

# Package ‘TestDimorph’

October 12, 2022

**Type** Package

**Title** Analysis of the Interpopulation Difference in Degree of Sexual Dimorphism Using Summary Statistics

**Version** 0.5.5

**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 Relethford, J. H., & Hodges, D. C. (1985) <doi:10.1002/ajpa.1330660105>, Greene, D. L. (1989) <doi:10.1002/ajpa.1330790113>, berg, L. W. (1991) <doi:10.1002/ajpa.1330840110>.

**License** GPL-3

**Depends** R (>= 3.5.0)

**Suggests** testthat (>= 2.1.0), knitr, rmarkdown

**VignetteBuilder** knitr

**Encoding** UTF-8

**Language** en-US

**LazyData** true

**RoxygenNote** 7.2.1

**Imports** corrplot, dplyr, ggplot2, Morpho, multcompView, stats, tidy,tmvtnorm ,truncnorm ,utils

**Date** 2022-08-27

**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>>), Iain R. Konigsberg [aut] (<<https://orcid.org/0000-0001-7356-100X>>), Lyle W. Konigsberg [aut] (<<https://orcid.org/0000-0003-4052-1575>>)

**Repository** CRAN

**Date/Publication** 2022-08-30 07:00:20 UTC

**R topics documented:**

aov_ss . . . . .	2
Australia . . . . .	4
baboon.parms_df . . . . .	5
baboon.parms_list . . . . .	6
baboon.parms_R . . . . .	6
Cre mains_ measurements . . . . .	7
D_index . . . . .	7
extract_sum . . . . .	9
FT . . . . .	10
Hedges_g . . . . .	11
Howells . . . . .	12
Howells_R . . . . .	13
Howells_summary . . . . .	13
Howells_summary_list . . . . .	14
Howells_V . . . . .	15
MI_index . . . . .	15
multivariate . . . . .	17
NHANES_1999 . . . . .	18
raw_gen . . . . .	19
SMO . . . . .	20
t_greene . . . . .	21
univariate . . . . .	23
van_vark . . . . .	24
<b>Index</b>	<b>27</b>

---

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-Kramer'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 "f2" for f squared, "eta2" for eta squared, "omega2" for omega 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.

## References

#For the femur head diameter data

F. Curate, C. Umbelino, A. Perinha, C. Nogueira, A.M. Silva, E. Cunha, Sex determination from the femur in Portuguese populations with classical and machinelearning classifiers, *J. Forensic Leg. Med.* (2017) , doi:<http://dx.doi.org/10.1016/j.jflm.2017.08.011>.

O. Gulhan, *Skeletal Sexing Standards of Human Remains in Turkey* (PhD thesis), Cranfield University, 2017 [Dataset].

P. Timonov, A. Fasova, D. Radoinova, A. Alexandrov, D. Delev, A study of sexual dimorphism in the femur among contemporary Bulgarian population, *Euras. J. Anthropol.* 5 (2014) 46–53.

E.F. Kranioti, N. Vorniotakis, C. Galiatsou, M.Y. Iscan , M. Michalodimitrakis, Sex identification and software development using digital femoral head radiographs, *Forensic Sci. Int.* 189 (2009) 113.e1–7.

## Examples

```
# Comparisons of femur head diameter in four populations
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. The data is from two regions (B1 and B19), see Gilligan and Bulbeck (2007) for a map of the regions. Data downloaded from Dr. Peter Brown's website: <https://www.peterbrown-palaeoanthropology.net/resource.html>

### 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.iliac** Bi-iliac breadth in millimeters

### References

Gilligan, I., & Bulbeck, D. (2007). Environment and morphology in Australian Aborigines: A re-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 and a high cholesterol, saturated fat diet. The baboons were classified into one of two subspecies and a hybrid of the two subspecies (*Papio hamadryas anubis*, *P.h. cynocephalus*, or hybrid). 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** Sub-species or hybrid

**M.mu** Means of LDL and apo B in different sub-species for males

**F.mu** Means of LDL and apo B in different sub-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 5 matrices (`R.res`, `M.mu`, `F.mu`, `M.sdev`, and `F.sdev`) and two vectors (`m` and `f`)

**R.res** pooled within group correlation matrix

**M.mu** Means of LDL and apo B in different sub-species for males

**F.mu** Means of LDL and apo B in different sub-species for females

**m** Male sample sizes

**f** Female sample sizes

**M.sdev** Standard deviations for males

**F.sdev** Standard deviations for females

**See Also**

[baboon.parms\\_df](#)

---

`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

**See Also**

[baboon.parms\\_list](#)

---

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,
  verbose = FALSE,
  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
verbose	logical; if TRUE number of bootstraps is displayed, Default: FALSE
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

