Package ‘qualvar’

January 3, 2018

Type Package
Title Implements Indices of Qualitative Variation Proposed by Wilcox (1973)
Version 0.2.0
URL http://joelgombin.github.io/qualvar/
BugReports https://www.github.com/joelgombin/qualvar/issues
Description Implements indices of qualitative variation proposed by Wilcox (1973).
License GPL-3
LazyData TRUE
Depends R (>= 2.10)
Imports stats
Suggests knitr,rmarkdown,plyr,tidyverse,ggplot2,GGally,DT
VignetteBuilder knitr,rmarkdown
RoxygenNote 6.0.1
NeedsCompilation no
Author Joel Gombin [aut, cre]
Maintainer Joel Gombin <joel.gombin@gmail.com>
Repository CRAN
Date/Publication 2018-01-03 00:12:05 UTC

R topics documented:

ADA .......................................................... 2
B ............................................................... 3
DM ............................................................. 4
HREL .......................................................... 5
MDA .......................................................... 6
VA ............................................................. 7
wilcox1973 .................................................. 8
**Description**

Computes the average deviation analog (ADA) for a vector of frequencies of categories.

**Usage**

ADA(x, na.rm = TRUE)

**Arguments**

- `x`: a vector of frequencies
- `na.rm`: if TRUE, missing values are removed. If FALSE, NA is returned if there is any NA value.

**Details**

According to Wilcox (1973, p. 328), the ADA is 'an analog of the average or mean deviation'. The formula for the ADA is:

\[ 1 - \frac{\sum_{i=1}^{k} |f_i - \frac{N}{K}|}{\frac{N}{K} (K - 1)} \]

**Value**

The value of the ADA statistics, which is normalised (varies between 0 and 1).

**References**


**Examples**

```r
x <- rmultinom(1, 100, rep_len(0.25, 4))
x <- as.vector(t(x))
ADA(x)

df <- rmultinom(10, 100, rep_len(0.25, 4))
df <- as.data.frame(t(df))
apply(df, 1, ADA)
```
Description

Computes the B index for a vector of frequencies of categories.

Usage

\[ B(x, \text{na.rm} = \text{TRUE}) \]

Arguments

- \( x \): a vector of frequencies
- \( \text{na.rm} \): if TRUE, missing values are removed. If FALSE, NA is returned if there is any NA value.

Details

According to Wilcox (1973, p. 330), and following Kaiser (1968), the B index relies on the geometric mean, but corrects it for undesirable properties. The formula for the B index is:

\[
1 - \sqrt{1 - \left( \frac{\prod_{i=1}^{k} f_i K}{N} \right)^2}
\]

Value

The value of the B statistics, which is normalised (varies between 0 and 1).

References


Examples

```r
x <- rmultinom(1, 100, rep_len(0.25, 4))
x <- as.vector(t(x))
B(x)

df <- rmultinom(10, 100, rep_len(0.25, 4))
df <- as.data.frame(t(df))
apply(df, 1, B)
```
**DM**

*Deviation from the mode (DM)*

**Description**

Computes the deviation from the mode for a vector of frequencies of categories.

**Usage**

`DM(x, na.rm = TRUE)`

**Arguments**

- `x`: a vector of frequencies
- `na.rm`: if TRUE, missing values are removed. If FALSE, NA is returned if there is any NA value.

**Details**

According to Wilcox (1973, p. 327), 'the measure can be thought of as an index of deviation from the modal frequency, analogous to the variance as a measure of deviation from the mean'. The formula for the DM is:

\[
1 - \frac{\sum_{i=1}^{k}(f_m - f_i)}{N(K - 1)}
\]

**Value**

The value of the DM statistics, which is normalised (varies between 0 and 1).

