

Package ‘FactorCopula’

December 19, 2021

Version 0.8

Date 2021-12-18

Title Factor, Bi-Factor and Second-Order Copula Models

Author Sayed H. Kadhem [aut, cre],
Aristidis K. Nikoloulopoulos [aut]

Maintainer Sayed H. Kadhem <s.kadhem@uea.ac.uk>

Depends R (>= 3.5.0), statmod, abind, utils, polycor, VineCopula,
matlab

Description Estimation, model selection and goodness-of-fit of (1) factor copula models for mixed continuous and discrete data in Kadhem and Nikoloulopoulos (2021) <[doi:10.1111/bmsp.12231](https://doi.org/10.1111/bmsp.12231)>; (2) bi-factor and second-order copula models for item response data in Kadhem and Nikoloulopoulos (2021) <[arXiv:2102.10660](https://arxiv.org/abs/2102.10660)>.

License GPL (>= 2)

NeedsCompilation yes

Repository CRAN

Date/Publication 2021-12-19 01:52:15 UTC

R topics documented:

FactorCopula-package	2
discrepancy	3
GSS	5
M2.Factor	6
M2.StructuredFactor	9
mapping	12
mle.Factor	13
mle.StructuredFactor	16
PE	19
rFactor	19
rStructuredFactor	22
Select.Factor	24
Select.StructuredFactor	26

semicorr	28
TAS	30
transformation	31
Vuong.Factor	31
Vuong.StructuredFactor	34

Index	38
--------------	-----------

FactorCopula-package *Factor, bi-factor and second-order copula models*

Description

Estimation, model selection and goodness-of-fit of (1) factor copula models for mixed continuous and discrete data in Kadhem and Nikoloulopoulos (2021a); (2) bi-factor and second-order copula models for item response data in Kadhem and Nikoloulopoulos (2021b).

Details

This package contains R functions for:

- diagnostics based on semi-correlations (Kadhem and Nikoloulopoulos, 2021a,b; Joe, 2014) to detect tail dependence or tail asymmetry;
- diagnostics to show that a dataset has a factor structure based on linear factor analysis (Kadhem and Nikoloulopoulos, 2021a,b ; Joe, 2014);
- estimation of the factor copula models in Krupskii and Joe (2013), Nikoloulopoulos and Joe (2015), and Kadhem and Nikoloulopoulos (2021a, 2021b);
- model selection of the factor copula models in Krupskii and Joe (2013), Nikoloulopoulos and Joe (2015) and Kadhem and Nikoloulopoulos (2021a, 2021b) using the heuristic algorithms in Kadhem and Nikoloulopoulos (2021a, 2021b) that automatically selects the bivariate parametric copula families that link the observed to the latent variables;
- goodness-of-fit of the factor copula models in Krupskii and Joe (2013), Nikoloulopoulos and Joe (2015) and Kadhem and Nikoloulopoulos (2021a, 2021b) using the M_2 statistic (Maydeu-Olivares and Joe, 2006). Note that the continuous and count data have to be transformed to ordinal.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>

Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

Joe, H. (2014). *Dependence Modelling with Copulas*. Chapman & Hall, London.

Maydeu-Olivares, A. and Joe, H. (2006). Limited information goodness-of-fit testing in multidimensional contingency tables. *Psychometrika*, **71**, 713–732. doi: [10.1007/s1133600512959](https://doi.org/10.1007/s1133600512959).

Kadhem, S.H. and Nikoloulopoulos, A.K. (2021a) Factor copula models for mixed data. *British Journal of Mathematical and Statistical Psychology*, **74**, 365–403. doi: [10.1111/bmsp.12231](https://doi.org/10.1111/bmsp.12231).

Kadhem, S.H. and Nikoloulopoulos, A.K. (2021b) Bi-factor and second-order copula models for item response data. *Arxiv e-prints*, <arXiv:2102.10660>. <https://arxiv.org/abs/2102.10660>.

Krupskii, P. and Joe, H. (2013) Factor copula models for multivariate data. *Journal of Multivariate Analysis*, **120**, 85–101. doi: [10.1016/j.jmva.2013.05.001](https://doi.org/10.1016/j.jmva.2013.05.001).

Nikoloulopoulos, A.K. and Joe, H. (2015) Factor copula models with item response data. *Psychometrika*, **80**, 126–150. doi: [10.1007/s1133601393874](https://doi.org/10.1007/s1133601393874).

discrepancy

Diagnostics to detect a factor dependence structure

Description

The diagnostic method in Joe (2014, pages 245-246) to show that each dataset has a factor structure based on linear factor analysis. The correlation matrix $\mathbf{R}_{\text{observed}}$ has been obtained based on the sample correlations from the bivariate pairs of the observed variables. These are the linear (when both variables are continuous), polychoric (when both variables are ordinal), and polyserial (when one variable is continuous and the other is ordinal) sample correlations among the observed variables. The resulting $\mathbf{R}_{\text{observed}}$ is generally positive definite if the sample size is not small enough; if not one has to convert it to positive definite. We calculate various measures of discrepancy between $\mathbf{R}_{\text{observed}}$ and $\mathbf{R}_{\text{model}}$ (the resulting correlation matrix of linear factor analysis), such as the maximum absolute correlation difference $D_1 = \max |\mathbf{R}_{\text{model}} - \mathbf{R}_{\text{observed}}|$, the average absolute correlation difference $D_2 = \text{avg} |\mathbf{R}_{\text{model}} - \mathbf{R}_{\text{observed}}|$, and the correlation matrix discrepancy measure $D_3 = \log(\det(\mathbf{R}_{\text{model}})) - \log(\det(\mathbf{R}_{\text{observed}})) + \text{tr}(\mathbf{R}_{\text{model}}^{-1} \mathbf{R}_{\text{observed}}) - d$.

Usage

```
discrepancy(cormat, n, f3)
```

Arguments

cormat	$\mathbf{R}_{\text{observed}}$.
n	Sample size.
f3	If TRUE, then the linear 3-factor analysis is fitted.

Value

A matrix with the calculated discrepancy measures for different number of factors.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
 Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

- Joe, H. (2014). *Dependence Modelling with Copulas*. Chapman & Hall, London.
- Kadhem, S.H. and Nikoloulopoulos, A.K. (2021) Factor copula models for mixed data. *British Journal of Mathematical and Statistical Psychology*, **74**, 365–403. doi: [10.1111/bmsp.12231](https://doi.org/10.1111/bmsp.12231).

Examples

```
#-----
#                               PE Data
#-----
data(PE)
#correlation
continuous.PE1 <- -PE[,1]
continuous.PE <- cbind(continuous.PE1, PE[,2])
u.PE <- apply(continuous.PE, 2, rank)/(nrow(PE)+1)
z.PE <- qnorm(u.PE)
categorical.PE <- data.frame(apply(PE[, 3:5], 2, factor))
nPE <- cbind(z.PE, categorical.PE)

#-----
# Discrepancy measures-----
#-----
#correlation matrix for mixed data
cormat.PE <- as.matrix(polycor::hetcor(nPE, std.err=FALSE))
#discrepancy measures
out.PE = discrepancy(cormat.PE, n = nrow(nPE), f3 = FALSE)

#-----
#-----
#                               GSS Data
#-----
data(GSS)
attach(GSS)
continuous.GSS <- cbind(INCOME,AGE)
continuous.GSS <- apply(continuous.GSS, 2, rank)/(nrow(GSS)+1)
z.GSS <- qnorm(continuous.GSS)
ordinal.GSS <- cbind(DEGREE,PINCOME,PDEGREE)
count.GSS <- cbind(CHILDREN,PCHILDREN)

# Transforming the count variables to ordinal
# count1 : CHILDREN
count1 = count.GSS[,1]
count1[count1 > 3] = 3

# count2: PCHILDREN
count2 = count.GSS[,2]
count2[count2 > 7] = 7

# Combining both transformed count variables
ncount.GSS = cbind(count1, count2)

# Combining ordinal and transformed count variables
```

```

categorical.GSS <- cbind(ordinal.GSS, ncount.GSS)
categorical.GSS <- data.frame(apply(categorical.GSS, 2, factor))

# combining continuous and categorical variables
nGSS = cbind(z.GSS, categorical.GSS)

#-----
# Discrepancy measures-----
#-----
#correlation matrix for mixed data
cormat.GSS <- as.matrix(polycor::hetcor(nGSS, std.err=FALSE))
#discrepancy measures
out.GSS = discrepancy(cormat.GSS, n = nrow(nGSS), f3 = TRUE)

```

GSS

The 1994 General Social Survey

Description

Hoff (2007) analysed seven demographic variables of 464 male respondents to the 1994 General Social Survey. Of these seven, two were continuous (income and age of the respondents), three were ordinal with 5 categories (highest degree of the survey respondent, income and highest degree of respondent's parents), and two were count variables (number of children of the survey respondent and respondent's parents).

