Package 'irrCAC'

May 8, 2026

Version 1.4

Date 2026-04-26

Title Computing the Extent of Agreement among Raters with
Chance-Corrected Agreement Coefficient (CAC)

Description Contains a series of R functions for calculating various chance-corrected agreement coefficients (CAC) among 2 or more raters. Among the CAC coefficients covered are Cohen's kappa, Conger's kappa,
Fleiss' kappa, Brennan-
Prediger coefficient, Gwet's AC1/AC2 coefficients, and Krippendorff's alpha. Multiple sets of weights are proposed for computing weighted analyses. Also included in this package is Bangdiwala's B
coefficient.

Depends R (>= 3.5.0)

Encoding UTF-8

Imports dplyr, magrittr, tidyr, stringr, tibble

Language en-US

LazyData true

License GPL (>= 2)

RoxygenNote 7.3.3

Suggests knitr, rmarkdown, spelling

VignetteBuilder knitr

NeedsCompilation no

Author Kilem L. Gwet [aut, cre]

Maintainer Kilem L. Gwet <gwet@agreestat.com>

Repository CRAN

Date/Publication 2026-04-27 07:20:08 UTC

Contents

`agree.cac3rd` *Dataset showing the distribution of 6 raters by psychiatric condition*

Description

This dataset contains rating data about 15 subjects or patients whose psychiatric condition was evaluated by 6 raters. Each column represents one condition and contains the number of raters who diagnosed the patient with that condition.

Usage

```
agree.cac3rd
```

Format

The first column is of character type whereas the remaining columns contains integers.

Gender Contains subject's gender (M or F)

Depression The interval upper bound

Pers.Disorder Contains the number of raters who labeled the subject as having personal disorder

Schizophrenia Contains the number of raters who labeled the subject as having schizophrenia

Neurosis Contains the number of raters who labeled the subject as having Neurosis

Other Contains the number of raters who labeled the subject as having Other psychiatric condition

Source

K. Gwet, PhD.

agree.contingency *Dataset representing a 10x10 contingency table*

Description

This dataset contains a 10x10 contingency table.

Usage

```
agree.contingency
```

Format

The first column is alphabetic whereas the remaining columns contains integers.

Categories Contains category's name C1,C2, ..., C10)

C1 Colum C1

C2 Colum C2

C3 Colum C3

C4 Colum C4

C5 Colum C5

C6 Colum C6

C7 Colum C7

C8 Colum C8

C9 Colum C9

C10 Colum C10

Source

K. Gwet, PhD.

agreeCAC *Ratings of 15 subjects from 4 raters*

Description

This dataset contains raw ratings of 15 subjects produced by 4 raters named as RaterA, RaterB, RaterC and RaterD.

Usage

```
agreeCAC
```

Format

All 5 columns are of alphabetic type.

Gender Contains the subject's gender (M or F)

RaterA Contains one of ratings a,b,c produced by RaterA

RaterB Contains one of ratings a,b,c produced by RaterB

RaterC Contains one of ratings a,b,c produced by RaterC

RaterD Contains one of ratings a,b,c produced by RaterD

Source

K. Gwet, PhD.

`altman` *Dataset describing the Altman's Benchmarking Scale*

Description

This dataset contains information describing the Altman scale for benchmarking chance-corrected agreement coefficients such as Gwet AC1/AC2, Kappa and many others.

Usage

```
altman
```

Format

Each row of this dataset describes an interval and the interpretation of the magnitude it represents.

lb.AL The interval lower bound

ub.AL The interval upper bound

interp.AL The interpretation

Source

Altman, D.G. (1991). *Practical Statistics for Medical Research*. Chapman and Hall.

`altman.bf` *Computing Altman's Benchmark Scale Membership Probabilities*

Description

Computing Altman's Benchmark Scale Membership Probabilities

Usage

```
altman.bf(coeff, se, BenchDF = altman)
```

Arguments

`coeff`	A mandatory parameter representing the estimated value of an agreement coefficient.
`se`	A mandatory parameter representing the agreement coefficient standard error.
`BenchDF`	An optional parameter that is a 3-column data frame containing the Altman's benchmark scale information. The 3 columns are the interval lower bound, upper bound, and their interpretation. The default value is a small file contained in the package and named *altman.RData*, which describes the official Altman's scale intervals and their interpretation.

Value

A one-column matrix containing the membership probabilities (c.f. https://agreestat.com/papers/inter-rater%20reliability%20study%20design1.pdf)

`bangdiwala.table` *Bangdiwala B coefficient for 2 raters*

Description

Bangdiwala B coefficient for 2 raters

Usage

```
bangdiwala.table(ratings, conflev = 0.95, N = Inf)
```

Arguments

`ratings`	A square table of ratings (assume no missing ratings).
`conflev`	An optional parameter that specifies the confidence level used for constructing confidence intervals. By default the function assumes the standard value of 95%.
`N`	An optional parameter representing the finite population size if any. It is used to perform the finite population correction to the standard error. It's default value is infinity.

Value

A data frame containing the following 5 variables: coeff.name coeff.val coeff.se coeff.ci coeff.pval.

Examples

```
#The dataset "cont3x3abstractors" comes with this package. Analyze it as follows:
bangdiwala.table(cont3x3abstractors) #Yields Bangdiwala coefficient along with precision measures
Bcoeff <- bangdiwala.table(cont3x3abstractors)$coeff.val #Yields Bangdiwala's coefficient alone.
Bcoeff
q <- nrow(cont3x3abstractors) #Number of categories
```

`bangdiwala2RR.fn` *Bangdiwala B coefficient for 2 raters when input dataset is made up of 2 columns of raw data.*

Description

Bangdiwala B coefficient for 2 raters when input dataset is made up of 2 columns of raw data.

Usage

```
bangdiwala2RR.fn(fra.ratings.raw, conflev = 0.95, N = Inf)
```

Arguments

`fra.ratings.raw`	A dataframe with 2 columns of raw ratings.
`conflev`	An optional parameter that specifies the confidence level used for constructing confidence intervals. By default the function assumes the standard value of 95%.
`N`	An optional parameter representing the finite population size if any. It is used to perform the finite population correction to the standard error. It's default value is infinity.

Value

A data frame containing the following 9 variables: coeff.name, b1, b2, coeff.val, coeff.se, conf.int, p.value, n and name of the weight used.

Examples

```
#The dataset cac.ben.gerry comes with this package. Analyze it as follows:
bangdi <- bangdiwala2RR.fn(cac.ben.gerry[,c(3,4)]) #using only the last 2 columns.
#The result will be following:
c(bangdi$coeff.name,bangdi$pa,bangdi$pe,bangdi$coeff.val,bangdi$coeff.se,
  bangdi$conf.int,bangdi$p.val,bangdi$tot.obs)
#1 Bangdiwala''s B 0.1322314 0.2066116      0.64 0.2158518 (0.159,1) 7.083e-03      11
bangdi$w.name
#1 Identity
```

`bipolar.weights` *Function for computing the Bipolar Weights*

Description

Function for computing the Bipolar Weights

Usage

```
bipolar.weights(categ)
```

Arguments

`categ`	A mandatory parameter representing the vector of all possible ratings.

Value

A square matrix of quadratic weights to be used for calculating the weighted coefficients.