```
# plot and calculation of D
run.D <- function() {
  print(D_index(Cremains_measurements[1, ], plot = TRUE))
  cat("Published D value: ", Cremains_measurements[1, 8], "\n")
}
run.D()

## Not run:
# confidence interval with bootstrapping
D_index(Cremains_measurements[1, ], rand = FALSE, B = 1000)

## End(Not run)
```

---

extract_sum	<i>Summary Statistics Extraction</i>
-------------	--------------------------------------

---

## Description

Extract summary data needed for other functions from raw data.

## Usage

```
extract_sum(x, Sex = 1, Pop = 2, firstX = 3, test = "tg", 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	'tg' for Greene t test <a href="#">t_greeney</a> , 'uni' for <a href="#">univariate</a> , 'aov' for sex specific ANOVA <a href="#">aov_ss</a> , 'multi' for <a href="#">multivariate</a> , and 'van' for <a href="#">van_vark</a> , 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
## Not run:
extract_sum(Howells, test = "multi")
# for univariate test on a specific parameter
extract_sum(Howells, test = "uni", firstX = 4)

## End(Not run)
```

---

FT

*Heuristic data*

---

### Description

Heuristic data from Fidler and Thompson (2001)

### Usage

FT

### Format

A data frame with 24 rows and 3 variables:

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

**Pop** Populations' names

**x** Dependent variable

### References

Fidler, Fiona, and Bruce Thompson. "Computing correct confidence intervals for ANOVA fixed-and random-effects effect sizes." *Educational and Psychological Measurement* 61.4 (2001): 575-604.

---

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,
  verbose = FALSE,
  rand = TRUE,
  digits = 4
)
```

**Arguments**

<code>x</code>	A data frame containing summary statistics.
<code>Trait</code>	Number of the column containing names of measured parameters, Default: 1
<code>CI</code>	confidence interval coverage takes value from 0 to 1, Default: 0.95.
<code>B</code>	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
<code>verbose</code>	logical; if TRUE number of bootstraps is displayed, Default: FALSE
<code>rand</code>	logical; if TRUE, uses random seed. If FALSE, then <code>set.seed(42)</code> for repeatability, Default: TRUE
<code>digits</code>	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("Cre mains_measurements")
# Confidence intervals with non-central t distribution
Hedges_g(Cre mains_measurements[1, ])
## Not run:
# confidence interval with bootstrapping
Hedges_g(Cre mains_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, W. W. (1996). Notes and Comments: Howells' craniometric data on the internet. American Journal of Physical Anthropology, 101(3), 441-442

---

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 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 5 matrices (R.res, M.mu, F.mu, M.sdev, and F.sdev) and two vectors (m and f) with structure similar to [baboon.parms\\_list](#)

---

Howells_V	<i>Pooled within-group variance-covariance matrix for Howells' data</i>
-----------	---

---

**Description**

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

**Usage**

Howells\_V

**Format**

A 8\*8 numerical matrix

**See Also**

[Howells](#)

---

MI_index	<i>Mixture Index ("MI")</i>
----------	-----------------------------

---

**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 (each integrating to 1.0), equivalent to Inman and Bradley's (1989) "overlap coefficient." As a result of this rescaling, the "MI" and "NI" plots will appear identical save for the scale on the y-axis.

**Usage**

```
MI_index(
  x,
  plot = FALSE,
  Trait = 1,
  B = NULL,
  verbose = FALSE,
  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
verbose	logical; if TRUE number of bootstraps is displayed, Default: FALSE
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 Ipiná 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.

Ipiná, 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**

```
# plot and calculation of MI
MI_index(Cremains_measurements[1, ], plot = TRUE)
#' #NI index
MI_index(Cremains_measurements[1, ], index_type = "NI")
1 - D_index(Cremains_measurements[1, ])$D

