Pooled within-group variance-covariance matrix
`q`	Number of canonical variates to retain for chi square test, Default: 2
`Trait`	number of column containing names of traits Default: 1.
`Pop`	Number of the column containing populations' names, Default: 2
`plot`	Logical; if TRUE returns a graphical representation of dimorphism differences, Default: TRUE
`lower.tail`	Logical; if TRUE probabilities are 'P[X <= x]', otherwise, 'P[X > x]'., Default: FALSE
`digits`	Number of significant digits, Default: 4

Details

Input is a data frame of means and sample sizes similar to Howells_summary with the same naming conventions used throughout the functions but with the standard deviation columns removed.

Value

The output includes a two-dimensional plot that illustrate the existing differences between tested populations and a statistical test of significance for the difference in dimorphism using chi square distribution.

References

van Vark, G. N., et al. (1989). van Vark, G. N., et al. "Some multivariate tests for differences in sexual dimorphism between human populations." Annals of human biology 16.4: 301-310.

Examples

```
library(TestDimorph)
# selecting means and sample sizes
van_vark_data <- Howells_summary[!endsWith(
  x = names(Howells_summary),
  suffix = "dev"
)]
# running the function
van_vark(van_vark_data, Howells_V)
```

Index