`bp.coeff.dist` *Brennan-Prediger's agreement coefficient among multiple raters (2, 3, +) when the input dataset is the distribution of raters by subject and category.*

Description

Brennan-Prediger's agreement coefficient among multiple raters (2, 3, +) when the input dataset is the distribution of raters by subject and category.

Usage

```
bp.coeff.dist(
  ratings,
  weights = "unweighted",
  categ = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	An *nxq* matrix / data frame containing the distribution of raters by subject and category. Each cell *(i,k)* contains the number of raters who classsified subject *i* into category *k*.
weights	is an optional parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parametr categ. Otherwise, only the categories reported will be used.
categ	An optional parameter representing all categories available to raters during the experiment. This parameter may be useful if some categories were not used by any rater inspite of being available to the raters.
conflev	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
N	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A vector containing the following information: pa(the percent agreement),pe(the percent chance agreement),coeff(Brennan-Prediger coefficient), stderr(the standard error of Brennan-Prediger coefficient),conf.int(the p-value of Brennan-Prediger coefficient), p.value(the p-value of Brennan-Prediger coefficient),coeff.name ("Brennan-Prediger").

Source

Brennan, R.L., and Prediger, D. J. (1981). "Coefficient Kappa: some uses, misuses, and alternatives," *Educational and Psychological Measurement*, 41, 687-699.

Examples

```
#The dataset "distrib.6raters" comes with this package. It represents the distribution of 6 raters
#by subject and by category. Note that each row of this dataset sums to the number of raters, which
#is 6. You may this dataset as follows:
distrib.6raters
bp.coeff.dist(distrib.6raters) #BP coefficient, precision measures, weights & list of categories
bp <- bp.coeff.dist(distrib.6raters)$coeff #Yields Brennan-Prediger coefficient alone.
bp
q <- ncol(distrib.6raters) #Number of categories
bp.coeff.dist(distrib.6raters,weights = quadratic.weights(1:q)) #Weighted BP with quadratic weights
```

`bp.coeff.raw` *Brennan & Prediger's (BP) agreement coefficient for an arbitrary number of raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.*

Description

Brennan & Prediger's (BP) agreement coefficient for an arbitrary number of raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.

Usage

```
bp.coeff.raw(
  ratings,
  weights = "unweighted",
  categ.labels = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

`ratings`	An nxr matrix / data frame of ratings where each column represents one rater and each row one subject.
`weights`	is a mandatory parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parametr categ.labels. Otherwise, the program may not work.
`categ.labels`	An optional vector parameter containing the list of all possible ratings. It may be useful in case some of the possibe ratings are not used by any rater, they will still be used when calculating agreement coefficients. The default value is NULL. In this case, only categories reported by the raters are used in the calculations.
`conflev`	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
`N`	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A data list containing 3 objects: (1) a one-row data frame containing various statistics including the requested agreement coefficient, (2) the weight matrix used in the calculations if any, and (3) A

vector of categories used in the analysis. These could be categories reported by the raters, or those available to the raters whether they used them or not. The output data frame contains the following variables: "coeff.name" (coefficient name), "pa" (the percent agreement), "pe" (the percent chance agreement), coeff.val (Brennan-Prediger coefficient estimate), "coeff.se" (standard error), "conf.int" (the confidence interval), "p.value"(Brennan-Prediger coefficient's p-value), "w.name"(the weights' identification).

References

Brennan, R.L., & Prediger, D. J. (1981). "Coefficient Kappa: some uses, misuses, and alternatives." *Educational and Psychological Measurement*, 41, 687-699.

Examples

```
#The dataset "cac.raw4raters" comes with this package. Analyze it as follows:
cac.raw4raters
bp.coeff.raw(cac.raw4raters) #BP coefficient, precision measures, weights & categories
bp.coeff.raw(cac.raw4raters)$est #Brennan-Prediger coefficient with precision measures
bp <- bp.coeff.raw(cac.raw4raters)$est$coeff.val #Yields Brennan-Prediger coefficient alone.
bp
bp.coeff.raw(cac.raw4raters, weights = "quadratic") #weighted Brennan-Prediger coefficient
```

---

bp2.table *Brenann-Prediger coefficient for 2 raters*

Description

Brenann-Prediger coefficient for 2 raters

Usage

```
bp2.table(
  ratings,
  weights = identity.weights(1:ncol(ratings)),
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	A square table of ratings (assume no missing ratings).
weights	An optional matrix that contains the weights used in the weighted analysis. By default, this parameter contaings the identity weight matrix, which leads to the unweighted analysis.
conflev	An optional parameter that specifies the confidence level used for constructing confidence intervals. By default the function assumes the standard value of 95%.
N	An optional parameter representing the finite population size if any. It is used to perform the finite population correction to the standard error. It's default value is infinity.

Value

A data frame containing the following 5 variables: coeff.name coeff.val coeff.se coeff.ci coeff.pval.

Examples

```
#The dataset "cont3x3abstractors" comes with this package. Analyze it as follows:
bp2.table(cont3x3abstractors) #Yields Brennan-Prediger's coefficient along with precision measures
bp <- bp2.table(cont3x3abstractors)$coeff.val #Yields Brennan-Prediger coefficient alone.
bp
q <- nrow(cont3x3abstractors) #Number of categories
bp2.table(cont3x3abstractors,weights = quadratic.weights(1:q)) #Weighted BP coefficient
```

---

cac.ben.gerry *Ratings of 12 units from 2 raters named Ben and Gerry*

Description

This dataset contains ratings that 2 raters named Ben and Gerry assigned to 12 units distributed in 2 groups "G1" and "G2".

Usage

```
cac.ben.gerry
```

Format

Each row of this dataset describes an interval and the interpretation of the magnitude it represents.

Group Group Name

Units Unit number

Ben Ben's Ratings

Gerry Gerry's Ratings

The first 2 columns "Group" and "Units" play a descriptive role here and are not used by any fucntion included in this package. One will typically use cac.ben.gerry[,c(3,4)] or cac.ben.gerry[,c("Ben","Gerry")] as input dataset.

`cac.dist.g1g2` *Distribution of 4 raters by subject and by category, for 14 Subjects that belong to 2 groups "G1" and "G2"*

Description

This dataset contains rating data in the form of a subject-level distribution of 4 raters by category the subject was classified into. A total of 4 raters had to classify 14 subjects into one of 5 categories "a", "b", "c", "d", and "e". This dataset is different version of the more detailed "cac.raw.g1g2" dataset. While "cac.raw.g1g2" tells you about the exact category into which each rater classified all subjects, "cac.dist.g1g2" on the other hand, can only tell you how many raters classified a given subject into a particular category.

Usage

```
cac.dist.g1g2
```

Format

This dataset contains ratings obtained from an experiment where 4 raters classified 14 subjects into 5 possible categories labeled as a, b, c, d, and e. None of the 4 raters scored all 14 units. Therefore, some missing ratings appear in each of the columns associated with the 4 raters.

Note that only the the 4 last columns are to be used with the functions included in this package. The first 2 columns only play a descriptive role and are not used in any calculation.

Group This variable represents the group name.

Units This variable represents the unit number.

a Number of raters who classified the subject represented by the row into category "a"

b Number of raters who classified the subject represented by the row into category "b"

c Number of raters who classified the subject represented by the row into category "c"

d Number of raters who classified the subject represented by the row into category "d"

e Number of raters who classified the subject represented by the row into category "e"

`cac.dist4cat` *Distribution of 4 raters by Category and Subject - Subjects allocated in 2 groups A and B.*

Description

This dataset summarizes the ratings assigned by 4 raters who classified 15 subjects into one of 3 categories named "a", "b", and "c".

Usage