Usage

```
data(GSS)
```

Format

A data frame with 464 observations on the following 7 variables:

INCOME Income of the respondent in 1000s of dollars, binned into 21 ordered categories.

DEGREE Highest degree ever obtained (0:None, 1:HS, 2:Associates, 3:Bachelors, 4:Graduate).

CHILDREN Number of children of the survey respondent.

PINCOME Financial status of respondent's parents when respondent was 16 (on a 5-point scale).

PDEGREE Highest degree of the survey respondent's parents (0:None, 1:HS, 2:Associates, 3:Bachelors, 4:Graduate).

PCHILDREN Number of children of the survey respondent's parents - 1.

AGE Age of the respondents in years.

Source

Hoff, P. D. (2007). Extending the rank likelihood for semiparametric copula estimation. *The Annals of Applied Statistics*, **1**, 265–283.

M2.Factor

*Goodness-of-fit of factor copula models for mixed data***Description**

The limited information M_2 statistic (Maydeu-Olivares and Joe, 2006) of factor copula models for mixed continuous and discrete data.

Usage

```
M2.1F(tcontinuous, ordinal, tcount, cpar, copF1, gl)
```

```
M2.2F(tcontinuous, ordinal, tcount, cpar, copF1, copF2, gl, SpC)
```

Arguments

tcontinuous	$n \times d_1$ matrix with the transformed continuous to ordinal reponse data, where n and d_1 is the number of observations and transformed continuous variables, respectively.
ordinal	$n \times d_2$ matrix with the ordinal reponse data, where n and d_2 is the number of observations and ordinal variables, respectively.
tcount	$n \times d_3$ matrix with the transformed count to ordinal reponse data, where n and d_3 is the number of observations and transformed count variables, respectively.
cpar	A list of estimated copula parameters.
copF1	$(d_1 + d_2 + d_3)$ -vector with the names of bivariate copulas that link the each of the observed variables with the 1st factor. Choices are "bvn" for BVN, "bvt ν " with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, "frk" for Frank, "gum" for Gumbel, "rgum" for reflected Gumbel, "1rgum" for 1-reflected Gumbel, "2rgum" for 2-reflected Gumbel, "joe" for Joe, "rjoe" for reflected Joe, "1rjoe" for 1-reflected Joe, "2rjoe" for 2-reflected Joe, "BB1" for BB1, "rBB1" for reflected BB1, "BB7" for BB7, "rBB7" for reflected BB7, "BB8" for BB8, "rBB8" for reflected BB8, "BB10" for BB10, "rBB10" for reflected BB10.
copF2	$(d_1 + d_2 + d_3)$ -vector with the names of bivariate copulas that link the each of the observed variables with the 2nd factor. Choices are "bvn" for BVN, "bvt ν " with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, "frk" for Frank, "gum" for Gumbel, "rgum" for reflected Gumbel, "1rgum" for 1-reflected Gumbel, "2rgum" for 2-reflected Gumbel, "joe" for Joe, "rjoe" for reflected Joe, "1rjoe" for 1-reflected Joe, "2rjoe" for 2-reflected Joe, "BB1" for BB1, "rBB1" for reflected BB1, "BB7" for BB7, "rBB7" for reflected BB7, "BB8" for BB8, "rBB8" for reflected BB8, "BB10" for BB10, "rBB10" for reflected BB10.
gl	Gauss legendre quadrature nodes and weights.
SpC	Special case for the 2-factor copula model with BVN copulas. Select a bivariate copula at the 2nd factor to be fixed to independence. e.g. "SpC = 1" to set the first copula at the 2nd factor to independence.

Details

The M_2 statistic has been developed for goodness-of-fit testing in multidimensional contingency tables by Maydeu-Olivares and Joe (2006). Nikoloulopoulos and Joe (2015) have used the M_2 statistic to assess the goodness-of-fit of factor copula models for ordinal data. We build on the aforementioned papers and propose a methodology to assess the overall goodness-of-fit of factor copula models for mixed continuous and discrete responses. Since the M_2 statistic has been developed for multivariate ordinal data, we propose to first transform the continuous and count variables to ordinal and then calculate the M_2 statistic at the maximum likelihood estimate before transformation.

Value

A list containing the following components:

M2	The M_2 statistic which has a null asymptotic distribution that is χ^2 with $s - q$ degrees of freedom, where s is the number of univariate and bivariate margins that do not include the category 0 and q is the number of model parameters.
df	$s - q$.
p-value	The resultant p -value.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

- Kadhem, S.H. and Nikoloulopoulos, A.K. (2021) Factor copula models for mixed data. *British Journal of Mathematical and Statistical Psychology*, **74**, 365–403. doi: [10.1111/bmsp.12231](https://doi.org/10.1111/bmsp.12231).
- Maydeu-Olivares, A. and Joe, H. (2006). Limited information goodness-of-fit testing in multidimensional contingency tables. *Psychometrika*, **71**, 713–732. doi: [10.1007/s1133600512959](https://doi.org/10.1007/s1133600512959).
- Nikoloulopoulos, A.K. and Joe, H. (2015) Factor copula models with item response data. *Psychometrika*, **80**, 126–150. doi: [10.1007/s1133601393874](https://doi.org/10.1007/s1133601393874).

Examples

```
#-----
# Setting quadreture points
nq <- 25
gl <- gauss.quad.prob(nq)
#-----
#                               PE Data
#-----
data(PE)
continuous.PE1 = -PE[,1]
continuous.PE2 = PE[,2]
continuous.PE <- cbind(continuous.PE1, continuous.PE2)
```

```

categorical.PE <- PE[, 3:5]
#-----
#           Estimation
#-----
#----- One-factor -----
# one-factor copula model
cop1f.PE <- c("joe", "joe", "rjoe", "joe", "gum")
est1factor.PE <- mle1factor(continuous.PE, categorical.PE,
                           count=NULL, copF1=cop1f.PE, gl, hessian = T)
#-----
#           M2
#-----
#Transforming the continuous to ordinal data:
ncontinuous.PE = continuous2ordinal(continuous.PE, 5)
# M2 statistic for the one-factor copula model:

m2.1f.PE <- M2.1F(ncontinuous.PE, categorical.PE, tcount=NULL,
                 cpar=est1factor.PE$cpar, copF1=cop1f.PE, gl)

#-----
#           GSS Data
#-----
data(GSS)
attach(GSS)
continuous.GSS <- cbind(INCOME,AGE)
ordinal.GSS <- cbind(DEGREE,PINCOME,PDEGREE)
count.GSS <- cbind(CHILDREN,PCHILDREN)

#-----
#           Estimation
#-----
# one-factor copula model
cop1f.GSS <- c("joe", "2rjoe", "bvt3", "bvt3",
              "rgum", "2rjoe", "2rgum")
est1factor.GSS <- mle1factor(continuous.GSS, ordinal.GSS,
                            count.GSS, copF1=cop1f.GSS, gl, hessian = T)

#two-factor copula model
cop1.2f <- c("rgum", "rjoe", "bvn", "1rjoe",
            "1rjoe", "rjoe", "gum")
cop2.2f <- c("gum", "2rjoe", "rjoe", "gum",
            "bvt5", "bvn", "2rgum")
est2factor.GSS <- mle2factor(continuous.GSS, ordinal.GSS,
                            count.GSS, copF1=cop1.2f, copF2=cop2.2f, gl, hessian = T)

#-----
#           Transformation
#-----
# Transforming the continuous to ordinal data:

# continuous1: Income
continuous1 = as.integer(cut(continuous.GSS[,1],
c(0,10,19,29,40,100), include.lowest = T))

