Package ‘GWI’

February 26, 2021

Type Package

Title Count and Continuous Generalized Variability Indexes

Version 1.0.2

Date 2021-02-25

Author Aboubacar Y. Touré and Célestin C. Kokonendji

Maintainer Aboubacar Y. Touré <aboubacaryacoubatoure.ussgb@gmail.com>

Description Firstly, both functions of the univariate Poisson dispersion index (DI) for count data and the univariate exponential variation index (VI) for nonnegative continuous data are performed. Next, other functions of univariate indexes such the binomial dispersion index (DIb), the negative binomial dispersion index (DInb) and the inverse Gaussian variation index (VIiG) are given. Finally, we are computed some multivariate versions of these functions such that the generalized dispersion index (GDI) with its marginal one (MDI) and the generalized variation index (GVI) with its marginal one (MVI) too.

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2021-02-26 00:00:05 UTC

R topics documented:

`GWI-package` *Count and continuous generalized variability indexes*

Description

Univariate Poisson dispersion index `di.fun`, univariate exponential variation index `vi.fun` functions are performed. Next, the univariate binomial dispersion index `dib.fun`, the univariate negative binomial dispersion index `dinb.fun` and the univariate inverse Gaussian variation index `viiG.fun` functions are given. Finally, the generalized dispersion index and its marginal one `gmdi.fun` , the generalized variation index and its marginal one `gmvi.fun` functions are displayed.

Details

The univariate Poisson dispersion index (DI) and its relative versions with respect to binomial and negative binomial

The Poisson dispersion phenomenon is well-known and very widely used in practice; see, e.g., Kokonendji (2014) for a review of count (or discrete integer-valued) models. There are many interpretable mechanisms leading to this phenomenon which makes it possible to classify count distributions and make inference; see, e.g., Mizère et al. (2006) and Touré et al. (2020) for approximative statistical tests. Introduced from Fisher (1934), the Poisson dispersion index, also called the Fisher dispersion index, of a count random variable X on $S = \{0, 1, 2, \ldots\} =: N_0$ can be defined as

$$DI(X) = \frac{VarX}{EX},$$

the ratio of variance to mean. In fact, the positive quantity $DI(X)$ is the ratio of two variances since EX is the expected variance under the Poisson distribution. Hence, one easily deduces the concept of the relative dispersion index (denoted by RDI) by choosing another reference than the Poisson distribution. Indeed, if X and Y are two count random variables on the same support $S \subseteq N_0$ such that $EX = EY$, then

$$RDI_Y(X) := \frac{VarX}{VarY} = \frac{DI(X)}{DI(Y)} >=< 1;$$

i.e. X is over-, equi- and under-dispersed compared to Y if $VarX > VarY$, $VarX = VarY$ and $VarX < VarY$, respectively.

For instance, the binomial dispersion index is defined as

$$RDI_B(X) = \frac{varX}{EX(1 - EX/N)},$$

where $N \in \{1, 2, \ldots\}$ is the fixed number of trials. Also, the negative binomial dispersion index is defined as

$$RDI_N B(X) = \frac{varX}{EX(1 + EX/\lambda)},$$

where $\lambda > 0$ is the dispersion parameter. See also, Weiss (2018, page 15) and Abid et al. (2021) for more details.

The univariate variation index (VI) and its relative version with respect to inverse Gaussian distribution: More recently, Abid et al. (2020) have introduced the exponential variation index for positive continuous random variable X on $[0,\infty)$ as

$$VI(X) = \frac{VarX}{(EX)^2}.$$

It can be viewed as the squared coefficient of variation. It is used in the framework of reliability to discriminate distribution of increasing/decreasing failure rate on the average (IFRA/DFRA); see, e.g., Barlow and Proschan (1981) in the sense of the coefficient of variation. See also Touré et al. (2020) for more details. Following RDI, the relative variation index (RVI) is defined, for two continuous random variables X and Y on the same support $S = [0,\infty)$ with $EX = EY$, by

$$RVI_Y(X) := \frac{VarX}{VarY} = \frac{VI(X)}{VI(Y)} >=< 1;$$

i.e. X is over-, equi- and under-varied compared to Y if $VarX > VarY$, $VarX = VarY$ and $VarX < VarY$, respectively. For instance, the inverse Gaussian variation index is defined as

$$RVI_IG(X) = \lambda^2 \frac{varX}{(EX)^3},$$

where $\lambda > 0$ is the shape parameter.

