Package 'goweragreement'

February 2, 2024

Type Package

Title Bayesian Gower Agreement for Categorical Data

Version 1.0

Date 2024-02-01

Depends R (>= 3.5.0)

Suggests testthat

Description Provides tools for applying the Bayesian Gower agreement methodology (pre-sented in the package vignette) to nominal or ordinal data. The framework can accommo-date any number of units, any number of coders, and missingness; and can handle both one-way and two-way random study designs. Influential units and/or coders can be identified eas-ily using leave-one-out statistics.

License GPL (>= 2)

URL <http://www.johnhughes.org>

RoxygenNote 7.2.3

Encoding UTF-8

NeedsCompilation no

Author John Hughes [aut, cre]

Maintainer John Hughes <drjphughesjr@gmail.com>

Repository CRAN

Date/Publication 2024-02-02 19:50:12 UTC

R topics documented:

---

confint.gower   *Compute a credible interval for a Bayesian Gower fit.*

Description

Compute a credible interval for a Bayesian Gower fit.

Usage

```
## S3 method for class 'gower'
confint(object, parm = "mu", level = 0.95, ...)
```

Arguments

object	an object of class "gower", the result of a call to gower.agree.
parm	always ignored since there is only one parameter, mu.
level	the desired confidence level for the interval. The default is 0.95.
...	additional arguments. These are passed to quantile.

Details

This function computes a credible interval for mu, the agreement parameter.

This function uses the bootstrap sample to compute a credible interval. If type = "HPD" is not an element of . . . , the lower and upper limits of the interval are appropriate quantiles of the bootstrap sample. Otherwise the limits are for an HPD interval.

Value

A vector with entries giving the lower and upper limits of the interval. These will be labelled as (1 - level) / 2 and 1 - (1 - level) / 2 if the quantile method was used. They will be labeled "Lower" and "Upper" if the interval is an HPD interval.

See Also

gower.agree

Examples

```
# Fit the liver data, using the mean distance for each row of the data matrix.
# The range (which is equal to 4) must be passed to \code{\link{gower.agree}}
# since these data are ordinal and the L1 distance function is used. We assume
# a one-way sampling design for these data, i.e., units are random and coders
# are fixed. Also compute a 95\% credible interval using the quantile method
# and then the HPD method.
```

```
data(liver)
liver = as.matrix(liver)
fit = gower.agree(liver, data.type = "ordinal", range = 4)
confint(fit)
confint(fit, type = "HPD")
```

---

discrete.dist *Apply the discrete metric to two scores.*

Description

Apply the discrete metric to two scores.

Usage

```
discrete.dist(x, y)
```

Arguments

x   a score.

y   a score.

Details

This function applies the discrete metric to two scores. This is an appropriate distance function for nominal data.

Value

0 if x is equal to y, 1 if x is not equal to y, NA if either score is NA.

See Also

L1.dist

---

gower.agree   *Apply the Bayesian Gower agreement methodology to nominal or or-*
                                dinal data.

Description

Apply the Bayesian Gower agreement methodology to nominal or ordinal data.

Usage

```
gower.agree(
  data,
  data.type = c("nominal", "ordinal"),
  dist.type = c("mean", "max"),
  design = c("one-way", "two-way"),
  iter = 10000,
  ...
)
```

Arguments

data	a matrix of scores. Each row corresponds to a unit, each column to a coder.
data.type	the type of scores to be analyzed, either "nominal" or "ordinal".
dist.type	for ordinal data, whether the row statistics are computed using the mean of the pairwise distances or the maximum pairwise distance.
design	the sampling design, either "one-way" or "two-way". For the former, the units are random and the coders are fixed. For the latter, both the units and the coders are random.
iter	the desired size of the posterior sample. The default value is 10,000.
...	additional arguments for the distance function. These are ignored for nominal data. For ordinal data the range of the scores must be provided via argument range.

Details

This is the package's flagship function. It applies the Bayesian Gower methodology to nominal or ordinal data, and provides an estimate of the posterior mean along with a credible interval.

Value

Function gower.agree returns an object of class "gower", which is a list comprising the following elements.

mu.hat	the estimate of the posterior mean.
mu.sample	the posterior sample.
call	the matched call.

units	the number of units.
coders	the number of coders.
data	the data matrix, sans rows that were removed due to missigness.
data.type	the type of scores, nominal or ordinal.
dist.type	for ordinal data, the manner in which the row statistics were computed.
design	the sampling design, one-way or two-way.
row.stats	the vector of row statistics.
del	the number of rows that were deleted due to missingness.