```



```

continuous1 = continuous1 - 1

# continuous2: AGE
continuous2 = as.integer(cut(continuous.GSS[,2] ,
c(0, 24, 44, 64, 100), include.lowest = T))
continuous2 = continuous2 - 1

# Combining the transformed continuous variables.
ncontinuous.GSS <- cbind(continuous1, continuous2)

#----- COUNT VARIABLE -----
# count1 : CHILDREN
count1 = count.GSS[,1]
count1[count1 > 3] = 3

# count2: PCHILDREN
count2 = count.GSS[,2]
count2[count2 > 7] = 7

# Combining both transformed count variables
ncount.GSS = cbind(count1, count2)

#-----
#                               M2
#-----
# M2 statistic for the one-factor copula model:

m2.1f.GSS <- M2.1F(ncontinuous.GSS, ordinal.GSS,
                  ncount.GSS, cpar = est1factor.GSS$cpar,
                  copF1 = cop1f.GSS, gl)

#-----
# M2 statistic for the two-factor copula model:

m2.2f.GSS <- M2.2F(ncontinuous.GSS, ordinal.GSS, ncount.GSS,
                  cpar = est2factor.GSS$cpar, copF1 = cop1.2f,
                  copF2 = cop2.2f, gl)

```

M2.StructuredFactor *Goodness-of-fit of bi-factor and second-order copula models for item response data*

Description

The limited information M_2 statistic (Maydeu-Olivares and Joe, 2006) of bi-factor and second-order copula models for item response data.

Usage

```
M2Bifactor(y,cpar, copnames1, copnames2, gl, ngrp, grpsize)
M2Second_order(y,cpar, copnames1, copnames2, gl, ngrp, grpsize)
```

Arguments

y	$n \times d$ matrix with the ordinal response data, where n and d is the number of observations and variables, respectively.
cpar	A list of estimated copula parameters.
copnames1	For the bi-factor copula: d -vector with the names of bivariate copulas that link each of the observed variables with the common factor. For the second-order factor copula: G -vector with the names of bivariate copulas that link the each of the group-specific factors with the common factor, where G is the number of groups of items. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{2, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel.
copnames2	For the bi-factor copula: d -vector with the names of bivariate copulas that link the each of the observed variables with the group-specific factor. For the second-order factor copula: d -vector with the names of bivariate copulas that link the each of the observed variables with the group-specific factor. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{2, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel.
gl	Gauss legendre quadrature nodes and weights.
ngrp	number of non-overlapping groups.
grpsize	vector indicating the size for each group, e.g., c(4,4,4) indicating four items in all three groups.

Details

The M_2 statistic has been developed for goodness-of-fit testing in multidimensional contingency tables by Maydeu-Olivares and Joe (2006). We use the M_2 to assess the overall fit for the bi-factor and second-order copula models for item response data (Kadhem & Nikoloulopoulos, 2021).

Value

A list containing the following components:

M2	The M_2 statistic which has a null asymptotic distribution that is χ^2 with $s - q$ degrees of freedom, where s is the number of univariate and bivariate margins that do not include the category 0 and q is the number of model parameters.
df	$s - q$.
p-value	The resultant p -value.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
 Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

Kadhem, S.H. and Nikoloulopoulos, A.K. (2021) Bi-factor and second-order copula models for item response data. *Arxiv e-prints*, <arXiv:2102.10660>. <https://arxiv.org/abs/2102.10660>.
 Maydeu-Olivares, A. and Joe, H. (2006). Limited information goodness-of-fit testing in multidimensional contingency tables. *Psychometrika*, **71**, 713–732. doi: [10.1007/s1133600512959](https://doi.org/10.1007/s1133600512959).

Examples

```
#-----
# Setting quadrature points
nq <- 25
g1 <- gauss.quad.prob(nq)
#-----
#                               TAS Data
#-----
data(TAS)
grp1=c(1,3,6,7,9,13,14)
grp2=c(2,4,11,12,17)
grp3=c(5,8,10,15,16,18,19,20)
#Rearrange items within testlets
ydat=TAS[,c(grp1,grp2,grp3)]

d=ncol(ydat);d
n=nrow(ydat);n

#size of each group
g1=length(grp1)
g2=length(grp2)
g3=length(grp3)

grpsize=c(g1,g2,g3)#group size
#number of groups
ngrp=length(grpsize)

#-----
#                               M2
#-----
#BI-FACTOR
tauX0 = c(0.49,0.16,0.29,0.09,0.47,0.49,0.30,
          0.46,0.41,0.33,0.29,0.24,0.10,0.16,
          0.14,0.12,0.03,0.03,0.10,0.10)
tauXg = c(0.09,0.37,0.23,0.53,0.24,0.32,0.27,
          0.53,0.58,0.20,0.23,0.25,0.34,0.33,
          0.30,0.19,0.24,0.29,0.43,0.26)
copX0 = rep("bvt2", d)
```

```

copXg = c(rep("rgum", g1), rep("bvt3", g2+g3))
#converting taus to cpar
cparX0=mapply(function(x,y) tau2par(x,y),x=copX0,y=tauX0)
cparXg=mapply(function(x,y) tau2par(x,y),x=copXg,y=tauXg)
cpar=c(cparX0,cparXg)

m2_Bifactor = M2Bifactor(y=ydat, cpar, copX0, copXg, g1, ngrp, grpssize)

#SECOND-ORDER
tauX0Xg=c(0.60,0.74,0.18)
tauXgY=c(0.48,0.23,0.34,0.25,0.51,0.56,0.37,0.64,0.57,
0.37,0.35,0.32,0.33,0.33,0.29,0.23,0.23,0.25,0.39,0.28)
cparX0Xg=tau2par("bvn", tauX0Xg)
cparXgY=tau2par("bvn", tauXgY)

cpar=c(cparX0Xg,cparXgY)
copX0Xg = rep("bvn", ngrp)
copXgY = rep("bvn", g1+g2+g3)
m2_Second_order = M2Second_order(y=ydat,cpar, copX0Xg, copXgY, g1, ngrp, grpssize)

```

mapping

Mapping of Kendall's tau and copula parameter

Description

Bivariate copulas: mapping of Kendall's tau and copula parameter.

Usage

```

par2tau(copulaname, cpar)
tau2par(copulaname, tau)

```

Arguments

copulaname	Choices are "bvn" for BVN, "bvt ν " with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, "frk" for Frank, "gum" for Gumbel, "rgum" for reflected Gumbel, "1rgum" for 1-reflected Gumbel, "2rgum" for 2-reflected Gumbel, "joe" for Joe, "rjoe" for reflected Joe, "1rjoe" for 1-reflected Joe, "2rjoe" for 2-reflected Joe, "BB1" for BB1, "rBB1" for reflected BB1, "BB7" for BB7, "rBB7" for reflected BB7, "BB8" for BB8, "rBB8" for reflected BB8, "BB10" for BB10, "rBB10" for reflected BB10.
cpar	Copula parameter(s).
tau	Kendall's tau.

Value

Kendall's tau or copula parameter.