Next, consider the following notations. Let $Y = (Y_1, \ldots, Y_k)^\top$ be a nondegenerate count or continuous k-variate random vector, $k \geq 1$. Let also EY be the mean vector of Y and $covY$= $(cov(Y_i, Y_j))_{i,j\in\{1,\ldots,k\}}$ the covariance matrix of Y.

The generalized dispersion index (GDI) and marginal dispersion index (MVI): Kokonendji and Puig (2018) have introduced the *generalized dispersion index* for count vector Y on $\{0, 1, 2, \ldots\}^k$ by

$$GDI(Y) = \frac{\sqrt{EY}^\top (covY)\sqrt{EY}}{EY^\top EY}.$$

Note that when $k = 1$, $GDI(Y)$ is just the classical Fisher dispersion index DI. $GDI(Y)$ makes it possible to compare the full variability of Y (in the numerator) with respect to its expected uncorrelated Poissonian variability (in the denominator) which depends only on EY. $GDI(Y)$ takes into account the correlation between variables. For only taking into account the dispersion information coming from the margins, the authors defined the "marginal dispersion index":

$$MDI(Y) = \frac{\sqrt{EY}^\top (diagvarY)\sqrt{EY}}{EY^\top EY} = \sum_{j=1}^{k} \frac{\{E(Y_j)\}^2}{EY^\top EY} DI(Y_j).$$

The generalized variation index (GVI) and marginal variation index (MVI): Similarly, Kokonendji et al. (2020) defined the *generalized variation index* for positive continuous vector Y on $[0,\infty)^k$ by

$$GVI(Y) = \frac{EY^\top (covY)EY}{(EY^\top EY)^2}.$$

Remark that when $k = 1$, $GVI(Y)$ is the univariate variation index VI. $GVI(Y)$ makes it possible to compare the full variability of Y (in the numerator) with respect to its expected uncorrelated exponential variability (in the denominator) which depends only on EY. Also, $GVI(Y)$ takes into account the correlation between variables. To only take into account the variation information coming from the margins, Kokonendji et al. (2020) defined the "marginal variation index":

$$MVI(Y) = \frac{EY^\top (diagvarY) EY}{(EY^\top EY)^2} = \sum_{j=1}^{k} \frac{(EY_j)^4}{(EY^\top EY)^2} VI(Y_j).$$

Author(s)

Aboubacar Y. Touré and Célestin C. Kokonendji

Maintainer: Aboubacar Y. Touré <aboubacaryacoubatoure.ussgb@gmail.com>

References

Abid, R., Kokonendji, C.C. and Masmoudi, A. (2020). Geometric Tweedie regression models for continuous and semicontinuous data with variation phenomenon, *AStA Advances in Statistical Analysis* **104**, 33-58.

Abid, R.,Kokonendji, C.C. and Masmoudi, A. (2021). On Poisson-exponential-Tweedie models for ultra-overdispersed count data, *AStA Advances in Statistical Analysis* **105**, 1-23.

Barlow, R.A. and Proschan, F. (1981). Statistical Theory of Reliability and Life Testing : Probability Models, *Silver Springs*, Maryland.

Fisher, R.A. (1934). The effects of methods of ascertainment upon the estimation of frequencies, *Annals of Eugenics* **6**, 13-25.

Kokonendji, C.C., Over- and underdispersion models. In: N. Balakrishnan (Ed.) The Wiley Encyclopedia of Clinical Trials- Methods and Applications of Statistics in Clinical Trials, **Vol.2** (Chap.30), pp. 506-526. *Wiley*, New York (2014).

Kokonendji, C.C. and Puig, P. (2018). Fisher dispersion index for multivariate count distributions : A review and a new proposal, *Journal of Multivariate Analysis* **165**, 180-193.

Kokonendji, C.C., Touré, A.Y. and Sawadogo, A. (2020). Relative variation indexes for multivariate continuous distributions on $[0,\infty)^k$ and extensions, *AStA Advances in Statistical Analysis* **104**, 285-307.