## Not run:
# confidence interval was bootstrapping
MI_index(Cremains_measurements[1, ], rand = FALSE, B = 1000)

## End(Not run)
```



---

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 correlation 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 <a href="#">p.adjust</a> Default: "none".

...	Additional arguments that could be passed to <a href="#">univariate</a>
lower.tail	Logical; if TRUE probabilities are 'P[X <= x]', otherwise, 'P[X > x]', Default: FALSE
CI	confidence interval coverage for the chosen effect size 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 correlation matrix 'R.res' should be entered as a separate argument as in [baboon.parms\\_R](#). Another acceptable format is a named list of matrices and vectors containing different summary statistics as well as the correlation matrix as in [baboon.parms\\_list](#). By setting the option 'univariate' to 'TRUE', multiple 'ANOVA's can be run on each parameter independently.

### Value

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

### See Also

[baboon.parms\\_df](#)

### Examples

```
# x is a data frame with separate correlation matrix
multivariate(baboon.parms_df, R.res = baboon.parms_R)
# x is a list with the correlation matrix included
multivariate(baboon.parms_list, univariate = TRUE)
# 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 Normal Or Truncated Normal Distribution*


---

**Description**

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

**Usage**

```
raw_gen(
  x,
  Trait = 1,
  Pop = 2,
  R.res = NULL,
  lower = -Inf,
  upper = Inf,
  verbose = 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 correlation matrix, Default: NULL
lower	scalar of lower bounds, Default: -Inf
upper	scalar of upper bounds, Default: Inf
verbose	Logical; if TRUE displays a message with the method used for generation , 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 correlation 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 correlation matrix as in `baboon.parms_R`. If data frame is entered without a correlation matrix, data generation is carried out using univariate distribution.

### Value

a data frame of raw data

### Examples

```
# Data generation using univariate distributions
raw_gen(baboon.parms_df, lower = 0)

# another univariate example
raw_gen(Cremains_measurements[1, ])[, -2]

# Data generation using multivariate distribution
raw_gen(baboon.parms_list, lower = 0)
```

---

SMO

*Hypothetical set of unbalanced data*

---

### Description

Example data set from Shaw and Mitchell-Olds (1993)

### Usage

SMO

### Format

A data frame with 11 rows and 3 variables:

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

**Pop** Populations' names

**x** Dependent variable

### References

Shaw, Ruth G., and Thomas Mitchell-Olds. "ANOVA for unbalanced data: an overview." *Ecology* 74.6 (1993): 1638-1645.

---

t_greene	<i>Greene t test of Sexual Dimorphism</i>
----------	---

---

### 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,
  plot = FALSE,
  colors = c("#DD5129", "#985F51", "#536D79", "#0F7BA2", "#208D98", "#319F8E", "#43B284",
    "#7FB274", "#BCB264", "#FAB255"),
  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
plot	Logical; if TRUE graphical matrix of p values, Default: FALSE
colors	color palette used in the corplot
alternative	a character string specifying the alternative hypothesis, must be one of "two.sided", "greater" or "less", Default: "two.sided"
padjust	Method of p.value adjustment for multiple comparisons following <a href="#">p.adjust</a> 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](#). For the visualization of pairwise comparisons using the corplot, the rounder the image in the plot grid the lower the p-value (see the color scale for similar information). The default colors used in the corplot are from the "MetBrewer" "Egypt"

palette which is listed under the "colorblind\_palettes". Different colors palettes can be selected from "RColorBrewer" package.

## Value

data frame of t.test results

## References

# for the t-test

Greene, David Lee. "Comparison of t-tests for differences in sexual dimorphism between populations." *American Journal of Physical Anthropology* 79.1 (1989): 121-125.

Relethford, John H., and Denise C. Hodges. "A statistical test for differences in sexual dimorphism between populations." *American Journal of Physical Anthropology* 66.1 (1985): 55-61.

#For the femur head diameter data

F. Curate, C. Umbelino, A. Perinha, C. Nogueira, A.M. Silva, E. Cunha, Sex determination from the femur in Portuguese populations with classical and machinelearning classifiers, *J. Forensic Leg. Med.* (2017) , doi:<http://dx.doi.org/10.1016/j.jflm.2017.08.011>.

O. Gulhan, *Skeletal Sexing Standards of Human Remains in Turkey* (PhD thesis), Cranfield University, 2017 [Dataset].

P. Timonov, A. Fasova, D. Radoinova, A. Alexandrov, D. Delev, A study of sexual dimorphism in the femur among contemporary Bulgarian population, *Euras. J. Anthropol.* 5 (2014) 46–53.

E.F. Kranioti, N. Vorniotakis, C. Galiatsou, M.Y. Iscan , M. Michalodimitrakis, Sex identification and software development using digital femoral head radiographs, *Forensic Sci. Int.* 189 (2009) 113.e1–7.

## Examples