References

- Joe H (1997) *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
- Joe H (2014) *Dependence Modeling with Copulas*. Chapman & Hall, London.
- Joe H (2014) *CopulaModel: Dependence Modeling with Copulas*. Software for book: *Dependence Modeling with Copulas*, Chapman & Hall, London 2014.

Examples

```
# 1-param copulas
#BVN copula
cpar.bvn = tau2par("bvn", 0.5)
tau.bvn = par2tau("bvn", cpar.bvn)

#Frank copula
cpar.frk = tau2par("frk", 0.5)
tau.frk = par2tau("frk", cpar.frk)

#Gumbel copula
cpar.gum = tau2par("gum", 0.5)
tau.gum = par2tau("gum", cpar.gum)

#Joe copula
cpar.joe = tau2par("joe", 0.5)
tau.joe = par2tau("joe", cpar.joe)

# 2-param copulas
#BB1 copula
tau.bb1 = par2tau("bb1", c(0.5,1.5))

#BB7 copula
tau.bb7 = par2tau("bb7", c(1.5,1))

#BB8 copula
tau.bb8 = par2tau("bb8", c(3,0.8))

#BB10 copula
tau.bb10 = par2tau("bb10", c(3,0.8))
```

mle.Factor

Maximum likelihood estimation of factor copula models for mixed data

Description

We use a two-stage estimation approach toward the estimation of factor copula models for mixed continuous and discrete data.

Usage

```
mle1factor(continuous, ordinal, count, copF1, gl, hessian, print.level)
mle2factor(continuous, ordinal, count, copF1, copF2, gl, hessian, print.level)
mle2factor.bvn(continuous, ordinal, count, copF1, copF2, gl, SpC, print.level)
```

Arguments

continuous	$n \times d_1$ matrix with the continuous reponse data, where n and d_1 is the number of observations and continuous variables, respectively.
ordinal	$n \times d_2$ matrix with the ordinal reponse data, where n and d_2 is the number of observations and ordinal variables, respectively.
count	$n \times d_3$ matrix with the count reponse data, where n and d_3 is the number of observations and count variables, respectively.
copF1	$(d_1 + d_2 + d_3)$ -vector with the names of bivariate copulas that link the each of the observed variabls with the 1st factor. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel, “joe” for Joe, “rjoe” for reflected Joe, “1rjoe” for 1-reflected Joe, “2rjoe” for 2-reflected Joe, “BB1” for BB1, “rBB1” for reflected BB1, “BB7” for BB7, “rBB7” for reflected BB7, “BB8” for BB8, “rBB8” for reflected BB8, “BB10” for BB10, “rBB10” for reflected BB10.
copF2	$(d_1 + d_2 + d_3)$ -vector with the names of bivariate copulas that link the each of the observed variabls with the 2nd factor. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel, “joe” for Joe, “rjoe” for reflected Joe, “1rjoe” for 1-reflected Joe, “2rjoe” for 2-reflected Joe, “BB1” for BB1, “rBB1” for reflected BB1, “BB7” for BB7, “rBB7” for reflected BB7, “BB8” for BB8, “rBB8” for reflected BB8, “BB10” for BB10, “rBB10” for reflected BB10.
gl	Gauss legendre quadrature nodes and weights.
SpC	Special case for the 2-factor copula model with BVN copulas. Select a bivariate copula at the 2nd factor to be fixed to independence. e.g. “SpC = 1” to set the first copula at the 2nd factor to independence.
hessian	If TRUE, the hessian of the negative log-likelihood is calculated during the minimization process.
print.level	Determines the level of printing which is done during the minimization process; same as in nlm.

Details

Estimation is achieved by maximizing the joint log-likelihood over the copula parameters with the univariate parameters/distributions fixed as estimated at the first step of the proposed two-step estimation approach.

Value

A list containing the following components:

cutpoints	The estimated univariate cutpoints (fitting the univariate probit model).
negbinest	The estimated univariate parameters for the count responses (fitting the negative binomial distribution).
loglik	The maximized joint log-likelihood.
cpar	Estimated copula parameters in a list form.
taus	The estimated copula parameters in Kendall's tau scale.
SEs	The SEs of the Kendall's tau estimates.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

- Kadhem, S.H. and Nikoloulopoulos, A.K. (2021) Factor copula models for mixed data. *British Journal of Mathematical and Statistical Psychology*, **74**, 365–403. doi: [10.1111/bmsp.12231](https://doi.org/10.1111/bmsp.12231).
- Krupskii, P. and Joe, H. (2013) Factor copula models for multivariate data. *Journal of Multivariate Analysis*, **120**, 85–101. doi: [10.1016/j.jmva.2013.05.001](https://doi.org/10.1016/j.jmva.2013.05.001).
- Nikoloulopoulos, A.K. and Joe, H. (2015) Factor copula models with item response data. *Psychometrika*, **80**, 126–150. doi: [10.1007/s1133601393874](https://doi.org/10.1007/s1133601393874).

Examples

```
#-----
# Setting quadrature points
nq <- 25
gl <- gauss.quad.prob(nq)
#-----
#                               PE Data
#-----
data(PE)
continuous.PE1 = -PE[,1]
continuous.PE2 = PE[,2]
continuous.PE <- cbind(continuous.PE1, continuous.PE2)

categorical.PE <- PE[, 3:5]
#-----
#                               Estimation
#-----
#----- One-factor -----
# one-factor copula model
cop1f.PE <- c("joe", "joe", "rjoe", "joe", "gum")
est1factor.PE <- mle1factor(continuous.PE, categorical.PE,
                           count=NULL, copF1=cop1f.PE, gl, hessian = T)
```

```

est1factor.PE
#-----
#-----
#           GSS Data
#-----
data(GSS)
attach(GSS)
continuous.GSS <- cbind(INCOME, AGE)
ordinal.GSS <- cbind(DEGREE, PINCOME, PDEGREE)
count.GSS <- cbind(CHILDREN, PCHILDREN)

#-----
#           Estimation
#-----
#----- One-factor -----
# one-factor copula model
cop1f.GSS <- c("joe", "2rjoe", "bvt3", "bvt3",
              "rgum", "2rjoe", "2rgum")
est1factor.GSS <- mle1factor(continuous.GSS, ordinal.GSS,
                             count.GSS, copF1 = cop1f.GSS, gl, hessian = T)

#----- Two-factor -----
# two-factor copula model
cop1.2f <- c("rgum", "rjoe", "bvn", "1rjoe",
            "1rjoe", "rjoe", "gum")
cop2.2f <- c("gum", "2rjoe", "rjoe", "gum",
            "bvt5", "bvn", "2rgum")
est2factor.GSS <- mle2factor(continuous.GSS, ordinal.GSS,
                             count.GSS, copF1 = cop1.2f, copF2 = cop2.2f, gl, hessian = T)

```

mle.StructuredFactor *Maximum likelihood estimation of the bi-factor and second-order copula models for item response data*

Description

We approach the estimation of the bi-factor and second-order copula models for item response data with the IFM method of Joe (2005).

Usage

```

mleBifactor(y, copnames1, copnames2, gl, ngrp, grpsetSize,
            hessian, print.level)
mleSecond_order(y, copnames1, copnames2, gl, ngrp, grpsetSize,
                hessian, print.level)

```


Arguments

<code>y</code>	$n \times d$ matrix with the item response data, where n and d is the number of observations and variables, respectively.
<code>copnames1</code>	For the bi-factor copula: d -vector with the names of bivariate copulas that link the each of the observed variables with the common factor. For the second-order factor copula: G -vector with the names of bivariate copulas that link the each of the group-specific factors with the common factor, where G is the number of groups of items. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel.
<code>copnames2</code>	For the bi-factor copula: d -vector with the names of bivariate copulas that link the each of the observed variables with the group-specific factor. For the second-order factor copula: d -vector with the names of bivariate copulas that link the each of the observed variables with the group-specific factor. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel.
<code>gl</code>	Gauss legendre quadrature nodes and weights.
<code>ngrp</code>	number of non-overlapping groups.
<code>grpsize</code>	vector indicating the size for each group, e.g., <code>c(4,4,4)</code> indicating four items in all three groups.
<code>hessian</code>	If TRUE, the hessian of the negative log-likelihood is calculated during the minimization process.
<code>print.level</code>	Determines the level of printing which is done during the minimization process; same as in <code>nlm</code> .

