

Package ‘p2distance’

February 20, 2015

Type Package

Title Welfare's Synthetic Indicator

Version 1.0.1

Date 2012-05-2

Author A.J. Perez-Luque; R. Moreno; R. Perez-Perez and F.J. Bonet

Maintainer R. Perez-Perez <ramon@ugr.es>

Description The welfare's synthetic indicator provides an ideal tool for measuring multi-dimensional concepts such as welfare, development, living standards, etc. It enables information from the various indicators to be aggregated into a single synthetic measure.

License GPL

LazyLoad yes

Repository CRAN

Date/Publication 2012-08-09 10:27:55

NeedsCompilation no

R topics documented:

p2distance-package	2
loadCSVtoP2distance	3
makeReferenceVector	4
p2distance	5
welfare	8

Index	10
--------------	-----------

p2distance-package *Welfare's Synthetic Indicator* ~-~ p2distance ~-~

Description

The welfare's synthetic indicator (see references) provides an ideal tool for measuring multi-dimensional concepts such as welfare, development, living standards, etc. It enables information from the various indicators to be aggregated into a single synthetic measure. This synthetic indicator, also called P_2 distance, overcomes several limitations of others traditional approach as Principal Components Analysis (PCA), Data Envelopment Analysis (DEA), etc. and offers some good statistical properties and allows inter-spatial and inter-temporal comparisons.

Details

Package:	p2distance
Type:	Package
Version:	1.0.1
Date:	2012-05-02
License:	GPL
LazyLoad:	yes

~-~ p2distance is the main function ~-~

Author(s)

A.J. Perez-Luque; R. Moreno; R. Perez-Perez and F.J. Bonet

References

- Pena, J. B. (1977). *Problemas de la medición del bienestar y conceptos afines (una aplicación al caso Español)*. Madrid: INE.
- Pena, J. B. (2009). La medición del bienestar social: una revisión crítica. *Estudios de Economía Aplicada*, **27**(2), 299–324.
- Somarriba, N. (2008). *Aproximación a la medición de la calidad de vida en la Unión Europea*. Doctoral Thesis. University of Valladolid. <http://www.eumed.net/tesis/2010/mnsa/index.htm>
- Zarzosa, P. (1992). *Aproximación a la medición del bienestar social, estudio de la idoneidad del indicador sintético Distancia P2*. Doctoral Thesis. University of Valladolid.
- Zarzosa, P. (1996). *Aproximación a la medición del Bienestar social*. Valladolid: University of Valladolid. 248 pp.

loadCSVtoP2distance *Load CSV file to matrix object*

Description

Load a CSV file to matrix object.

Usage

```
loadCSVtoP2distance(path, header=TRUE, sep="\t", dec=".", quote="\\"",  
  na.strings="NA", fileEncoding = "", encoding = "unknown")
```

Arguments

path	the path of the file which the data are to be read from. Each row of the table appears as one line of the file.
header	a logical value indicating whether the file contains the names of the variables as its first line. If missing, the value is determined from the file format: header is set to TRUE if and only if the first row contains one fewer field than the number of columns.
sep	the field separator character. Values on each line of the file are separated by this character.
dec	the character used in the file for decimal points.
quote	the set of quoting characters. To disable quoting altogether, use quote = "".
na.strings	a character vector of strings which are to be interpreted as NA values. Blank fields are also considered to be missing values in logical, integer, numeric and complex fields.
fileEncoding	character string: if non-empty declares the encoding used on a file (not a connection) so the character data can be re-encoded. See the 'Encoding' section of the help for file, the 'R Data Import/Export Manual' and 'Note'.
encoding	encoding to be assumed for input strings. It is used to mark character strings as known to be in Latin-1 or UTF-8 (see Encoding): it is not used to re-encode the input, but allows R to handle encoded strings in their native encoding (if one of those two).

Details

This function return a matrix object ready for p2distance function. It read a text file and it use the first column to give a row name. It uses the read.table function to read the file.