Mizère, D., Kokonendji, C.C. and Dossou-Gbété, S. (2006). Quelques tests de la loi de Poisson contre des alternatives géenérales basées sur l'indice de dispersion de Fisher, *Revue de Statistique Appliquée* **54**, 61-84.

Touré, A.Y., Dossou-Gbété, S. and Kokonendji, C.C. (2020). Asymptotic normality of the test statistics for relative dispersion and relative variation indexes, *Journal of Applied Statistics* **47**, 2479-2491.

Weiss, C.H. (2018). An Introduction to Discrete-Valued Times Series. *Wiley*, Hoboken NJ.

di.fun   *Function for DI*

Description

The function computes the univariate Poisson dispersion index for a count random variable.

Usage

```
di.fun(X)
```

Arguments

X	A count random variable

Details

`di.fun` provides the univariate Poisson dispersion index (Fisher, 1934). We can refer to Touré et al. (2020) for more details on the Poisson dispersion index.

Value

Returns

di	The Poisson dispersion index

Author(s)

Aboubacar Y. Touré and Célestin C. Kokonendji

References

Fisher, R.A. (1934). The effects of methods of ascertainment upon the estimation of frequencies, *Annals of Eugenics* **6**, 13-25.

Touré, A.Y., Dossou-Gbété, S. and Kokonendji, C.C. (2020). Asymptotic normality of the test statistics for relative dispersion and relative variation indexes, *Journal of Applied Statistics* **47**, 2479-2491.

Examples

```
X<-c(6,7,8,9,8,4,7,6,12,8,0)
di.fun(X)
T<-c(61,72,83,94,85,46,77,68,129,80,10,12,12,3,4,5)
di.fun(T)
```

`dib.fun` *Function for DIb*

Description

The function computes the binomial dispersion index for a given number of trials $N \in \{1, 2, \ldots\}$.

Usage

```
dib.fun(X, N)
```

Arguments

`X`	A count random variable
`N`	The number of trials of binomial distribution

Details

`dib.fun` computes the dispersion index with respect to the binomial distribution. See Touré et al. (2020) and Weiss (2018) for more details.

Value

Returns

`dib`	The binomial dispersion index

Author(s)

Aboubacar Y. Touré and Célestin C. Kokonendji

References

Touré, A.Y., Dossou-Gbété, S. and Kokonendji, C.C. (2020). Asymptotic normality of the test statistics for relative dispersion and relative variation indexes, *Journal of Applied Statistics* **47**, 2479-2491.

Weiss, C.H. (2018). An Introduction to Discrete-Valued Times Series. *Wiley*, Hoboken NJ.

Examples

```
X<-c(12,9,0,8,5,7,6,5,3,4,9,4)
dib.fun(X,12)
Y<-c(0,0,1,1,0,1,1)
dib.fun(Y,7)
```

dinb.fun *Function for DInb*

Description

The function computes the negative binomial dispersion index for a given dispersion parameter $l \in (0, \infty)$.

Usage

```
dinb.fun(X, l)
```

Arguments

`X`	A count random variable
`l`	The dispersion parameter of negative binomial distribution

Details

`dinb.fun` computes the dispersion index with respect to negative binomial distribution. See Touré et al. (2020) and Abid et al. (2021) for more details.

Value

Returns

`dinb`	The negative binomial dispersion index

Author(s)

Aboubacar Y. Touré and Célestin C. Kokonendji

References

Abid, R.,Kokonendji, C.C. and Masmoudi, A. (2021). On Poisson-exponential-Tweedie models for ultra-overdispersed count data, *AStA Advances in Statistical Analysis* **105**, 1-23.

Touré, A.Y., Dossou-Gbété, S. and Kokonendji, C.C. (2020). Asymptotic normality of the test statistics for relative dispersion and relative variation indexes, *Journal of Applied Statistics* **47**, 2479-2491.

Examples

```
X<-c(12,9,0,8,5,7,6,5,3,4,9,4)
dinb.fun(X,12)
Y<-c(0,6,1,3,4,2,5)
dinb.fun(Y,7)
```

`gmdi.fun` *Function for GDI and MDI*

Description

The function computes the GDI and MDI indexes for multivariate count data.

Usage

```
gmdi.fun(Y)
```

Arguments

`Y`	A matrix of count random variables

Details

`gmdi.fun` computes GDI and MDI indexes introduced by Kokonendji and Puig (2018).