Details

Estimation is achieved by maximizing the joint log-likelihood over the copula parameters with the univariate cutpoints fixed as estimated at the first step of the proposed two-step estimation approach.

Value

A list containing the following components:

<code>cutpoints</code>	The estimated univariate cutpoints (fitting the univariate probit model).
<code>taus</code>	The estimated copula parameters in Kendall’s tau scale.
<code>SEs</code>	The SEs of the Kendall’s tau estimates.
<code>loglik</code>	The maximized joint log-likelihood.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
 Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

 PE

Political-economic risk of 62 countries for the year 1987

Description

Quinn (2004) used 5 mixed variables, namely the continuous variable black-market premium in each country (used as a proxy for illegal economic activity), the continuous variable productivity as measured by real gross domestic product per worker in 1985 international prices, the binary variable independence of the national judiciary (1 if the judiciary is judged to be independent and 0 otherwise), and the ordinal variables measuring the lack of expropriation risk and lack of corruption.

Usage

data(PE)

Format

A data frame with 62 observations (countries) on the following 5 variables:

BM Black-market premium.

GDP Gross domestic product.

IJ Independent judiciary.

XPR Lack of expropriation risk.

CPR Lack of corruption.

Source

Quinn, K. M. (2004). Bayesian factor analysis for mixed ordinal and continuous responses. *Political Analysis*, **12**, 338–353.

 rFactor

Simulation of factor copula models for mixed continuous and discrete data

Description

Simulating dependent standard uniform and ordinal response data from factor copula models.

Usage

r1factor(n, d1, d2, categ, theta, copF1)

r2factor(n, d1, d2, categ, theta, delta, copF1, copF2)

Arguments

n	Sample size.
d1	Number of standard uniform variables.
d2	Number of ordinal variables.
categ	A vector of categories for the ordinal variables.
theta	Copula parameters for the 1st factor.
delta	Copula parameters for the 2nd factor.
copF1	$(d_1 + d_2)$ -vector with the names of bivariate copulas that link the each of the observed variables with the 1st factor. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel, “joe” for Joe, “rjoe” for reflected Joe, “1rjoe” for 1-reflected Joe, “2rjoe” for 2-reflected Joe, “BB1” for BB1, “rBB1” for reflected BB1, “BB7” for BB7, “rBB7” for reflected BB7, “BB8” for BB8, “rBB8” for reflected BB8, “BB10” for BB10, “rBB10” for reflected BB10.
copF2	$(d_1 + d_2)$ -vector with the names of bivariate copulas that link the each of the observed variables with the 2nd factor. Choices are “bvn” for BVN, “bvt $[\nu]$ ” with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel, “joe” for Joe, “rjoe” for reflected Joe, “1rjoe” for 1-reflected Joe, “2rjoe” for 2-reflected Joe, “BB1” for BB1, “rBB1” for reflected BB1, “BB7” for BB7, “rBB7” for reflected BB7, “BB8” for BB8, “rBB8” for reflected BB8, “BB10” for BB10, “rBB10” for reflected BB10.

Value

Data matrix of dimension $n \times d$, where n is the sample size, and $d = d_1 + d_2$ is the total number of variables.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

Kadhem, S.H. and Nikoloulopoulos, A.K. (2021) Factor copula models for mixed data. *British Journal of Mathematical and Statistical Psychology*, **74**, 365–403. doi: [10.1111/bmsp.12231](https://doi.org/10.1111/bmsp.12231).

Examples

```
# -----
# -----
#           One-factor copula model
# -----
```

```

# -----
#Sample size -----
n = 100

#Continuous Variables -----
d1 = 5

#Ordinal Variables -----
d2 = 3

#Categories for ordinal -----
categ = c(3,4,5)

#Copula parameters -----
theta = rep(2, d1+d2)

#Copula names -----
copnamesF1 = rep("gum", d1+d2)

#----- Simulating data -----
datF1 = r1factor(n, d1=d1, d2=d2, categ, theta, copnamesF1)

#----- Plotting continuous data -----
pairs(qnorm(datF1[, 1:d1]))

# -----
# -----
#           Two-factor copula model
# -----
# -----
#Sample size -----
n = 100

#Continuous Variables -----
d1 = 5

#Ordinal Variables -----
d2 = 3

#Categories for ordinal -----
categ = c(3,4,5)

#Copula parameters -----
theta = rep(2.5, d1+d2)
delta = rep(1.5, d1+d2)

#Copula names -----
copnamesF1 = rep("gum", d1+d2)
copnamesF2 = rep("gum", d1+d2)

#----- Simulating data -----
datF2 = r2factor(n, d1=d1, d2=d2, categ, theta, delta,
                copnamesF1, copnamesF2)

```

```
#----- Plotting data -----
pairs(qnorm(datF2[,1:d1]))
```

rStructuredFactor	<i>Simulation of bi-factor and second-order copula models for item response data</i>
-------------------	--

Description

Simulating dependent item response data from the bi-factor and second-order copula models for item response data.

Usage

```
rBifactor(n, d, grpssize, categ, copnames1, copnames2, theta1, theta2)
rSecond_order(n, d, grpssize, categ, copnames1, copnames2, theta1, theta2)
```

Arguments

n	Sample size.
d	Number of observed variables/items.
grpssize	vector indicating the size for each group, e.g., c(4,4,4) indicating four items in all three groups.
categ	A vector of categories for the observed variables/items.
theta1	For the bi-factor model: copula parameter vector of size d for items with the common factor. For the second-order copulas: copula parameter vector of size G for the common factor and group-specific factors.
theta2	For the bi-factor model: copula parameter vector of size d for items with the group-specific factor. For the second-order copulas: copula parameter vector of size d for items with the group-specific factor.
copnames1	For the bi-factor copula: d -vector with the names of bivariate copulas that link the each of the observed variables with the common factor. For the second-order factor copula: G -vector with the names of bivariate copulas that link the each of the group-specific factors with the common factor, where G is the number of groups of items. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel.
copnames2	For the bi-factor copula: d -vector with the names of bivariate copulas that link the each of the observed variables with the group-specific factor. For the second-order factor copula: d -vector with the names of bivariate copulas that link the each of the observed variables with the group-specific factor. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel.

Value

Data matrix of dimension $n \times d$, where n is the sample size, and d is the total number of observed variables/items.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

Kadhem, S.H. and Nikoloulopoulos, A.K. (2021) Bi-factor and second-order copula models for item response data. *Arxiv e-prints*, <arXiv:2102.10660>. <https://arxiv.org/abs/2102.10660>.

Examples

```
# -----
# -----
#Sample size
n = 500

#Ordinal Variables -----
d = 9
grpsize=c(3,3,3)
ngrp=length(grpsize)

#Categories for ordinal -----
categ = rep(3,d)

# -----
# -----
#           Bi-factor copula model
# -----
# -----
#Copula parameters
theta = rep(2.5, d)
delta = rep(1.5, d)

#Copula names
copulanames1 = rep("gum", d)
copulanames2 = rep("gum", d)

#----- Simulating data -----
data_Bifactor = rBifactor(n, d, grpsize, categ, copulanames1,
copulanames2, theta, delta)

# -----
# -----
#           Second-order copula model
# -----
# -----
#Copula parameters
```

```

theta= rep(1.5, ngrp)
delta = rep(2.5, d)

#Copula names
copulanames1 = rep("gum", ngrp)
copulanames2 = rep("gum", d)

data_Second_order = rSecond_order(n, d, grpsize, categ,
copulanames1, copulanames2, theta, delta)

```

Select.Factor

Model selection of the factor copula models for mixed data

Description

A heuristic algorithm that automatically selects the bivariate parametric copula families that link the observed to the latent variables.