```
cac.dist4cat
```

Format

This dataset has 15 rows (for the 15 subjects) and 4 columns. Only the last 3 columns representing the categories into which subjects are classified are used in the calculations - unless the sub-group analysis is required.

Group This variable repsents the subject number.

a category a

b Category b

c Category c

`cac.raw.g1g2` *Dataset of raw ratings from 4 Raters on 14 Subjects that belong to 2 groups named "G1" and "G2"*

Description

This dataset contains data from a reliability experiment where 4 raters identified as Rater1, Rater2, Rater3 and Rater4 scored 14 units on a 5-point alphabetical scale based on the values a, b, c, d and e. These 14 units are allocated to 2 groups named G1 and G2.

Usage

```
cac.raw.g1g2
```

Format

This dataset contains ratings obtained from an experiment where 4 raters classified 14 subjects into 5 possible categories labeled as a, b, c, d, and e. None of the 4 raters scored all 14 units. Therefore, some missing ratings appear in each of the columns associated with the 4 raters.

Note that only the the 4 last columns are to be used with the functions included in this package. The first 2 columns only play a descriptive role and are not used in any calculation.

Group This variable repsents the unit number.

Units This variable repsents the unit number.

Rater1 All ratings from rater 1

Rater2 All ratings from rater 2

Rater3 All ratings from rater 3

Rater4 All ratings from rater 4

cac.raw.gender *Rating Data from 4 Raters and 15 human Subjects, 9 of whom are female and 6 males.*

Description

This dataset contains data from a reliability experiment where 4 raters scored 15 units on a 3-point alphabetic scale based on the values a, b, and c.

Usage

```
cac.raw.gender
```

Format

This dataset contains ratings obtained from an experiment where 4 raters classiffied 15 subjects into 3 possible categories labeled as a, b, and c.

Note that only the the 4 last columns are to be used with the functions included in this package. The first column only plays a descriptive role and is not to be used in any calculation.

Group This variable repsents the unit number.

RaterA All ratings from rater 1

RaterB All ratings from rater 2

RaterC All ratings from rater 3

RaterD All ratings from rater 4

cac.raw2raters *Dataset of raw ratings by 2 raters and 12 subjects.*

Description

This dataset contains 12 rows and 2 columns. Each row represents a subject and the 2 columns labeled as "rater1" and "rater2" contain the ratings produced by the 2 raters.

Usage

```
cac.raw2raters
```

Format

This dataset contains ratings obtained from an experiment where 2 raters classified 12 subjects into 3 categories A, B, and C. Two of the 12 subjects were rated by a single rater. Consequently, this dataset contains 2 missing ratings that are identified with the symbol <NA>.

rater1 All ratings from rater 1

rater2 All ratings from rater 2

Source

Gwet, K.L. (2021) *Handbook of Inter-Rater Reliability: Volume 1*, 5th Edition. AgreeStat Analytics.

`cac.raw4raters` *Rating Data from 4 Raters and 12 Subjects.*

Description

This dataset contains data from a reliability experiment where 5 observers scored 15 units on a 4-point numeric scale based on the values 0, 1, 2 and 3.

Usage

```
cac.raw4raters
```

Format

This dataset contains ratings obtained from an experiment where 4 raters classified 12 subjects into 5 possible categories labeled as 1, 2, 3, 4, and 5. None of the 4 raters scored all 12 units. Therefore, some missing ratings in the form of "NA" appear in each of the columns associated with the 4 raters.

Note that only the the 4 last columns are to be used with the functions included in this package. The first column only plays a descriptive role and is not used in any calculation.

Units This variable repsents the unit number.

Rater1 All ratings from rater 1

Rater2 All ratings from rater 2

Rater3 All ratings from rater 3

Rater4 All ratings from rater 4

Source

Gwet, K.L. (2014) *Handbook of Inter-Rater Reliability*, 4th Edition, page #120. Advanced Analytics, LLC.

cac.raw5obser *Scores assigned by 5 observers to 20 experimental units.*

Description

This dataset contains data from a reliability experiment where 5 observers scored 15 units on a 4-point numeric scale based on the values 0, 1, 2 and 3.

Usage

```
cac.raw5obser
```

Format

This dataset has 15 rows (for the 15 subjects) and 6 columns. Only the last 5 columns associated with the 5 observers are used in the calculations. Of the 5 observers, only observer 3 scored all 15 units. Therefore, some missing ratings in the form of "NA" appear in the columns associated with the remaining 4 observers.

Unit This variable repsents the unit number.

Observer1 All ratings from Observer 1

Observer2 All ratings from Observer 2

Observer3 All ratings from Observer 3

Observer4 All ratings from Observer 4

Observer5 All ratings from Observer 5

Source

Gwet, K.L. (2014) *Handbook of Inter-Rater Reliability*, 4th Edition. Advanced Analytics, LLC. *A larger version of this table can be found on page #125*

circular.weights *Function for computing the Circular Weights*

Description

Function for computing the Circular Weights

Usage

```
circular.weights(categ)
```

Arguments

categ	A mandatory parameter representing the vector of all possible ratings.

Value

A square matrix of quadratic weights to be used for calculating the weighted coefficients.

`conger.kappa.raw` *Conger's generalized kappa coefficient for an arbitrary number of raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.*

Description

Conger's generalized kappa coefficient for an arbitrary number of raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.

Usage

```
conger.kappa.raw(
  ratings,
  weights = "unweighted",
  categ.labels = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

`ratings`	An nxr matrix / data frame of ratings where each column represents one rater and each row one subject.
`weights`	is a mandatory parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parametr categ.labels. Otherwise, the program may not work.
`categ.labels`	An optional vector parameter containing the list of all possible ratings. It may be useful in case some of the possibe ratings are not used by any rater, they will still be used when calculating agreement coefficients. The default value is NULL. In this case, only categories reported by the raters are used in the calculations.
`conflev`	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
`N`	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A data list containing 3 objects: (1) a one-row data frame containing various statistics including the requested agreement coefficient, (2) the weight matrix used in the calculations if any, and (3) A vector of categories used in the analysis. These could be categories reported by the raters, or those available to the raters whether they used them or not. The output data frame contains the following variables: "coeff.name" (coefficient name), "pa" (the percent agreement), "pe" (the percent chance agreement), coeff.val (Conger's Kappa estimate), "coeff.se" (standard error), "conf.int" (Conger Kappa's confidence interval), "p.value"(agreement coefficient's p-value), "w.name"(the weights' identification).

References

Conger, A. J. (1980), "Integration and Generalization of Kappas for Multiple Raters," *Psychological Bulletin*, 88, 322-328.