```
# Comparisons of femur head diameter in four populations
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,
  padjust = "none"
)
```

**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,
  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 "f2" for f squared, "eta2" for eta squared, "omega2" for omega squared or "none", Default: "none".
pairwise	Logical; if TRUE runs multiple pairwise comparisons on different populations using <a href="#">t_greene</a> Default: FALSE
padjust	Method of p.value adjustment for multiple comparisons following <a href="#">p.adjust</a> Default: "none".
...	Additional arguments that could be passed to the <a href="#">t_greene</a> 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.
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 table.

**References**

Hector, Andy, Stefanie Von Felten, and Bernhard Schmid. "Analysis of variance with unbalanced data: an update for ecology & evolution." *Journal of animal ecology* 79.2 (2010): 308-316.

**Examples**

```
#'
# See Tables 6 and 8 and from Fidler and Thompson (2001).
# The "eta2" and "omega2" CIs match those in Table 8.
# See "FT" dataset for Fidler and Thompson (2001) reference

# acquiring summary data
FT_sum <- extract_sum(FT, test = "uni", run = FALSE)
# univariate analysis on summary data
univariate(FT_sum, CI = 0.90, es_anova = "eta2", digits = 5)
univariate(FT_sum, CI = 0.90, es_anova = "omega2", digits = 5)

# Reproduces Table 2 from Shaw and Mitchell-Olds (1993) using their Table 1.
# See "SMO" dataset for Shaw and Mitchell-Olds (1993) reference
# Note that Table 2 residual df is incorrectly given as 6,
# but is correctly given as 7 in Hector et al. (2010)

# acquiring summary data
univ_SMO <- extract_sum(SMO, test = "uni", run = FALSE)
# univariate analysis on summary data
print(univariate(univ_SMO, type_anova = "I")[[1]])
print(univariate(univ_SMO, type_anova = "II"))
univariate(univ_SMO, type_anova = "III")
```

**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 = NULL,
  q = 2,
  Trait = 1,
  Pop = 2,
  plot = TRUE,
  lower.tail = FALSE,
  digits = 4
)
```

**Arguments**

x	A Data frame of means and sample sizes for different populations or a list of the summary data frame with Pooled within-group variance-covariance matrix.
W	Pooled within-group variance-covariance matrix supplied if x is a dataframe , Default:NULL
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.

**Note**

For plot labels to be fully visualized, maximizing image size is advised.

**References**

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

**Examples**

```
# 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

## \* datasets

- Australia, [4](#)
  - baboon.parms\_df, [5](#)
  - baboon.parms\_list, [6](#)
  - baboon.parms\_R, [6](#)
  - Cremains\_measurements, [7](#)
  - FT, [10](#)
  - Howells, [12](#)
  - Howells\_R, [13](#)
  - Howells\_summary, [13](#)
  - Howells\_summary\_list, [14](#)
  - Howells\_V, [15](#)
  - NHANES\_1999, [18](#)
  - SMO, [20](#)
- [aov\\_ss, 2, 9](#)
- [Australia, 4](#)
- [baboon.parms\\_df, 3, 5, 6, 18, 20, 21, 24](#)
- [baboon.parms\\_list, 6, 6, 14, 18, 20](#)
- [baboon.parms\\_R, 6, 18, 20](#)
- [Cremains\\_measurements, 7](#)
- [D\\_index, 7, 11, 16](#)
- [extract\\_sum, 9](#)
- [FT, 10](#)
- [Hedges\\_g, 11](#)
- [Howells, 10, 12, 14, 15](#)
- [Howells\\_R, 13](#)
- [Howells\\_summary, 13, 14, 25](#)
- [Howells\\_summary\\_list, 14](#)
- [Howells\\_V, 15](#)
- [MI\\_index, 15](#)
- [multivariate, 9, 17](#)
- [NHANES\\_1999, 18](#)
- [p.adjust, 17, 21, 23](#)
- [raw\\_gen, 19](#)
- [SMO, 20](#)
- [t\\_greene, 9, 17, 21, 23](#)
- [univariate, 9, 18, 23](#)
- [van\\_vark, 9, 24](#)