Package 'IVCor'

January 9, 2025

Type Package

Title A Robust Integrated Variance Correlation

Version 0.1.0

Description A integrated variance correlation is proposed to measure the dependence between a categorical or continuous random variable and a continuous random variable or vector.
This package is designed to estimate the new correlation coefficient with parametric and non-parametric approaches.
Test of independence for different problems can also be implemented via the new correlation coefficient with this package.

License GPL-3

Encoding UTF-8

Imports splines, quantreg, BwQuant, quantdr, stats

RoxygenNote 7.2.3

Suggests knitr, mvtnorm, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

NeedsCompilation no

Author Wei Xiong [aut],
Han Pan [aut, cre],
Hengjian Cui [aut]

Maintainer Han Pan <scott_pan@163.com>

Repository CRAN

Date/Publication 2025-01-09 18:00:02 UTC

Contents

IVC *Integrated Variance Correlation*

Description

This function is used to calculate the integrated variance correlation between two random variables or between a random variable and a multivariate random variable

Usage

```
IVC(y, x, K, NN = 3, type)
```

Arguments

y	is a numeric vector
x	is a numeric vector or a data matrix
K	is the number of quantile levels
NN	is the number of B spline basis, default is 3
type	is an indicator for measuring linear or nonlinear correlation, "linear" represents linear correlation and "nonlinear" represents linear or nonlinear correlation using B splines

Value

The value of the corresponding sample statistic

Examples

```
# linear model
n=100
x=rnorm(n)
y=3*x+rnorm(n)

IVC(y,x,K=5,type="linear")
# nonlinear model
n=100
p=3
x=matrix(NA,nrow=n,ncol=p)
for(i in 1:p){
 x[,i]=rnorm(n)
```

```
}
y=cos(x[,1]+x[,2])+x[,3]^2+rnorm(n)
IVC(y,x,K=5,type="nonlinear")
```

IVCCA *Integrated Variance Correlation with Discrete Response Variable*

Description

This function is used to calculate the integrated variance correlation between a discrete response variable and a continuous random variable

Usage

```
IVCCA(y, x, K)
```

Arguments

y	is the categorical response vector
x	is a numeric vector
K	is the number of quantile levels

Value

The value of the corresponding sample statistic

Examples

```
n=100
y=sample(rep(1:3), n, replace = TRUE, prob = c(1/3,1/3,1/3))
x=c()
for(i in 1:n){
 x[i]=rnorm(1,mean=2*y[i],sd=1)
}

IVCCA(y,x,K=5)
```

IVCCAT *Integrated Variance Correlation Based Hypothesis Test for Discrete Response*

Description

This function is used to test independence between a categorical variable and a continuous variable using integrated variance correlation

Usage

```
IVCCAT(y, x, K, num_per, type)
```

Arguments

y	is a categorical response vector
x	is a numeric vector
K	is the number of quantile levels
num_per	is the number of permutation times
type	is an indicator for fixed number of categories or infinity number of categories, "fixed" represents number of categories is fixed, then a permutation test is used, "infinity" represents number of categories is infinite, then an asymptotic normal distribution is used to calculate p values

Value

The p-value of the corresponding hypothesis test

Examples

```
# small R
n=100
x=runif(n,0,1)
y=sample(rep(1:3), n, replace = TRUE, prob = c(1/3,1/3,1/3))

IVCCAT(y,x,K=5,num_per=20,type = "fixed")
# large R
n=200
y=sample(rep(1:20), n, replace = TRUE, prob = rep(1/20,20))
mu_x=sample(c(1,2,3,4),20,replace = TRUE,prob = c(1/4,1/4,1/4,1/4))
x=c()
for (i in 1:n) {
 x[i]=2*mu_x[y[i]]+rcauchy(1)
}

IVCCAT(y,x,K=10,type = "infinity")
```

IVCCA_crit *Critical Values for Integrated Variance Correlation Based Hypothesis Test with Discrete Response*

Description

This function is used to calculate the critical values for integrated variance correlation test with discrete response at significance level 0.1, 0.05 and 0.01

Usage

```
IVCCA_crit(R, N = 500, realizations)
```

Arguments

R	is the number of categories
N	is a integer as large as possible, default is 500
realizations	is the the number of replication times for simulating the distribution under the null hypothesis

Value

The critical values at significance level 0.1, 0.05 and 0.01

Examples

```
IVCCA_crit(R=5,N=500,realizations=100)
```

IVCLLQ *Integrated Variance Correlation with Local Linear Estimation*

Description

This function is used to calculate the integrated variance correlation between two random variables with local linear estimation

Usage

```
IVCLLQ(y, x, K)
```

Arguments

y	is a numeric vector
x	is a numeric vector
K	is the number of quantile levels

Value

The value of the corresponding sample statistic

Examples

```
n=100
x=rnorm(n)
y=exp(x)+rnorm(n)

IVCLLQ(y,x,K=4)
```

IVCT *Integrated Variance Correlation Based Hypothesis Test*

Description

This function is used to test significance of linear or nonlinear correlation using integrated variance correlation