Examples

```
#The dataset "cac.raw4raters" comes with this package. Analyze it as follows:
cac.raw4raters
conger.kappa.raw(cac.raw4raters) #Conger's kappa, precision stats, weights & categories
conger.kappa.raw(cac.raw4raters)$est #Conger's kappa with precision measures
conger <- conger.kappa.raw(cac.raw4raters)$est$coeff.val #Yields Conger's kappa alone.
conger
conger.kappa.raw(cac.raw4raters, weights = "quadratic") #weighted Conger's kappa
```

---

cont3x3abstractors *Distribution of 100 pregnant women by pregnancy type and by abstractor.*

Description

This dataset contains pregnancy type data collected from 100 women who entered an Emergency Room with a positive pregnancy test and a second condition, which is either abdominal pain or vaginal bleeding. After reviewing their medical records, 2 reviewers (also referred to as abstractors) classified them into one of the following three pregnancy categories: Ectopic Pregnancy (Ectopic), Abnormal Intrauterine pregnancy (AIU) and Normal Intrauterine Pregnancy (NIU).

Usage

```
cont3x3abstractors
```

Format

Each row of this dataset describes an interval and the interpretation of the magnitude it represents.

Type Pregnancy Type. This variable is shown here for information only and is never used by any function in the irrCAC package.

Ectopic Ectopic Pregnancy

AIU Abnormal Intrauterine Pregnancy

NIU Normal Intrauterine Pregnancy

Source

Gwet, K.L. (2014). *Handbook of Inter-Rater Reliability*, 4th Edition. Advanced Analytics, LLC.

`cont4x4diagnosis` *Distribution of 223 Psychiatric Patients by Type of of Psychiatric Disorder and Diagnosis Method.*

Description

This dataset shows the distribution of 223 psychiatric patients by diagnosis category and by the method used to obtain the diagnosis. The first method named “Clinical Diagnosis" (also known as “Facility Diagnosis") is used in a service facility (e.g. public hospital, or a community unit) and does not rely on a rigorous application of research criteria. The second method known as “Research Diagnosis" is based on a strict application of research criteria. Column 1 contains the diagnosis categories into which patients are classified with Method 1. The first row on the other hand, shows categories into which patients are classified with Method 2.

Usage

```
cont4x4diagnosis
```

Format

This dataset contains a 4x4 squared table. The first column is never used in the calculations and only contains row names. Only the last 4 columns are used for computing agreement coefficients.

Diagnosis Pregnancy Type. This variable is shown here for information only and is never used by any function in the irrCAC package.

Schizophrenia Ectopic Pregnancy

Bipolar.Disorder Abnormal Intrauterine Pregnancy

Depression Normal Intrauterine Pregnancy

Other Normal Intrauterine Pregnancy

Source

Gwet, K.L. (2014). *Handbook of Inter-Rater Reliability*, 4th Edition. Advanced Analytics, LLC.

`distrib.6raters` *Distribution of 6 psychiatrists by Subject/patient and diagnosis Category.*

Description

This dataset summarizes the ratings assigned by 6 psychiatrists classifying 15 patients into one of five categories named "Depression", "Personal Disorder", "Schizophrenia", "Neurosis" and "Other".

Usage

```
distrib.6raters
```

Format

This dataset has 15 rows (for the 15 subjects) and 7 columns. Only the last 6 columns representing the categories into which subjects are classified are used in the calculations.

Subject This variable repsents the subject number.

Personality.Disorder Personality disorder category

Schizophrenia Schizophrenia Category

Neurosis Neurosis category

Other "Other" category

Source

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters, *Psychological Bulletin*, 76, 378-382.

`fleiss` *Dataset describing Fleiss' Benchmarking Scale*

Description

This dataset contains information describing Fleiss' scale for benchmarking chance-corrected agreement coefficients such as Gwet AC1/AC2, Kappa and many others.

Usage

```
fleiss
```

Format

Each row of this dataset describes an interval and the interpretation of the magnitude it represents.

lb.FL The interval lower bound

ub.FL The interval upper bound

interp.FL The interpretation

Source

Fleiss, J. L. (1981). Statistical Methods for Rates and Proportions. John Wiley & Sons.

`fleiss.bf` *Computing Fleiss Benchmark Scale Membership Probabilities*

Description

Computing Fleiss Benchmark Scale Membership Probabilities

Usage

```
fleiss.bf(coeff, se, BenchDF = fleiss)
```

Arguments

`coeff`	A mandatory parameter representing the estimated value of an agreement coefficient.
`se`	A mandatory parameter representing the agreement coefficient standard error.
`BenchDF`	An optional parameter that is a 3-column data frame containing the Fleiss' benchmark scale information. The 3 columns are the interval lower bound, upper bound, and their interpretation. The default value is a small file contained in the package and named *fleiss.RData*, which describes the fleiss' scale intervales and their interpretation.

Value

A one-column matrix containing the membership probabilities (c.f. https://agreestat.com/papers/inter-rater%20reliability%20study%20design1.pdf)

`fleiss.kappa.dist` *Fleiss' agreement coefficient among multiple raters (2, 3, +) when the input dataset is the distribution of raters by subject and category.*

Description

Fleiss' agreement coefficient among multiple raters (2, 3, +) when the input dataset is the distribution of raters by subject and category.

Usage

```
fleiss.kappa.dist(
  ratings,
  weights = "unweighted",
  categ = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

`ratings`	An *nxq* matrix / data frame containing the distribution of raters by subject and category. Each cell *(i,k)* contains the number of raters who classsified subject *i* into category *k*.
`weights`	is an optional parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parametr categ. Otherwise, only the categories reported will be used.
`categ`	An optional parameter representing all categories available to raters during the experiment. This parameter may be useful if some categories were not used by any rater inspite of being available to the raters.
`conflev`	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
`N`	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A vector containing the following information: pa(the percent agreement),pe(the percent chance agreement),coeff(Fleiss' agreement coefficient), stderr(the standard error of Fleiss' coefficient),conf.int(the confidence interval of Fleiss Kappa coefficient), p.value(the p-value of Fleiss' coefficient),coeff.name ("Fleiss").

Source

Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions*. John Wiley & Sons.

Examples

```
#The dataset "distrib.6raters" comes with this package. It represents the distribution of 6 raters
#by subject and by category. Note that each row of this dataset sums to the number of raters, which
#is 6. You may this dataset as follows:
distrib.6raters
fleiss.kappa.dist(distrib.6raters) #Fleiss' kappa, precision measures, weights & list of categories
fleiss <- fleiss.kappa.dist(distrib.6raters)$coeff #Yields Fleiss' kappa alone.
fleiss
q <- ncol(distrib.6raters) #Number of categories
fleiss.kappa.dist(distrib.6raters,weights = quadratic.weights(1:q)) #Weighted fleiss/quadratic wts
```

`fleiss.kappa.raw` *Fleiss' generalized kappa among multiple raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.*

Description

Fleiss' generalized kappa among multiple raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.

Usage

```
fleiss.kappa.raw(
  ratings,
  weights = "unweighted",
  categ.labels = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

`ratings`	An nxr matrix / data frame of ratings where each column represents one rater and each row one subject.
`weights`	is a mandatory parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parametr categ.labels. Otherwise, the program may not work.

categ.labels	An optional vector parameter containing the list of all possible ratings. It may be useful in case some of the possibe ratings are not used by any rater, they will still be used when calculating agreement coefficients. The default value is NULL. In this case, only categories reported by the raters are used in the calculations.
conflev	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
N	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A data list containing 3 objects: (1) a one-row data frame containing various statistics including the requested agreement coefficient, (2) the weight matrix used in the calculations if any, and (3) the categories used in the analysis. These could be categories reported by the raters, or those that were available to the raters whether they used them or not. The output data frame contains the following variables: "coeff.name" (coefficient name-here it will be "Fleiss' Kappa"), "pa" (the percent agreement), "pe" (the percent chance agreement), coeff.val (the agreement coefficient estimate-Fleiss' Kappa), "coeff.se" (the standard error), "conf.int" (Fleiss Kappa's confidence interval), "p.value"(Fleiss Kappa's p-value), "w.name"(the weights' identification).