Value

A matrix object containing the data of csv file

Author(s)

A.J. Perez-Luque; R. Moreno; R. Perez-Perez and F.J. Bonet

makeReferenceVector *Make a reference vector*

Description

Make a reference vector.

Usage

```
makeReferenceVector(X, reference_vector_function = min)
```

Arguments

X Array with n rows (entities) and m columns (values of variables)

reference_vector_function
Function to make the reference vector. Common functions used are: min, max, mean, median, etc. See `apply` for further details. Minimum es the default function

Details

A reference vector must be defined for each partial indicator so as to compare different spatial entities. This vector is used by `p2distance` function (as base reference) to calculate distances of each spatial entities with this reference base. It is quite common to consider the minimum value as the base reference (see references).

Value

Vector with the reference value (base reference) of each variable. The vector length equals to number of variables

Author(s)

A.J. Perez-Luque; R. Moreno; R. Perez-Perez and F.J. Bonet

References

- Pena, J. B. (1977). *Problemas de la medición del bienestar y conceptos afines (una aplicación al caso Español)*. Madrid: INE.
- Pena, J. B. (2009). La medición del bienestar social: una revisión crítica. *Estudios de Economía Aplicada*, **27(2)**, 299–324.
- Somarriba, N. (2008). *Aproximación a la medición de la calidad de vida en la Unión Europea*. Doctoral Thesis. University of Valladolid. <http://www.eumed.net/tesis/2010/mnsa/index.htm>
- Zarzosa, P. (1992). *Aproximación a la medición del bienestar social, estudio de la idoneidad del indicador sintético Distancia P2*. Doctoral Thesis. University of Valladolid.

Examples

```
## Create a data frame of 3 variables (indicators) for 6 entities (rows)
dat <- data.frame(
  x1=c(10,12,13,14,12,11),
  x2=c(40,51,61,68,34,44),
  x3=c(0.43, 0.41, 0.39, 0.55, 0.60, 0.38))

## Calculate the reference vector (base reference) using minimum value of each variable
makeReferenceVector(dat, reference_vector_function=min)
```

p2distance

*Welfare's Synthetic Indicator function***Description**

This function calculates the P_2 distance synthetic indicator for a set of variables.

Usage

```
p2distance(matrix, reference_vector = NULL, reference_vector_function = min,
  iterations = 20, umbral = 1e-04)
```

Arguments

matrix	An object of matrix type with spatial entities in rows and variables in columns
reference_vector	Optionally. A reference vector defined for each partial indicator so as to compare different spatial entities
reference_vector_function	Optionally. Function to make the reference vector. Minimum es the default. Others common functions used: min, max, mean, median, etc. See makeReferenceVector for further details
iterations	Numbers of maximum iterations for the computational process until reach the convergence
umbral	The algorithm stop when the difference between iterations is lower than this umbral

Details

This is the main function on package. It calculates the Pena distance indicator, also called DP2, which is used to measure welfare in quality-of-life applications, to create Environmental Quality Indexes, etc. (see references). It is a multidimensional indicator capable to aggregate various partial indicators (variables) in a unique measure to compare the state of different spatial entities. The P2 Distance from a spatial entity r is defined as

$$DP_2 = \sum_{i=1}^n \left\{ \left(\frac{d_i}{\sigma_i} \right) (1 - R_{i,i-1,i-2,\dots,1}^2) \right\}$$

with $R_1^2 = 0$; where $d_i = |x_{ri} - x_{*i}|$ with the reference base $X_* = (x_{*1}, x_{*2}, \dots, x_{*n})$ where:

- n is the number of variables
- x_{ri} , is the value of the variable i in the spatial entity r
- σ_i is the standard deviation of variable i
- $R_{i,i-1,\dots,1}^2$ is the coefficient of determination in the regression of X_i over $X_{i-1}, X_{i-2}, \dots, X_1$ already included.

The numerical value of the DP2 index has no real meaning, but its is useful for comparing the state of different spatial entities in terms of welfare, environmental conditions, etc.