Usage

```
IVCT(y, x, K, num_per, NN = 3, type)
```

Arguments

y	is the response vector
x	is a numeric vector or a data matrix
K	is the number of quantile levels
num_per	is the number of permutation times
NN	is the number of B spline basis, default is 3
type	is an indicator for measuring linear or nonlinear correlation, "linear" represents linear correlation and "nonlinear" represents linear or nonlinear correlation using B splines

Value

The p-value of the corresponding hypothesis test

Examples

```
# linear model
n=100
x=rnorm(n)
y=rnorm(n)

IVCT(y,x,K=5,num_per=20,type = "linear")
# nonlinear model
```

```
n=100
p=4
x=matrix(NA,nrow=n,ncol=p)
for(i in 1:p){
 x[,i]=runif(n,0,1)
}
y=3*ifelse(x[,1]>0.5,1,0)*x[,2]+3*cos(x[,3])^2*x[,1]+3*(x[,4]^2-1)*x[,1]+rnorm(n)

IVCT(y,x,K=5,num_per=20,type = "nonlinear")
```

IVCTLLQ *Integrated Variance Correlation Based Hypothesis Test with Local Linear Estimation*

Description

This function is used to test significance using integrated variance correlation with local linear estimation

Usage

```
IVCTLLQ(y, x, K, num_per)
```

Arguments

y	is a numeric vector
x	is a numeric vector
K	is the number of quantile levels
num_per	is the number of permutation times

Value

The p-value of the corresponding hypothesis test

Examples

```
n=100
x=runif(n,-1,1)
y=2*cos(2*x)+rnorm(n)


IVCTLLQ(y,x,K=5,num_per=100)
```

IVCT_Interval *Integrated Variance Correlation Based Interval Independence Hypothesis Test*

Description

This function is used to test interval independence using integrated variance correlation

Usage

```
IVCT_Interval(y, x, tau1, tau2, K, num_per, NN = 3, type)
```

Arguments

y	is the response vector
x	is a numeric vector or a data matrix
tau1	is the minimum quantile level
tau2	is the maximum quantile level
K	is the number of quantile levels
num_per	is the number of permutation times
NN	is the number of B spline basis, default is 3
type	is an indicator for measuring linear or nonlinear correlation, "linear" represents linear correlation and "nonlinear" represents linear or nonlinear correlation using B splines

Value

The p-value of the corresponding hypothesis test

Examples

```
require("mvtnorm")
n=100
p=3
pho1=0.5
mean_x=rep(0,p)
sigma_x=matrix(NA,nrow = p,ncol = p)
for (i in 1:p) {
 for (j in 1:p) {
   sigma_x[i,j]=pho1^(abs(i-j))
 }
}
x=rmvnorm(n, mean = mean_x, sigma = sigma_x,method = "chol")
y=rnorm(n)

IVCT_Interval(y,x,tau1=0.5,tau2=0.75,K=5,num_per=20,type = "linear")
```

```
n=100
x_til=runif(n,min=-1,max=1)
y_til=rnorm(n)
epsilon=rnorm(n)
x=x_til+2*epsilon*ifelse(x_til<=-0.5&y_til<=-0.675,1,0)
y=y_til+2*epsilon*ifelse(x_til<=-0.5&y_til<=-0.675,1,0)

IVCT_Interval(y,x,tau1=0.6,tau2=0.8,K=5,num_per=20,type = "nonlinear")
```

`IVC_crit` *Critical Values for Integrated Variance Correlation Based Hypothesis Test*

Description

This function is used to calculate the critical values for integrated variance correlation test at significance level 0.1, 0.05 and 0.01

Usage

```
IVC_crit(N = 500, realizations)
```

Arguments

`N`	is a integer as large as possible, default is 500
`realizations`	is the the number of replication times for simulating the distribution under the null hypothesis

Value

The critical values at significance level 0.1, 0.05 and 0.01

Examples

```
IVC_crit(N=500,realizations=100)
```

IVC_Interval *Integrated Variance Correlation for Interval Independence*

Description

This function is used to calculate the integrated variance correlation to measure interval independence

Usage

```
IVC_Interval(y, x, K, tau1, tau2, NN = 3, type)
```

Arguments

y	is a numeric vector
x	is a numeric vector or a data matrix
K	is the number of quantile levels
tau1	is the minimum quantile level
tau2	is the maximum quantile level
NN	is the number of B spline basis, default is 3
type	is an indicator for measuring linear or nonlinear correlation, "linear" represents linear correlation and "nonlinear" represents linear or nonlinear correlation using B splines

Value

The value of the corresponding sample statistic for interval independence

Examples

```
# linear model
require("mvtnorm")
n=100
p=3
pho1=0.5
mean_x=rep(0,p)
sigma_x=matrix(NA,nrow = p,ncol = p)
for (i in 1:p) {
 for (j in 1:p) {
   sigma_x[i,j]=pho1^(abs(i-j))
 }
}
x=rmvnorm(n, mean = mean_x, sigma = sigma_x,method = "chol")
y=2*(x[,1]+x[,2]+x[,3])+rnorm(n)

IVC_Interval(y,x,K=5,tau1=0.4,tau2=0.6,type="linear")
# nonlinear model
```

```
n=100
x=runif(n,min=-2,max=2)
y=exp(x^2)*rnorm(n)

IVC_Interval(y,x,K=5,tau1=0.4,tau2=0.6,type="nonlinear")
```

Index