References

Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions*. John Wiley & Sons.

Examples

```
#The dataset "cac.raw4raters" comes with this package. Analyze it as follows:
cac.raw4raters
fleiss.kappa.raw(cac.raw4raters) #Fleiss' kappa, precision measures, weights & categories
fleiss.kappa.raw(cac.raw4raters)$est #Yields Fleiss' kappa with precision measures
fleiss <- fleiss.kappa.raw(cac.raw4raters)$est$coeff.val #Yields Fleiss' kappa alone.
fleiss
fleiss.kappa.raw(cac.raw4raters, weights = "quadratic") #weighted Fleiss' kappa/quadratic wts
```

---

freq.supp.fn *freq.supp.fn: This function reads a 3-variable input data file containing unique pairs of categories along with their frequency of occurrences, and outputs a similar file where all possible pairs of categories are represented, some with a frequency of occurrence of 0.*

Description

freq.supp.fn: This function reads a 3-variable input data file containing unique pairs of categories along with their frequency of occurrences, and outputs a similar file where all possible pairs of categories are represented, some with a frequency of occurrence of 0.

Usage

```
freq.supp.fn(freq.data, categories.vec)
```

Arguments

freq.data	The input data file containing all unique combinations of reported categories.
categories.vec	A vector containing the complete set of categories available to raters (e.g. "a", "b", "c", "d", "e"). The raters will not necessarily use all of these categories.

Value

This function returns a complete data frame containing all possible combinations of of categories in the categories.vec vector. Newly-added combinations of categories will have a frequency occurrence of 0.

Examples

```
#The dataset "freqs.data" comes with this package. Analyze it as follows:
freq.supp.fn(freqs.data)
#Executing this command will yield the following data frame:
#  Ben   Gerry n
# <chr> <chr> <chr>
#   a     b     1
#   a     d     1
#   b     b     2
#   c     c     3
#   d     b     1
#   d     d     1
#   e     e     1
#   a     a     0
```

---

freqs.data *Distribution of 10 subjects by rater (Ben and Gerry) and by category.*

Description

This dataset contains rating data collected from 10 human subjects by 2 raters named Ben and Gerry. While the 5 categories a, b, c, d, and e are available for rating, Gerry never used category a.

Usage

```
freqs.data
```

Format

Each row of this dataset describes a unique pair of categories and the number of subjects that both Ben and Gerry classified into the 2 categories respectively.

Ben Categories that Ben used to classify the subjects.

Gerry Categories that Gerry used to classify the subjects

n Number of subjects that Ben and Gerry classified into the associated pair of categories.

Source

N/A.

`gwet.ac1.dist` *Gwet's AC1/AC2 agreement coefficient among multiple raters (2, 3, +) when the input dataset is the distribution of raters by subject and category.*

Description

Gwet's AC1/AC2 agreement coefficient among multiple raters (2, 3, +) when the input dataset is the distribution of raters by subject and category.

Usage

```
gwet.ac1.dist(
  ratings,
  weights = "unweighted",
  categ = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

`ratings`	An *nxq* matrix / data frame containing the distribution of raters by subject and category. Each cell *(i,k)* contains the number of raters who classsified subject *i* into category *k*.
`weights`	is an optional parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parameter categ. Otherwise, only the categories reported will be used.

categ	An optional parameter representing all categories available to raters during the experiment. This parameter may be useful if some categories were not used by any rater inspite of being available to the raters.
conflev	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
N	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A vector containing the following information: pa(the percent agreement),pe(the percent chance agreement), coeff(Gwet's AC1 or AC2 dependending on whether weights are used or not),stderr(the standard error of Gwet's coefficient), conf.int(the confidence interval of Gwet's coefficient), p.value(the p-value of Gwet's coefficient),coeff.name (AC1/AC2).

Source

Gwet, K. L. (2008). "Computing inter-rater reliability and its variance in the presence of high agreement," *British Journal of Mathematical and Statistical Psychology*, 61, 29-48.

Examples

```
#The dataset "distrib.6raters" comes with this package. It represents the distribution of 6 raters
#by subject and by category. Note that each row of this dataset sums to the number of raters, which
#is 6. You may this dataset as follows:
distrib.6raters
gwet.ac1.dist(distrib.6raters) #AC1 coefficient, precision measures, weights & list of categories
ac1 <- gwet.ac1.dist(distrib.6raters)$coeff #Yields AC1 coefficient alone.
ac1
q <- ncol(distrib.6raters) #Number of categories
gwet.ac1.dist(distrib.6raters,weights = quadratic.weights(1:q)) #AC2 with quadratic weights
```

---

gwet.ac1.raw *Gwet's AC1/AC2 agreement coefficient among multiple raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.*

Description

Gwet's AC1/AC2 agreement coefficient among multiple raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.

Usage

```
gwet.ac1.raw(
  ratings,
  weights = "unweighted",
  categ.labels = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	An nxr matrix / data frame of ratings where each column represents one rater and each row one subject.
weights	is a mandatory parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parametr categ.labels. Otherwise, the program may not work.
categ.labels	An optional vector parameter containing the list of all possible ratings. It may be useful in case some of the possibe ratings are not used by any rater, they will still be used when calculating agreement coefficients. The default value is NULL. In this case, only categories reported by the raters are used in the calculations.
conflev	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
N	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A data list containing 3 objects: (1) a one-row data frame containing various statistics including the requested agreement coefficient, (2) the weight matrix used in the calculations if any, and (3) the categories used in the analysis. These could be categories reported by the raters, or those that were available to the raters whether they used them or not. The output data frame contains the following variables: "coeff.name" (coefficient name), "pa" (the percent agreement), "pe" (the percent chance agreement), coeff.val (the agreement coefficient estimate-AC1 or AC2), "coeff.se" (the standard error), "conf.int" (AC1/AC2 confidence interval), "p.value" (Gwet AC1/AC2 p-value), "w.name"(the weights' identification).

References

Gwet, K. L. (2008). "Computing inter-rater reliability and its variance in the presence of high agreement." *British Journal of Mathematical and Statistical Psychology*, 61, 29-48.