Usage

```

select1F(continuous, ordinal, count, copnamesF1, g1)
select2F(continuous, ordinal, count, copnamesF1, copnamesF2, g1)

```

Arguments

continuous	$n \times d_1$ matrix with the continuous reponse data, where n and d_1 is the number of observations and continuous variables, respectively.
ordinal	$n \times d_2$ matrix with the ordinal reponse data, where n and d_2 is the number of observations and ordinal variables, respectively.
count	$n \times d_3$ matrix with the count reponse data, where n and d_3 is the number of observations and count variables, respectively.
copnamesF1	A vector with the names of possible candidates of bivariate copulas that link the each of the observed variables with the 1st factor. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel, “joe” for Joe, “rjoe” for reflected Joe, “1rjoe” for 1-reflected Joe, “2rjoe” for 2-reflected Joe, “BB1” for BB1, “rBB1” for reflected BB1, “BB7” for BB7, “rBB7” for reflected BB7, “BB8” for BB8, “rBB8” for reflected BB8, “BB10” for BB10, “rBB10” for reflected BB10.
copnamesF2	A list with the names of possible candidates of bivariate copulas that link the each of the observed variables with the 1st and 2nd factors. Choices are “bvn” for BVN, “bvt ν ” with $\nu = \{1, \dots, 9\}$ degrees of freedom for t-copula, “frk” for Frank, “gum” for Gumbel, “rgum” for reflected Gumbel, “1rgum” for 1-reflected Gumbel, “2rgum” for 2-reflected Gumbel, “joe” for Joe, “rjoe” for reflected Joe, “1rjoe” for 1-reflected Joe, “2rjoe” for 2-reflected Joe, “BB1” for BB1, “rBB1” for reflected BB1, “BB7” for BB7, “rBB7” for reflected BB7, “BB8” for BB8, “rBB8” for reflected BB8, “BB10” for BB10, “rBB10” for reflected BB10.

gl Gauss legendre quadrature nodes and weights.

Details

The linking copulas at each factor are selected with a sequential algorithm under the initial assumption that linking copulas are Frank, and then sequentially copulas with non-tail quadrant independence are assigned to any of pairs where necessary to account for tail asymmetry (discrete data) or tail dependence (continuous data).

Value

A list containing the following components:

‘‘1st factor’’ The selected bivariate linking copulas for the 1st factor.
 ‘‘2nd factor’’ The selected bivariate linking copulas for the 2nd factor.
 AIC Akaike information criterion.
 taus The estimated copula parameters in Kendall’s tau scale.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
 Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

Kadhem, S.H. and Nikoloulopoulos, A.K. (2021) Factor copula models for mixed data. *British Journal of Mathematical and Statistical Psychology*, **74**, 365–403. doi: [10.1111/bmsp.12231](https://doi.org/10.1111/bmsp.12231).

Examples

```
#-----
#                               Estimation
#-----
# Setting quadrature points
nq<-25
gl<-gauss.quad.prob(nq)
#-----
#                               PE Data
#-----
data(PE)
continuous.PE1 = -PE[,1]
continuous.PE <- cbind(continuous.PE1, PE[,2])
categorical.PE <- PE[, 3:5]

#----- One-factor -----
# listing the possible copula candidates:
d <- ncol(PE)
copulasF1 <- rep(list(c("bvn", "bvt3", "bvt5", "frk", "gum",
"rgum", "rjoe", "joe", "1rjoe", "2rjoe", "1rgum", "2rgum")), d)
out1F.PE <- select1F(continuous.PE, categorical.PE,
```

```

count=NULL, copulasF1, gl)

#-----
#-----
#           GSS Data
#-----
data(GSS)
attach(GSS)
continuous.GSS <- cbind(INCOME, AGE)
ordinal.GSS <- cbind(DEGREE, PINCOME, PDEGREE)
count.GSS <- cbind(CHILDREN, PCHILDREN)

#----- One-factor -----
# listing the possible copula candidates:
d <- ncol(GSS)
copulasF1 <- rep(list(c("bvn", "bvt3", "bvt5", "frk", "gum",
"rgum", "rjoe", "joe", "1rjoe", "2rjoe", "1rgum", "2rgum")), d)
out1F.GSS <- select1F(continuous.GSS, ordinal.GSS, count.GSS, copulasF1, gl)

#----- two-factor -----
# listing the possible copula candidates:
copulasF1 = copulasF2 = rep(list(c("bvn", "bvt3", "bvt5", "frk",
"rgum", "rgum", "rjoe", "joe", "1rjoe", "2rjoe", "1rgum", "2rgum")), d)
out2F.GSS <- select2F(continuous.GSS, ordinal.GSS,
count.GSS, copulasF1, copulasF2, gl)

```

Select.StructuredFactor

Model selection of the bi-factor and second-order copula models for item response data

Description

A heuristic algorithm that automatically selects the bivariate parametric copula families for the bi-factor and second-order copula models for item response data.

Usage

```

selectBifactor(y, grpssize, copnames, gl)
selectSecond_order(y, grpssize, copnames, gl)

```

Arguments

y $n \times d$ matrix with the item response data, where n and d is the number of observations and variables, respectively.


```

grp1=c(1,3,6,7,9,13,14)
grp2=c(2,4,11,12,17)
grp3=c(5,8,10,15,16,18,19,20)
ydat=TAS[,c(grp1,grp2,grp3)]

#size of each group
g1=length(grp1)
g2=length(grp2)
g3=length(grp3)
grpsize=c(g1,g2,g3)

# listing the possible copula candidates:
copnames=c("bvn", "bvt2", "bvt3",
" gum", "rgum")

Bifactor_model = selectBifactor(ydat, grpsize, copnames, gl)
Second_order_model = selectSecond_order(ydat, grpsize, copnames, gl)

```

semicorr

Diagnostics to detect tail dependence or tail asymmetry.

Description

The sample versions of the correlation ρ_N , upper semi-correlation ρ_N^+ (correlation in the joint upper quadrant) and lower semi-correlation ρ_N^- (correlation in the joint lower quadrant). These are sample linear (when both variables are continuous), polychoric (when both variables are ordinal), and polyserial (when one variable is continuous and the other is ordinal) correlations.

Usage

```
semicorr(dat, type)
```

Arguments

dat	Data frame of mixed continuous and ordinal data.
type	A vector with 1's for the location of continuous data and 2's for the location of ordinal data.

Value

A matrix containing the following components for semicorr():

rho	ρ_N .
lrho	ρ_N^- .
urho	ρ_N^+ .

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
 Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

Joe, H. (2014). *Dependence Modelling with Copulas*. Chapman and Hall/CRC.
 Kadhem, S.H. and Nikoloulopoulos, A.K. (2021) Factor copula models for mixed data. *British Journal of Mathematical and Statistical Psychology*, **74**, 365–403. doi: [10.1111/bmsp.12231](https://doi.org/10.1111/bmsp.12231).