Value

Returns:

`gdi`	The generalized dispersion index
`mdi`	The marginal dispersion index

Author(s)

Aboubacar Y. Touré and Célestin C. Kokonendji

References

Kokonendji, C.C. and Puig, P. (2018). Fisher dispersion index for multivariate count distributions : A review and a new proposal, *Journal of Multivariate Analysis* **165**, 180-193.

Examples

```
Y<-cbind(c(1,2,3,4,5,6,7,8),c(1,2,3,4,5,6,7,8))
gmdi.fun(Y)
Z<-cbind(c(1,2,3,4,5,6,7,8),c(1,2,3,4,5,6,7,8),c(1,2,3,4,5,6,7,8),c(1,2,3,4,5,6,7,8))
gmdi.fun(Z)
```

gmvi.fun *Function for GVI and MVI*

Description

The function computes GVI and MVI indexes for multivariate positive continuous data.

Usage

```
gmvi.fun(Y)
```

Arguments

Y	A matrix of positive continuous random variables

Details

`gmvi.fun` computes the GVI and MVI indexes defined in Kokonendji et al. (2020).

Value

Returns:

gvi	The generalized variation index
mvi	The marginal variation index

Author(s)

Aboubacar Y. Touré and Célestin C. Kokonendji

References

Kokonendji, C.C., Touré, A.Y. and Sawadogo, A. (2020). Relative variation indexes for multivariate continuous distributions on $[0,\infty)^k$ and extensions, *AStA Advances in Statistical Analysis* **104**, 285-307.

Examples

```
Y<-cbind(c(2.3 ,26.1 ,8.7 ,10.9 ,1.2,1.4),c(9.7 ,7.3,9.3 ,9.4 ,10.5 ,9.8))
gmvi.fun(Y)
Z<-cbind(c(2.3 ,26.1 ,8.7),c(9.7 ,7.3,9.3),c(9.7 ,7.3,9.3),c(9.7 ,7.3,9.3))
gmvi.fun(Z)
```

`vi.fun` *Function for VI*

Description

The function calculates the univariate exponential variation index for a positive continuous random variable.

Usage

```
vi.fun(X)
```

Arguments

`X`	A positive continuous random variable

Details

`vi.fun` computes the univariate exponential variation index defined by Abid et al. (2020). See also Touré et al. (2020) for more details on this index.

Value

Returns

`vi`	The exponential variation index

Author(s)

Aboubacar Y. Touré and Célestin C. Kokonendji

References

Abid, R., Kokonendji, C.C. and Masmoudi, A. (2020). Geometric Tweedie regression models for continuous and semicontinuous data with variation phenomenon, *AStA Advances in Statistical Analysis* **104**, 33-58.

Touré, A.Y., Dossou-Gbété, S. and Kokonendji, C.C. (2020). Asymptotic normality of the test statistics for relative dispersion and relative variation indexes, *Journal of Applied Statistics* **47**, 2479-2491.

Examples

```
X<-c(6.23,7.02,8.94,9.56,8.01,4.34,7.44,6.66,12.72,8.34,0)
vi.fun(X)
T<-c(6.231,7.022,8.943,9.789,8.014,4.423)
vi.fun(T)
```

viiG.fun *Function for VIiG*

Description

The function computes the inverse Gaussian variation index with shape parameter $l \in (0, \infty)$.

Usage

```
viiG.fun(X, l)
```

Arguments

X	A positive continuous random variable
l	The shape parameter of the inverse Gaussian distribution

Details

`viiG.fun` computes the variation index with respect to the inverse Gaussian distribution. See Touré et al. (2020) for more details.

Value

Returns

viiG	The inverse Gaussian variation index

Author(s)

Aboubacar Y. Touré and Célestin C. Kokonendji

References

Touré, A.Y., Dossou-Gbété, S. and Kokonendji, C.C. (2020). Asymptotic normality of the test statistics for relative dispersion and relative variation indexes, *Journal of Applied Statistics* **47**, 2479-2491.

Examples

```
X<-c(0.12,9.11,0.03,8.71,5.02,7.12,6.42,5.73)
viiG.fun(X,0.05)
Y<-c(0.003,6.283,1.001,3.112,4.342,2.890,5.005)
viiG.fun(Y,0.3)
```

Index