Examples

```
#The dataset "cac.raw4raters" comes with this package. Analyze it as follows:
cac.raw4raters
gwet.ac1.raw(cac.raw4raters) #AC1 coefficient, precision measures, weights & categories
gwet.ac1.raw(cac.raw4raters)$est #Yields AC1 coefficient with precision measures
ac1 <- gwet.ac1.raw(cac.raw4raters)$est$coeff.val #Yields AC1 coefficient alone.
ac1
gwet.ac1.raw(cac.raw4raters, weights = "quadratic") #AC2 coefficient with quadratic wts
```

---

gwet.ac1.table *Gwet's AC1/AC2 coefficient for 2 raters*

Description

Gwet's AC1/AC2 coefficient for 2 raters

Usage

```
gwet.ac1.table(
  ratings,
  weights = identity.weights(1:ncol(ratings)),
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	A square table of ratings (assume no missing ratings).
weights	An optional matrix that contains the weights used in the weighted analysis. By default, this parameter contaings the identity weight matrix, which leads to the unweighted analysis.
conflev	An optional parameter that specifies the confidence level used for constructing confidence intervals. By default the function assumes the standard value of 95%.
N	An optional parameter representing the finite population size if any. It is used to perform the finite population correction to the standard error. It's default value is infinity.

Value

A data frame containing the following 5 variables: coeff.name coeff.val coeff.se coeff.ci coeff.pval.

Examples

```
#The dataset "cont3x3abstractors" comes with this package. Analyze it as follows:
gwet.ac1.table(cont3x3abstractors) #Yields AC1 along with precision measures
ac1 <- gwet.ac1.table(cont3x3abstractors)$coeff.val #Yields AC1 coefficient alone.
ac1
q <- nrow(cont3x3abstractors) #Number of categories
gwet.ac1.table(cont3x3abstractors,weights = quadratic.weights(1:q)) #AC2 with quadratic weights
```

identity.weights *Function for computing the Identity Weights*

Description

Function for computing the Identity Weights

Usage

```
identity.weights(categ)
```

Arguments

categ	A mandatory parameter representing the vector of all possible ratings.

Value

A square matrix of identity weights to be used for calculating the unweighted coefficients.

kappa2.table *Kappa coefficient for 2 raters*

Description

Kappa coefficient for 2 raters

Usage

```
kappa2.table(
  ratings,
  weights = identity.weights(1:ncol(ratings)),
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	A square or contingency table of ratings (assume no missing ratings). See the 2 datasets "cont3x3abstractors" and "cont4x4diagnosis" that come with this package as examples.
weights	An optional matrix that contains the weights used in the weighted analysis.
conflev	An optional confidence level for confidence intervals. The default value is the traditional 0.95.
N	An optional population size. The default value is infinity.

Value

A data frame containing the following 5 variables: coeff.name coeff.val coeff.se coeff.ci coeff.pval.

Examples

```
#The dataset "cont3x3abstractors" comes with this package. Analyze it as follows:
kappa2.table(cont3x3abstractors) #Yields Cohen's kappa along with precision measures
kappa <- kappa2.table(cont3x3abstractors)$coeff.val #Yields Cohen's kappa alone.
kappa
q <- nrow(cont3x3abstractors) #Number of categories
kappa2.table(cont3x3abstractors,weights = quadratic.weights(1:q))#weighted kappa/quadratic wts
```

`krippen.alpha.dist` *Krippendorff's agreement coefficient among multiple raters (2, 3, +) when the input dataset is the distribution of raters by subject and category.*

Description

Krippendorff's agreement coefficient among multiple raters (2, 3, +) when the input dataset is the distribution of raters by subject and category.

Usage

```
krippen.alpha.dist(
  ratings,
  weights = "unweighted",
  categ = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

`ratings`	An *nxq* matrix / data frame containing the distribution of raters by subject and category. Each cell *(i,k)* contains the number of raters who classsified subject *i* into category *k*.
`weights`	is an optional parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parametr categ. Otherwise, only the categories reported will be used.

categ	An optional parameter representing all categories available to raters during the experiment. This parameter may be useful if some categories were not used by any rater inspite of being available to the raters.
conflev	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
N	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A vector containing the following information: pa(the percent agreement),pe(the percent chance agreement),coeff(Krippendorff's alpha), stderr(the standard error of Krippendorff's coefficient),conf.int(the confidence interval of Krippendorff's alpha coefficient), p.value(the p-value of Krippendorff's alpha), coeff.name ("krippen alpha").

Source

Gwet, K. (2014). *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Multiple Raters*, 4th Edition. Advanced Analytics, LLC Krippendorff (1970). "Bivariate agreement coefficients for reliability of data," *Sociological Methodology*,2,139-150 Krippendorff (1980). *Content analysis: An introduction to its methodology (2nd ed.)*, New-bury Park, CA: Sage.

Examples

```
#The dataset "distrib.6raters" comes with this package. It represents the distribution of 6 raters
#by subject and by category. Note that each row of this dataset sums to the number of raters, which
#is 6. You may this dataset as follows:
distrib.6raters
krippen.alpha.dist(distrib.6raters) #Krippendorff's alpha, precision measures, weights & categories
alpha <- krippen.alpha.dist(distrib.6raters)$coeff #Yields Krippendorff's alpha coefficient alone.
alpha
q <- ncol(distrib.6raters) #Number of categories
krippen.alpha.dist(distrib.6raters,weights = quadratic.weights(1:q)) #Weighted alpha
```

---

krippen.alpha.raw *Krippendorff's alpha coefficient for an arbitrary number of raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.*

Description

Krippendorff's alpha coefficient for an arbitrary number of raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.

Usage

```
krippen.alpha.raw(
  ratings,
  weights = "unweighted",
  categ.labels = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	An nxr matrix / data frame of ratings where each column represents one rater and each row one subject.
weights	is a mandatory parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parametr categ.labels. Otherwise, the program may not work.
categ.labels	An optional vector parameter containing the list of all possible ratings. It may be useful in case some of the possibe ratings are not used by any rater, they will still be used when calculating agreement coefficients. The default value is NULL. In this case, only categories reported by the raters are used in the calculations.
conflev	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
N	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A data list containing 3 objects: (1) a one-row data frame containing various statistics including the requested agreement coefficient-in this case, Krippendorff's alpha, (2) the weight matrix used in the calculations if any, and (3) the vector of categories used in the analysis. These could be categories reported by the raters, or those that were available to the raters whether they used them or not. The output data frame contains the following variables: "coeff.name" (coefficient name), "pa" (the percent agreement), "pe" (the percent chance agreement), coeff.val (Krippendorff's alpha estimate), "coeff.se (standard error), conf.int" (Krippendorff alpha's confidence interval),"p.value" (Krippendorff alpha's p-value), "w.name" (the weights' identification).

References

Gwet, K. (2014). *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Multiple Raters*, 4th Edition. Advanced Analytics, LLC.

Krippendorff (1970). “Bivariate agreement coefficients for reliability of data." *Sociological Methodology*,2,139-150.

Krippendorff (1980). *Content analysis: An introduction to its methodology* (2nd ed.), New-bury Park, CA: Sage.

Examples

```
#The dataset "cac.raw4raters" comes with this package. Analyze it as follows:
cac.raw4raters
krippen.alpha.raw(cac.raw4raters) #Alpha coeff. , precision measures, weights & categories
krippen.alpha.raw(cac.raw4raters)$est #Krippendorff's alpha with precision measures
alpha <- krippen.alpha.raw(cac.raw4raters)$est$coeff.val #Krippendorff's alpha alone.
alpha
krippen.alpha.raw(cac.raw4raters, weights = "quadratic") #weighted alpha/ quadratic wts
```

---

krippen2.table *Krippendorff's Alpha coefficient for 2 raters*

Description

Krippendorff's Alpha coefficient for 2 raters

Usage

```
krippen2.table(
  ratings,
  weights = identity.weights(1:ncol(ratings)),
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	A square table of ratings (assume no missing ratings).
weights	An optional matrix that contains the weights used in the weighted analysis. By default, this parameter contaings the identity weight matrix, which leads to the unweighted analysis.
conflev	An optional parameter that specifies the confidence level used for constructing confidence intervals. By default the function assumes the standard value of 95%.
N	An optional parameter representing the finite population size if any. It is used to perform the finite population correction to the standard error. It's default value is infinity.

Value

A data frame containing the following 5 variables: coeff.name coeff.val coeff.se coeff.ci coeff.pval.

Examples