**References**


**Examples**

```r
x <- rmultinom(1, 100, rep_len(0.25, 4))
x <- as.vector(t(x))
names(x) <- c("a", "b", "c", "d")
DM(x)

df <- rmultinom(10, 100, rep_len(0.25, 4))
df <- as.data.frame(t(df))
names(df) <- c("a", "b", "c", "d")
apply(df, 1, DM)
```
Description
Computes the HREL index for a vector of frequencies of categories.

Usage
HREL(x, na.rm = TRUE)

Arguments
x a vector of frequencies
na.rm if TRUE, missing values are removed. If FALSE, NA is returned if there is any NA value.

Details
According to Wilcox (1973, p. 329), and following Senders (1958), the HREL is 'a measure originally developed by engineers for use in specifying the properties of communications channels. The rationale for HREL is presented in terms of guessing by Virginia Senders (supplementing the mode as best guess): "What we need is a measure of uncertainty, or of "poorness of a guess," which will be high when the number of alternative possibilities is high, and low when some of the possibilities are much more likely than others. One possible measure is the average number of questions we have to ask to specify the correct alternative'. The formula for the HREL is:

\[-\sum_{i=1}^{k} \frac{f_i}{N} \log_2 \frac{f_i}{N} \log_2 K\]

Value
The value of the HREL statistics, which is normalised (varies between 0 and 1).

References

Examples
x <- rmultinom(1, 100, rep_len(0.25, 4))
x <- as.vector(t(x))
HREL(x)

df <- rmultinom(10, 100, rep_len(0.25, 4))
df <- as.data.frame(t(df))
apply(df, 1, HREL)
Mean Difference Analog (MDA or MNDIF)

Description
Computes the mean difference analog (MDA or MNDIF) for a vector of frequencies of categories.

Usage
```r
MDA(x, na.rm = TRUE)
```

Arguments
- `x`: a vector of frequencies
- `na.rm`: if TRUE, missing values are removed. If FALSE, NA is returned if there is any NA value.

Details
According to Wilcox (1973, p. 328), the MDA is 'an analog of the mean difference, a measure of variation that is discussed and used much less frequently than the average deviation or the standard deviation. It is defined as "the average of the differences of all the possible pairs of variate-values, taken regardless of sign"'. The formula for the MDA is:

\[
1 - \frac{\sum_{i=1}^{k-1} \sum_{j=i+1}^{k} |f_i - f_j|}{N(K - 1)}
\]

Value
The value of the MDA statistics, which is normalised (varies between 0 and 1).

References

Examples
```r
x <- rmultinom(1, 100, rep_len(0.25, 4))
x <- as.vector(t(x))
MDA(x)

df <- rmultinom(10, 100, rep_len(0.25, 4))
df <- as.data.frame(t(df))
apply(df, 1, MDA)
```
VA

Variance Analog (VA)

Description
Computes the variance analog (VA) for a vector of frequencies of categories.

Usage
VA(x, na.rm = TRUE)

Arguments
- x: a vector of frequencies
- na.rm: if TRUE, missing values are removed. If FALSE, NA is returned if there is any NA value.

Details
According to Wilcox (1973, p. 329), the VA is 'based on the variance, which is defined as the arithmetic mean of the squared differences of each value from the mean'. The formula for the VA is:

$$1 - \frac{\sum_{i=1}^{k} (f_i - \frac{N}{K})^2}{\frac{N^2(K-1)}{K}}$$

Value
The value of the VA statistics, which is normalised (varies between 0 and 1).

References

Examples
```r
x <- rmultinom(1, 100, rep_len(0.25, 4))
x <- as.vector(t(x))
VA(x)

df <- rmultinom(10, 100, rep_len(0.25, 4))
df <- as.data.frame(t(df))
apply(df, 1, VA)
```
Description

A dataset containing the returns of the 1968 US presidential election, by states.

Usage

wilcox1973

Format

a data_frame with 52 rows and 4 variables.

Source

Index

*Topic datasets
   wilcox1973, 8

ADA, 2
B, 3
DM, 4
HREL, 5
MDA, 6
VA, 7
wilcox1973, 8