Value

discrimination.coefficient

Vector of discrimination coefficients (DC) for each variable. The value of DC, defined by Ivanovic (1974) is

$$DC_i = \frac{2}{m(m-1)} \sum_{j,l>j}^{k_i} m_{ji}m_{li} \left| \frac{x_{ji} - x_{li}}{\bar{X}_i} \right|$$

where m is the number of spatial entities and m_{ji} is the absolute frequency of x_{ji} . This measure ranges between 0 and 2. If a variable takes the same values for all spatial entities, DC equals zero, indicating zero discriminant power. By contrast, if a variable only has a value other than zero for one spatial entity and in the remainder $m - 1$, is equal to zero, DC reaches its maximum value (2) and the variable has full discriminant power (see Zarzosa, 1996; Zarzosa and Somarriba, 2012). There is an alternative way of calculating the coefficient, by using the Gini index,

$$DC_i = 2 \frac{m}{m-1} G$$

where m is the number of spatial entities and G the Gini index

p2distance

Vector with the last P_2 distance value for each spatial entity

p2distances

Array with vectors of P_2 distances values resulting for each iteration

diff_p2distances

Array with differences between two contiguous P_2 distances

iteration

Number of calculated iterations

umbral

Threshold in difference for two contiguous P_2 distances

variables_sort

Vector with the variable names by entrance order determined by last iteration

correction_factors

Correction Factors for each variable

cor.coeff

Correlation coefficient for each variable with the synthetic indicator (P_2 distance) calculated

partial.Indicators

For each spatial entity the difference between the reference vector and the value of each variable divided by the standard deviation. For a spatial entity, the sum of all partial indicators is the Frechet Distance (DF), which is the maximum value that P_2 distance can reach.

Author(s)

A.J. Perez-Luque; R. Moreno; R. Perez-Perez and F.J. Bonet

References

- Ivanovic, B. (1974) Comment établir une liste des indicateurs de developpement. *Revue de Statistique Appliquée*, **XXII(2)**, 37–50
- Montero, J. M., Chasco, C. and Larraz, B. (2010). Building an environmental quality index for a big city: a spatial interpolation approach combined with a distance indicator. *Journal of Geographical Systems*, **12**, 435–459.
- Pena, J. B. (1977). *Problemas de la medición del bienestar y conceptos afines (una aplicación al caso Español)*. Madrid: INE.
- Pena, J. B. (2009). La medición del bienestar social: una revisión crítica. *Estudios de Economía Aplicada*, **27(2)**, 299–324.
- Zarzosa, P. (1996). *Aproximación a la medición del Bienestar social*. Valladolid: University of Valladolid. 248 pp.
- Zarzosa, P. and Somarriba, N. (2012). An assessment of social welfare in Spain: Territorial analysis using a synthetic welfare indicator. *Social Indicators Research*, doi: <http://dx.doi.org/10.1007/s11205-012-0005-0>

Examples

```
## Calculate a welfare indicator for 27 countries of Europe
data(welfare)

## Convert welfare dataframe to matrix object
welfare <- as.matrix(welfare)

## Calculate P2 Distance
ind <- p2distance(matrix=welfare, reference_vector_function = min,
                 iterations = 20)

## Examine the results
# P2 distance
ind$p2distance

# Iterations to achieve convergence
ind$iteration

# Order of entry of variables resulting the last iteration
ind$variables_sort

# Correction factors of each variable
ind$correction_factors

# Correlations between P2 distance indicator and variables
ind$cor.coeff

# Discrimination coefficient of each variable
```

```
ind$discrimination.coefficient

## Plot of P2 Distance Indicator for European countries
barplot(ind$p2distance, beside=TRUE, col="white", space=.3, ylab="P2 distance",
        ylim=c(0,20), names.arg=rownames(ind$p2distance), las=3, cex.names=0.8)
```

welfare

Welfare Indicators of European Countries

Description

Quality of life's indicators (*welfare indicators*) of 27 countries of European Union during 2002-2007 period. For each variable (i.e. *indicator*) both, definition and year is specified. The source is available in EurLIFE (see details)

Usage

```
data(welfare)
```

Format

A data frame with 27 observations on the following 20 variables.