```
#The dataset "cont3x3abstractors" comes with this package. Analyze it as follows:
krippen2.table(cont3x3abstractors) #Krippendorff's alpha along with precision measures
alpha <- krippen2.table(cont3x3abstractors)$coeff.val #Krippendorff's alpha alone.
alpha
q <- nrow(cont3x3abstractors) #Number of categories
krippen2.table(cont3x3abstractors,weights = quadratic.weights(1:q)) #Weighted alpha coefficient
```

---

landis.koch *Dataset describing the Landis & Koch Benchmarking Scale*

Description

This dataset contains information describing the Landis & Koch scale for benchmarking chance-corrected agreement coefficients such as Gwet AC1/AC2, Kappa and many others.

Usage

```
landis.koch
```

Format

Each row of this dataset describes an interval and the interpretation of the magnitude it represents.

lb.LK The interval lower bound

ub.LK The interval upper bound

interp.LK The interpretation

Source

Landis, J.R. & Koch G. (1977). The measurement of observer agreement for categorical data, *Biometrics*, 33, 159-174.

---

landis.koch.bf *Computing Landis-Koch Benchmark Scale Membership Probabilities*

Description

Computing Landis-Koch Benchmark Scale Membership Probabilities

Usage

```
landis.koch.bf(coeff, se, BenchDF = landis.koch)
```

Arguments

coeff	A mandatory parameter representing the estimated value of an agreement coefficient.
se	A mandatory parameter representing the agreement coefficient standard error.
BenchDF	An optional parameter that is a 3-column data frame containing the Landis & Koch's benchmark scale information. The 3 columns are the interval lower bound, upper bound, and their interpretation. The default value is a small file contained in the package and named *landis.koch.RData*, which describes the official Landis & Koch's scale intervals and their interpretation.

Value

A one-column matrix containing the membership probabilities (c.f. https://agreestat.com/papers/inter-rater%20reliability%20study%20design1.pdf)

---

linear.weights *Function for computing the Linear Weights*

Description

Function for computing the Linear Weights

Usage

```
linear.weights(categ)
```

Arguments

categ	A mandatory parameter representing the vector of all possible ratings.

Value

A square matrix of quadratic weights to be used for calculating the weighted coefficients.

`long2wide.fn` *long2wide.fn: This function transforms a 3-column dataset of frequencies to a square matrix or a contingency table. This function uses the freq.supp.fn() function.*

Description

long2wide.fn: This function transforms a 3-column dataset of frequencies to a square matrix or a contingency table. This function uses the freq.supp.fn() function.

Usage

```
long2wide.fn(freqs.long)
```

Arguments

`freqs.long`	A 3-column data frame, where the first 2 variables represent the categories that both raters have actually used when classifying the subjects. The third and last variable is generally named "n" and represents the count of subjects that classified into the 2 associated categories by both raters.

Value

A matrix that represents a contingency showing the distribution of subjects by rater and category.

Examples

```
#The dataset "freqs.data" comes with this package. Analyze it as follows:
long2wide.fn(freqs.data) #Yields a 5x5 matrix
#This will produce the following 5x5 matrix:
#> long2wide.fn(freqs.data)
#a b c d e
#a 0 1 0 1 0
#b 0 2 0 0 0
#c 0 0 3 0 0
#d 0 1 0 1 0
#e 0 0 0 0 1
```

`ordinal.weights` *Function for computing the Ordinal Weights*

Description

Function for computing the Ordinal Weights

Usage

```
ordinal.weights(categ)
```

Arguments

categ	A mandatory parameter representing the vector of all possible ratings.

Value

A square matrix of quadratic weights to be used for calculating the weighted coefficients.

---

pa.coeff.dist *Percent agreement coefficient among multiple raters (2, 3, +) when the input dataset is the distribution of raters by subject and category.*

Description

Percent agreement coefficient among multiple raters (2, 3, +) when the input dataset is the distribution of raters by subject and category.

Usage

```
pa.coeff.dist(
  ratings,
  weights = "unweighted",
  categ = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	An *nxq* matrix / data frame containing the distribution of raters by subject and category. Each cell *(i,k)* contains the number of raters who classsified subject *i* into category *k*.
weights	is an optional parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parametr categ. Otherwise, only the categories reported will be used.
categ	An optional parameter representing all categories available to raters during the experiment. This parameter may be useful if some categories were not used by any rater inspite of being available to the raters.

conflev	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
N	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A vector containing the following information: pa(the percent agreement),pe(the percent chance agreement),coeff(Brennan-Prediger coefficient), stderr(the standard error of Brennan-Prediger coefficient),conf.int(the p-value of Brennan-Prediger coefficient), p.value(the p-value of Brennan-Prediger coefficient),coeff.name ("Brennan-Prediger").

Source

Brennan, R.L., and Prediger, D. J. (1981). "Coefficient Kappa: some uses, misuses, and alternatives," *Educational and Psychological Measurement*, 41, 687-699.

Examples

```
#The dataset "distrib.6raters" comes with this package. It represents the distribution of 6 raters
#by subject and by category. Note that each row of this dataset sums to the number of raters, which
#is 6. You may this dataset as follows:
distrib.6raters
pa.coeff.dist(distrib.6raters) #percent agreement, precision measures, weights& list of categories
pa <- pa.coeff.dist(distrib.6raters)$coeff #Yields the percent agreement coefficient alone.
pa
q <- ncol(distrib.6raters) #Number of categories
pa.coeff.dist(distrib.6raters,weights = quadratic.weights(1:q)) #Weighted percent agreement
```

---

pa.coeff.raw *Percent agreement among multiple raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.*

Description

Percent agreement among multiple raters (2, 3, +) when the input data represent the raw ratings reported for each subject and each rater.

Usage

```
pa.coeff.raw(
  ratings,
  weights = "unweighted",
  categ.labels = NULL,
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	An nxr matrix / data frame of ratings where each column represents one rater and each row one subject.
weights	is a mandatory parameter that is either a string variable or a matrix. The string describes one of the predefined weights and must take one of the values ("quadratic", "ordinal", "linear", "radical", "ratio", "circular", "bipolar"). If this parameter is a matrix then it must be a square matri qxq where q is the number of posssible categories where a subject can be classified. If some of the q possible categories are not used, then it is strobgly advised to specify the complete list of possible categories as a vector in parametr categ.labels. Otherwise, the program may not work.
categ.labels	An optional vector parameter containing the list of all possible ratings. It may be useful in case some of the possibe ratings are not used by any rater, they will still be used when calculating agreement coefficients. The default value is NULL. In this case, only categories reported by the raters are used in the calculations.
conflev	An optional parameter representing the confidence level associated with the confidence interval. Its default value is 0.95.
N	An optional parameter representing the population size (if any). It may be use to perform the final population correction to the variance. Its default value is infinity.

Value

A data list containing 3 objects: (1) a one-row data frame containing the estimates, (2) the weight matrix used in the calculations, and (3) the categories used in the analysis. The data frame of estimates contains the following variables "coeff.name" (coefficient name), "pa" (the percent agreement), "pe" (percent chance-agreement-always equals 0), "coeff.val" (agreement coefficient = pa), coeff.se (the percent agreement standard error), "conf.int" (the percent agreement confidence interval), "p.value"(the percent agreement p-value), "w.name"(the weights' identification).

Examples