Examples

```
# Fit the liver data, using the mean distance for each row of the data matrix.
# The range (which is equal to 4) must be passed to \code{\link{gower.agree}}
# since these data are ordinal and the L1 distance function is used. We assume
# a one-way sampling design for these data, i.e., units are random and coders
# are fixed.

data(liver)
liver = as.matrix(liver)
fit = gower.agree(liver, data.type = "ordinal", range = 4)
summary(fit)
```

---

hpd	*Compute a highest posterior density (HPD) interval.*

Description

Compute a highest posterior density (HPD) interval.

Usage

```
hpd(x, alpha = 0.05)
```

Arguments

x	the posterior sample.
alpha	the desired significance level.

Details

This function uses a given posterior sample to compute an HPD interval at a given significance level.

Value

A 2-vector containing the lower endpoint and the upper endpoint, respectively.

---

influence.gower *Compute DFBETAs for units and/or coders.*

Description

Compute DFBETAs for units and/or coders.

Usage

```
## S3 method for class 'gower'
influence(model, units, coders, ...)
```

Arguments

model	a fitted model object, the result of a call to [gower.agree]().
units	a vector of integers. A DFBETA will be computed for each of the corresponding units.
coders	a vector of integers. A DFBETA will be computed for each of the corresponding coders.
...	additional arguments. These are ignored.

Details

This function computes DFBETAs for one or more units and/or one or more coders.

Value

A list comprising at most four elements.

dfbeta.units	a vector containing DFBETAs for the units specified via argument units.
dfbeta.coders	a vector containing DFBETAs for the coders specified via argument coders.
fits.units	a list containing fit objects for the omitted units specified via argument units.
fits.coders	a list containing fit objects for the omitted coders specified via argument coders.

Examples

```
# Analyze nominal data previously considered by Krippendorff.
# Assume a one-way design. Compute a DFBETA for unit 6, which
# should be rather influential.

kripp = matrix(c(1,2,3,3,2,1,4,1,2,NA,NA,NA,
                 1,2,3,3,2,2,4,1,2,5,NA,3,
                 NA,3,3,3,2,3,4,2,2,5,1,NA,
                 1,2,3,3,2,4,4,1,2,5,1,NA), 12, 4)
kripp
```

```
set.seed(12)
fit = gower.agree(kripp)
summary(fit)
inf = influence(fit, units = 6)

# Report the DFBETA for unit 6 and the estimate of mu when unit 6
# is ommitted.

inf$dfbeta.units
fit$mu.hat - inf$dfbeta.units
```

L1.dist   *Apply the L1 distance function to two scores.*

Description

Apply the L1 distance function to two scores.

Usage

```
L1.dist(x, y, range)
```

Arguments

x	a score.
y	a score.
range	the range of the scores, i.e., the difference between the maximum score and the minimum score.

Details

This function applies the L1 distance function to two scores. This is an appropriate distance function for ordinal data.

Value

$|x - y|/r$, where r is the range of the scores; or NA if either score is NA.

See Also

discrete.dist

---

liver	*Ordinal data from a radiological study of congenital diaphragmatic hernia.*

Description

This data frame has exactly four columns. The first two columns contain liver herniation grades assigned by the first radiologist, and the final two columns contain grades assigned by the second radiologist. Each radiologist graded each image twice so that intra-rater consistency can be assessed. The liver grades and their meanings are provided below.

Usage

```
data(liver)
```

Format

A data frame having 47 rows and four columns

Details

- 1. No herniation of the liver into the fetal chest
- 2. Less than half of the ipsilateral thorax is occupied by the fetal liver
- 3. Greater than half of the thorax is occupied by the fetal liver
- 4. The liver dome reaches the thoracic apex
- 5. The liver dome not only reaches the thoracic apex but also extends across the thoracic midline

---

summary.gower	*Print a summary of a Bayesian Gower fit.*

Description

Print a summary of a Bayesian Gower fit.

Usage

```
## S3 method for class 'gower'
summary(object, conf.level = 0.95, digits = 4, ...)
```

Arguments

object	an object of class "gower", the result of a call to gower.agree.
conf.level	the confidence level for the credible interval. The default is 0.95.
digits	the number of significant digits to display. The default is 4.
...	additional arguments. These are currently ignored.

Details

This function prints a summary of the fit.

Value

A list containing various elements of the summary is returned invisibly, but this function should be called for its side effects.

See Also

gower.agree

Examples

```
# Fit the liver data, using the mean distance for each row of the data matrix.
# The range (which is equal to 4) must be passed to \code{\link{gower.agree}}
# since these data are ordinal and the L1 distance function is used. We assume
# a one-way sampling design for these data, i.e., units are random and coders
# are fixed.

data(liver)
liver = as.matrix(liver)
fit = gower.agree(liver, data.type = "ordinal", range = 4)
summary(fit)
```

Index