Examples

```
#-----
#                               PE Data
#-----
data(PE)
#correlation
continuous.PE1 <- -PE[,1]
continuous.PE <- cbind(continuous.PE1, PE[,2])
categorical.PE <- data.frame(apply(PE[, 3:5], 2, factor))
nPE <- cbind(continuous.PE, categorical.PE)

#-----
# Semi-correlations-----
#-----
# Exclude the dichotomous variable
sem.PE = nPE[,-3]
semicorr.PE = semicorr(dat = sem.PE, type = c(1,1,2,2))
#-----
#-----
#                               GSS Data
#-----
data(GSS)
attach(GSS)
continuous.GSS <- cbind(INCOME,AGE)
ordinal.GSS <- cbind(DEGREE,PINCOME,PDEGREE)
count.GSS <- cbind(CHILDREN,PCHILDREN)

# Transforming the COUNT variables to ordinal
# count1 : CHILDREN
count1 = count.GSS[,1]
count1[count1 > 3] = 3

# count2: PCHILDREN
count2 = count.GSS[,2]
count2[count2 > 7] = 7

# Combining both transformed count variables
ncount.GSS = cbind(count1, count2)
```

```

# Combining ordinal and transformed count variables
categorical.GSS <- cbind(ordinal.GSS, ncount.GSS)
categorical.GSS <- data.frame(apply(categorical.GSS, 2, factor))

# combining continuous and categorical variables
nGSS = cbind(continuous.GSS, categorical.GSS)
#-----
# Semi-correlations-----
#-----
semicorr.GSS = semicorr(dat = nGSS, type = c(1, 1, rep(2,5)))

```

TAS

Toronto Alexithymia Scale (TAS)

Description

The Toronto Alexithymia Scale is the most utilized measure of alexithymia in empirical research and is composed of $d = 20$ items that can be subdivided into $G = 3$ non-overlapping groups: $d_1 = 7$ items to assess difficulty identifying feelings (DIF), $d_2 = 5$ items to assess difficulty describing feelings (DDF) and $d_3 = 8$ items to assess externally oriented thinking (EOT). Students were 17 to 25 years old and 58% of them were female and 42% were male. They were asked to respond to each item using one of $K = 5$ categories: “1 = completely disagree”, “2 = disagree”, “3 = neutral”, “4 = agree”, “5 = completely agree”.

Usage

```
data(TAS)
```

Format

A data frame with 1925 observations on the following 20 items:

DIF items: 1,3,6,7,9,13,14.

DDF items: 2,4,11,12,17.

EOT items: 5,8,10,15,16,18,19,20.

Source

Briganti, G. and Linkowski, P. (2020). Network approach to items and domains from the toronto alexithymia scale. *Psychological Reports*, **123**, 2038–2052.

Williams, D. and Mulder, J. (2020). *BGGM: Bayesian Gaussian Graphical Models*. R package version 1.0.0.

transformation	<i>Continuous/count to ordinal responses</i>
----------------	--

Description

Transforming a continuous/count to ordinal variable with K categories.

Usage

```
continuous2ordinal(continuous, categ)
count2ordinal(count, categ)
```

Arguments

continuous	Matrix of continuous data.
count	Matrix of count data.
categ	The number of categories K .

Examples

```
#-----
#           PE Data
#-----
data(PE)
continuous.PE <- PE[, 1:2]

#Transforming the continuous to ordinal data :
tcontinuous = continuous2ordinal(continuous.PE, categ=5)
table(tcontinuous)

#Transforming the count to ordinal data:
set.seed(12345)
count.data = rpois(1000, 3)
tcount = count2ordinal(count.data, 5)
table(tcount)
```

Vuong.Factor	<i>Vuong's test for the comparison of factor copula models</i>
--------------	--

Description

Vuong (1989)'s test for the comparison of non-nested factor copula models for mixed data. We compute the Vuong's test between the factor copula model with BVN copulas (that is the standard factor model) and a competing factor copula model to reveal if the latter provides better fit than the standard factor model.

Usage

```
vuong.1f(cpar.bvn, cpar, copF1, continuous, ordinal, count, gl, param)
vuong.2f(cpar.bvn, cpar, copF1, copF2, continuous, ordinal, count, gl, param)
```

Arguments

<code>cpar.bvn</code>	copula parameters of the factor copula model with BVN copulas.
<code>cpar</code>	copula parameters of the competing factor copula model.
<code>copF1</code>	copula names for the first factor of the competing factor copula model.
<code>copF2</code>	copula names for the second factor of the competing factor copula model.
<code>continuous</code>	matrix of continuous data.
<code>ordinal</code>	matrix of ordinal data.
<code>count</code>	matrix of count data.
<code>gl</code>	gauss-legendre quadrature points.
<code>param</code>	parameterization of estimated copula parameters. If FALSE, then <code>cpar</code> are the actual copula parameters without any transformation/reparameterization.

Value

A vector containing the following components:

<code>z</code>	the test statistic.
<code>p.value</code>	the p -value.
<code>CI.left</code>	lower/left endpoint of 95% confidence interval.
<code>CI.right</code>	upper/right endpoint of 95% confidence interval.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
 Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

Kadhem, S.H. and Nikoloulopoulos, A.K. (2021) Factor copula models for mixed data. *British Journal of Mathematical and Statistical Psychology*, **74**, 365–403. doi: [10.1111/bmsp.12231](https://doi.org/10.1111/bmsp.12231).

Vuong, Q.-H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, **57**, 307–333.

Examples

```
#-----
# Setting quadrature points
nq <- 25
gl <- gauss.quad.prob(nq)
#-----
```



```

#                               PE Data
#-----
data(PE)
continuous.PE1 = -PE[,1]
continuous.PE2 = PE[,2]
continuous.PE <- cbind(continuous.PE1, continuous.PE2)
categorical.PE <- PE[, 3:5]
d <- ncol(PE)
#-----
#                               Estimation
#-----
# factor copula model with BVN copulas
cop1f.PE.bvn <- rep("bvn", d)
PE.bvn1f <- mle1factor(continuous.PE, categorical.PE,
count=NULL, copF1=cop1f.PE.bvn, gl, hessian = T)

# Selected factor copula model
cop1f.PE <- c("joe", "joe", "rjoe", "joe", "gum")
PE.selected1f <- mle1factor(continuous.PE, categorical.PE,
count=NULL, copF1=cop1f.PE, gl, hessian = T)
#-----
#                               Vuong's test
#-----
v1f.PE.selected <- vuong.1f(PE.bvn1f$cpars$f1,
PE.selected1f$cpars$f1, cop1f.PE, continuous.PE,
categorical.PE, count=NULL, gl, param=F)

#-----
#-----
#                               GSS Data
#-----
data(GSS)
attach(GSS)
continuous.GSS <- cbind(INCOME, AGE)
ordinal.GSS <- cbind(DEGREE, PINCOME, PDEGREE)
count.GSS <- cbind(CHILDREN, PCHILDREN)
d <- ncol(GSS)

#-----
#                               Estimation
#-----
# factor copula model with BVN copulas
# one-factor copula model
cop1f.GSS.bvn <- rep("bvn", d)
GSS.bvn1f <- mle1factor(continuous.GSS, ordinal.GSS,
count.GSS, copF1 = cop1f.GSS.bvn, gl, hessian = T)

# two-factor copula model
cop1f.GSS.bvn = cop2f.GSS.bvn = rep("bvn", d)
GSS.bvn2f <- mle2factor.bvn(continuous.GSS, ordinal.GSS,
count.GSS, copF1 = cop1f.GSS.bvn, copF2 = cop2f.GSS.bvn, gl, SpC = 7)

# Selected factor copula model

```

```

# one-factor copula model
cop1f.GSS <- c("joe","2rjoe","bvt3","bvt3",
              "rgum","2rjoe","2rgum")
GSS.selected1f <- mle1factor(continuous.GSS, ordinal.GSS,
count.GSS, copF1 = cop1f.GSS, gl, hessian = T)

# two-factor copula model
cop2f1.GSS <- c("rgum","rjoe","bvn","1rjoe","1rjoe","rjoe","gum")
cop2f2.GSS <- c("gum","2rjoe","rjoe","gum","bvt5","bvn","2rgum")
GSS.selected2f <- mle2factor(continuous.GSS, ordinal.GSS,
count.GSS, copF1 = cop2f1.GSS, copF2 = cop2f2.GSS, gl, hessian = T)

#-----
#              Vuong's test
#-----
#1-factor
v1f.GSS.selected <- vuong.1f(GSS.bvn1f$cpar$f1,
GSS.selected1f$cpar$f1, cop1f.GSS, continuous.GSS,
ordinal.GSS, count=count.GSS, gl, param=F)

#2-factor
v2f.GSS.selected <- vuong.2f(GSS.bvn2f$cpar,
GSS.selected2f$cpar, cop2f1.GSS, cop2f2.GSS,
continuous.GSS, ordinal.GSS, count=count.GSS, gl, param=F)

```

Vuong.StructuredFactor

Vuong's test for the comparison of bi-factor and second-order copula models

Description

The Vuong's test (Vuong,1989) is the sample version of the difference in Kullback-Leibler divergence between two models and can be used to differentiate two parametric models which could be non-nested. For the Vuong's test we provide the 95% confidence interval of the Vuong's test statistic (Joe, 2014, page 258). If the interval does not contain 0, then the best fitted model according to the AICs is better if the interval is completely above 0.