```
#The dataset "cac.raw4raters" comes with this package. Analyze it as follows:
cac.raw4raters
pa.coeff.raw(cac.raw4raters) #Percent agreement, precision measures, weights & categories
pa.coeff.raw(cac.raw4raters)$est #Yields percent agreement with precision measures
pa <- pa.coeff.raw(cac.raw4raters)$est$coeff.val #Yields percent agreement alone.
pa
pa.coeff.raw(cac.raw4raters, weights = "quadratic") #weighted percent agreement/quadratic weights
```

---

pa2.table *Percent Agreement coefficient for 2 raters*

Description

Percent Agreement coefficient for 2 raters

Usage

```
pa2.table(
  ratings,
  weights = identity.weights(1:ncol(ratings)),
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	A square table of ratings (assume no missing ratings).
weights	An optional matrix that contains the weights used in the weighted analysis. By default, this parameter contains the identity weight matrix, which leads to the unweighted analysis.
conflev	An optional parameter that specifies the confidence level used for constructing confidence intervals. By default the function assumes the standard value of 95%.
N	An optional parameter representing the finite population size if any. It is used to perform the finite population correction to the standard error. It's default value is infinity.

Value

A data frame containing the following 5 variables: coeff.name coeff.val coeff.se coeff.ci coeff.pval.

Examples

```
#The dataset "cont3x3abstractors" comes with this package. Analyze it as follows:
pa2.table(cont3x3abstractors) #Yields percent agreement along with precision measures
pa <- pa2.table(cont3x3abstractors)$coeff.val #Yields percent agreement alone.
pa
q <- nrow(cont3x3abstractors) #Number of categories
pa2.table(cont3x3abstractors,weights = quadratic.weights(1:q)) #Weighted percent agreement
```

---

quadratic.weights *Function for computing the Quadratic Weights*

Description

Function for computing the Quadratic Weights

Usage

```
quadratic.weights(categ)
```

Arguments

categ	A mandatory parameter representing the vector of all possible ratings.

Value

A square matrix of quadratic weights to be used for calculating the weighted coefficients.

---

radical.weights *Function for computing the Radical Weights*

Description

Function for computing the Radical Weights

Usage

```
radical.weights(categ)
```

Arguments

categ	A mandatory parameter representing the vector of all possible ratings.

Value

A square matrix of quadratic weights to be used for calculating the weighted coefficients.

---

ratio.weights *Function for computing the Ratio Weights*

Description

Function for computing the Ratio Weights

Usage

```
ratio.weights(categ)
```

Arguments

categ	A mandatory parameter representing the vector of all possible ratings.

Value

A square matrix of quadratic weights to be used for calculating the weighted coefficients.

scott2.table *Scott's coefficient for 2 raters*

Description

Scott's coefficient for 2 raters

Usage

```
scott2.table(
  ratings,
  weights = identity.weights(1:ncol(ratings)),
  conflev = 0.95,
  N = Inf
)
```

Arguments

ratings	A square table of ratings (assume no missing ratings).
weights	An optional matrix that contains the weights used in the weighted analysis. By default, this parameter contaings the identity weight matrix, which leads to the unweighted analysis.
conflev	An optional parameter that specifies the confidence level used for constructing confidence intervals. By default the function assumes the standard value of 95%.
N	An optional parameter representing the finite population size if any. It is used to perform the finite population correction to the standard error. It's default value is infinity.

Value

A data frame containing the following 5 variables: coeff.name coeff.val coeff.se coeff.ci coeff.pval.

Examples

```
#The dataset "cont3x3abstractors" comes with this package. Analyze it as follows:
scott2.table(cont3x3abstractors) #Yields Scott's Pi coefficient along with precision measures
scott <- scott2.table(cont3x3abstractors)$coeff.val #Yields Scott's coefficient alone.
scott
q <- nrow(cont3x3abstractors) #Number of categories
scott2.table(cont3x3abstractors,weights = quadratic.weights(1:q)) #weighted Scott's coefficient
```

trim — *An r function for trimming leading and trealing blanks*

Description

An r function for trimming leading and trealing blanks

Usage

```
trim(x)
```

Arguments

x	is a string variable.

Value

A string variable where leading and trealing blanks are trimmed.

x.dist10x5 — *Dataset of categorical ratings assigned to 10 subjects and presented in the form of a distribution of 4 raters by subject and category*

Description

This dataset shows how 4 raters classified 10 subjects into 5 categories labeled as q1, q2, q3, q4 and q5. Each record is associated with a subject and shows how the 4 raters are distributed across the 5 categories they assigned the subject to.

Usage

```
x.dist10x5
```

Format

A data frame of 10 rows and 6 columns. integers.

subject Patient's identifier

q1 Number of raters who classified the subject into category q1

q2 Number of raters who classified the subject into category q2

q3 Number of raters who classified the subject into category q3

q4 Number of raters who classified the subject into category q4

q5 Number of raters who classified the subject into category q5

Source

K. Gwet, PhD.

x.dist6x5psy — *Dataset showing how 6 psychiatrists classified 15 patients by their mental health condition.*

Description

Dataset of 15 psychiatric patients where each row identifies a patient and shows how the 6 psychiatrists distrubte across 5 mental conditions. These 5 conditions are Depression, personality disorder, Schizophrenia, Neurosis, and Other.

Usage

```
x.dist6x5psy
```

Format

A data frame of 15 rows and 6 columns. integers.

Subject Patient's identifier

Depression Number of raters to have diagnosed the patient with depression

Personality.Disorder Number of raters to have diagnosed the patient with personality disorder

Schizophrenia Number of raters to have diagnosed the patient with schizophrenia

Neurosis Number of raters to have diagnosed the patient with neurosis

Other Number of raters to have diagnosed the patient with other condition

Source

K. Gwet, PhD.

x.raw10x4 — *Raw categorical ratings assigned to 10 subjects by 4 raters*

Description

Dataset of 10 subjects and the categorical ratings assigned to them by 4 raters. The 5 categories available to raters are labeled as 1, 2, 3, 4, 5.

Usage

```
x.raw10x4
```

Format

A data frame of 10 rows and 5 columns. integers.

subject Patient's identifier

rater1 Category into which rater1 classified the subject

rater2 Category into which rater2 classified the subject

rater3 Category into which rater3 classified the subject

rater4 Category into which rater4 classified the subject

Source

K. Gwet, PhD.

---

x.raw12x4 *This dataset contains raw categorical ratings that 4 raters assigned to 12 subjects.*

Description

Dataset of raw ratings assigned to 12 units by 4 raters. Each row is associated with a unit identifier and each column to a rater. The columns are named Rater1, Rater2, Rater3 and Rater4 and contain the categories into which the unit was assigned. Category values are 1, 2, 3, 4, 5 and some ratings are missing.

Usage

```
x.raw12x4
```

Format

A data frame of 12 rows and 5 columns containing integer values 1, 2, 3, 4 and 5. Missing data points are represented with a dot ("."). integers.

Units Patient's identifier

Rater1 Category into which Rater1 classified the unit

Rater2 Category into which Rater2 classified the unit

Rater3 Category into which Rater3 classified the unit

Rater4 Category into which Rater4 classified the unit

Source

K. Gwet, PhD.

Index