`happiness` Mean value on a scale of 1 'Very unhappy' to 10 'Very happy'. Year: 2007

`life.satis` Life satisfaction: Share of individuals aged 15 and over who are very satisfied or fairly satisfied with the life they lead. Year: 2006

`judicial` Trust in the judicial system: Percentage of people aged 15 and over who tend to trust in justice or the legal system. Year: 2005

`night` Unsafe to walk around at night: Percentage of individuals aged 18 and over who think that it is very unsafe or rather unsafe to walk around the area they live at night. Year: 2003

`social` Satisfaction with social life: Percentage of the population aged 15 and over, who are very or fairly satisfied with their social life. Year: 2004

`people` Trust in people: Mean value on a scale of 1 'You can't be too careful in dealing with people' to 10 'Most people can be trusted'. Year: 2007

`family` Satisfaction with family life: Mean value on a scale from 1 'very dissatisfied' to 10 'very satisfied' with the family life. Year: 2007

`health` Satisfaction with the national health care system: Percentage of individuals who are very or fairly satisfied with their national health system. Year: 2002

`life.65` Life expectancy at the age of 65: Average number of further years a person at the age of 65 would live if, for the time period, the observed age-specific mortality rates remains constant. Year: 2003

`life.0` Life expectancy at birth: average number of years a person would live if, for the time period, the observed age-specific mortality rate remains constant. Year: 2005

- inequality** Inequality of income distribution: The ratio of total income received by the 20% of the population with the highest income (top quintile) to that received by the 20% of the population with the lowest income (lowest quintile). Year: 2005
- hobbies** Too little time for hobbies and interests: Percentage of people aged 18 and over having too little time for hobbies and interests. Year: 2007
- education** Satisfaction with education: mean value on a scale of 1 'Very dissatisfied' to 10 'Very satisfied' with their own education. Year: 2007
- standard** Satisfaction with standard of living: Mean value on a scale of 1 'very dissatisfied' to 10 'very satisfied' with the own present standard of living. Year: 2007
- dist.school** Distance to the nearest primary school: Proportion of people aged 15 and over who live within walking distance or within 20 minutes of the nearest primary school. Year: 1999, 2005
- area** Satisfaction with the area you live in: Share of individuals aged 15 and over who are very satisfied or fairly satisfied with area they live on a five item scale. Year: 2004
- home** Satisfaction with the home: Share of individuals aged 15 and over who are very satisfied or fairly satisfied with their home on a four item scale. Year: 2004
- stress** Find work stressful: Share of people who 'strongly agree' or 'agree' that their work is too demanding and stressful. Year: 2007
- employment** Employment rate: Employed persons aged 15-64 as a percentage of the total population in that age group. Year: 2005
- job** Job Satisfaction: Percentage of employed people who are very or fairly satisfied with their job, measured on a four item scale. Year: 2005

Details

These quality of life indicators come from different sources (Standard Eurobarometer Surveys; European Quality of Life Surveys; Eurostat; European Foundation for Improvement of Living and Working Conditions) that are available from EurLIFE, a interactive database on quality of life in Europe (see source below).

Source

EurLIFE database. Interactive database on quality of life in Europe. European Foundation for the Improvement of Living and Working Conditions (Eurofound). <http://www.eurofound.europa.eu/areas/qualityoflife/eurlife/index.php> - Accessed 20 Feb 2012

References

Somarriba, N. and Pena, B. (2009). Synthetic Indicators of Quality of Life in Europe. *Social Indicators Research*, **94**, 115–133.

Examples

```
data(welfare)
```

Index

*Topic **datasets**

welfare, [8](#)

*Topic **p2distance**

loadCSVtoP2distance, [3](#)

makeReferenceVector, [4](#)

p2distance, [5](#)

p2distance-package, [2](#)

*Topic **synthetic indicators**

p2distance, [5](#)

p2distance-package, [2](#)

loadCSVtoP2distance, [3](#)

makeReferenceVector, [4](#), [5](#)

p2distance, [4](#), [5](#)

p2distance-package, [2](#)

welfare, [8](#)