Usage

```
vuong_structured(models, cpar.m1, copnames.m1, cpar.m2,
copnames.m2, y, grpsize)
```

Arguments

`models` choose a number from (1,2,3,4). **1:** Model1 is bifactor, Model2 is bifactor. **2:** Model1 is second-order, Model2 is second-order. **3:** Model1 is second-order, Model2 is bifactor. **4:** Model1 is bifactor, Model2 is nested.

<code>cpar.m1</code>	vector of copula parameters for model 1, starting with copula parameters that link the items with common factor (bifactor), or group factors with common factor (second-order).
<code>cpar.m2</code>	vector of copula parameters for model 2, starting with copula parameters that link the items with common factor (bifactor), or group factors with common factor (second-order).
<code>copnames.m1</code>	vector of names of copula families for model 1, starting with copulas that link the items with common factor (bifactor), or group factors with common factor (second-order).
<code>copnames.m2</code>	vector of names of copula families for model 2, starting with copulas that link the items with common factor (bifactor), or group factors with common factor (second-order).
<code>y</code>	matrix of ordinal data.
<code>grpsize</code>	vector indicating the size for each group, e.g., <code>c(4,4,4)</code> indicating four items in all three groups.

Value

A vector containing the following components:

<code>z</code>	the test statistic.
<code>p.value</code>	the p -value.
<code>CI.left</code>	lower/left endpoint of 95% confidence interval.
<code>CI.right</code>	upper/right endpoint of 95% confidence interval.

Author(s)

Sayed H. Kadhem <s.kadhem@uea.ac.uk>
 Aristidis K. Nikoloulopoulos <a.nikoloulopoulos@uea.ac.uk>

References

Joe, H. (2014). *Dependence Modelling with Copulas*. Chapman and Hall/CRC.

Kadhem, S.H. and Nikoloulopoulos, A.K. (2021) Bi-factor and second-order copula models for item response data. *Arxiv e-prints*, <arXiv:2102.10660>. <https://arxiv.org/abs/2102.10660>.

Vuong, Q.-H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, **57**, 307–333.

Examples

```
#-----
# Setting quadrature points
nq <- 25
gl <- gauss.quad.prob(nq)
#-----
#                               TAS Data
```

```

#-----
data(TAS)
grp1=c(1,3,6,7,9,13,14)
grp2=c(2,4,11,12,17)
grp3=c(5,8,10,15,16,18,19,20)
ydat=TAS[,c(grp1,grp2,grp3)]

d=ncol(ydat);d
n=nrow(ydat);n

#Rearrange items within testlets
g1=length(grp1)
g2=length(grp2)
g3=length(grp3)

grpsize=c(g1,g2,g3)#group size
#number of groups
ngrp=length(grpsize)
#-----
# M1 bifactor - M2 bifactor
cpar.m1 = rep(0.6,d*2)
copnames.m1 = rep("bvn",d*2)
cpar.m2 = rep(1.6,d*2)
copnames.m2 = rep("rgum",d*2)
vuong.bifactor = vuong_structured(models=1, cpar.m1, copnames.m1,
                                cpar.m2, copnames.m2,
                                y=ydat, grpsize)

# M1 second-order - M2 second-order
cpar.m1 = rep(0.6,d+ngrp)
copnames.m1 = rep("bvn",d+ngrp)
cpar.m2 = rep(1.6,d+ngrp)
copnames.m2 = rep("rgum",d+ngrp)
vuong.second_order = vuong_structured(models=2, cpar.m1,
                                copnames.m1, cpar.m2, copnames.m2, y=ydat, grpsize)

# M1 second-order - M2 bifactor
cpar.m1 = rep(0.6,d+ngrp)
copnames.m1 = rep("bvn",d+ngrp)
cpar.m2 = rep(1.6,d*2)
copnames.m2 = rep("rgum",d*2)
vuong.2nd0_bif = vuong_structured(models=3, cpar.m1, copnames.m1,
                                cpar.m2, copnames.m2,
                                y=ydat, grpsize)

# M1 bifactor - M2 second-order
cpar.m1 = rep(0.6,d*2)
copnames.m1 = rep("bvn",d*2)
cpar.m2 = rep(1.6,d+ngrp)
copnames.m2 = rep("rgum",d+ngrp)
vuong.bif_2nd0 = vuong_structured(models=4, cpar.m1, copnames.m1,
                                cpar.m2, copnames.m2,
                                y=ydat, grpsize)

```


Index

- * **correlation**
 - discrepancy, 3
 - semicorr, 28
 - * **datagen**
 - rFactor, 19
 - rStructuredFactor, 22
 - * **datasets**
 - GSS, 5
 - PE, 19
 - TAS, 30
 - * **maximum likelihood**
 - Vuong.Factor, 31
 - Vuong.StructuredFactor, 34
 - * **models**
 - mle.Factor, 13
 - mle.StructuredFactor, 16
 - * **multivariate**
 - M2.Factor, 6
 - M2.StructuredFactor, 9
 - mle.Factor, 13
 - mle.StructuredFactor, 16
 - Select.Factor, 24
 - Select.StructuredFactor, 26
 - * **package**
 - FactorCopula-package, 2
 - * **parameters**
 - mapping, 12
 - * **univar**
 - transformation, 31
- continuous2ordinal (transformation), 31
- count2ordinal (transformation), 31
- discrepancy, 3
- FactorCopula-package, 2
- GSS, 5
- M2.1F (M2.Factor), 6
- M2.2F (M2.Factor), 6
- M2.Factor, 6
- M2.StructuredFactor, 9
- M2Bifactor (M2.StructuredFactor), 9
- M2Second_order (M2.StructuredFactor), 9
- mapping, 12
- mle.Factor, 13
- mle.StructuredFactor, 16
- mle1factor (mle.Factor), 13
- mle2factor (mle.Factor), 13
- mleBifactor (mle.StructuredFactor), 16
- mleSecond_order (mle.StructuredFactor), 16
- par2tau (mapping), 12
- PE, 19
- r1factor (rFactor), 19
- r2factor (rFactor), 19
- rBifactor (rStructuredFactor), 22
- rFactor, 19
- rSecond_order (rStructuredFactor), 22
- rStructuredFactor, 22
- Select.Factor, 24
- Select.StructuredFactor, 26
- select1F (Select.Factor), 24
- select2F (Select.Factor), 24
- selectBifactor (Select.StructuredFactor), 26
- selectSecond_order (Select.StructuredFactor), 26
- semicorr, 28
- TAS, 30
- tau2par (mapping), 12
- transformation, 31
- vuong.1f (Vuong.Factor), 31
- vuong.2f (Vuong.Factor), 31
- Vuong.Factor, 31
- Vuong.StructuredFactor, 34

vuong_structured
 (Vuong.StructuredFactor), [34](#)