Package 'CorBin'

November 14, 2020

Type Package

Title Generate High-Dimensional Binary Data with Correlation Structures

Version 1.0.0

Author Wei Jiang [aut], Shuang Song [aut, cre], Lin Hou [aut] and Hongyu Zhao [aut]

Maintainer Shuang Song <song-s19@mails.tsinghua.edu.cn>

Description We design algorithms with linear time complexity with respect to the dimension for three commonly studied correlation structures, including exchangeable, decaying-product and K-dependent correlation structures, and extend the algorithms to generate binary data of general non-negative correlation matrices with quadratic time complexity. Jiang, W., Song, S., Hou, L. and Zhao, H. ``A set of efficient methods to generate high-dimensional binary data with specified correlation structures." The American Statistician. See <doi:10.1080/00031305.2020.1816213> for a detailed presentation of the method.

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2020-11-14 09:20:02 UTC

R topics documented:

cBern *Main function*

Description

The main function of our package, through which we can simulate correlated binary data under different settings.

Usage

```
cBern(n, p, rho, type, k = NULL)
```

Arguments

n	number of observations
p	the vector of marginal probabilities with dimension m
rho	For the first three types, rho is either a non-negative value indecating the shared correlation coefficient or and m-1 vector indicating the correlation coefficients between adjacent variables. For the general case, rho should be a list, the i-th element of which specifies the coefficients on the i-th minor diagnal.
type	including 4 types. type="exchange" type="DCP" type="1-dependent" type="General"
k	(for 'General' use only). The number of layers setting for k-dependent structure. k=m-1 for the general case.

Value

an n*p matrix of binary data

References

Jiang, W., Song, S., Hou, L. and Zhao, H. A set of efficient methods to generate high-dimensional binary data with specified correlation structures. *The American Statistician*. DOI:10.1080/00031305.2020.1816213

See Also

cBernEx, cBernDCP, cBern1dep

Examples

```
X <- cBern(10, rep(0.5,3), 0.5, type="exchange")

X <- cBern(10, rep(0.5,3), c(0.2,0.2), type="DCP")

X <- cBern(5, c(0.4,0.5,0.6), c(0.2,0.3), type="1-dependent")

rho <- list()
rho[[1]] <- c(0.2,0.3)
rho[[2]] <- 0.1
X <- cBern(2, c(0.7,0.8,0.9),rho=rho,type="General", k=2)
```

---

cBern1dep *Generate binary data with 1-dependent correlated structure*

Description

Equivalent to cBern(n, p, rho, type="1-dependent")

Usage

```
cBern1dep(n, p, rho)
```

Arguments

n	number of observations
p	the vector of marginal probabilities with dimension m
rho	either a non-negative value indecating the shared correlation coefficient or and m-1 vector indicating the correlation coefficients between adjacent variables.

Value

an n*p matrix of binary data

Examples

```
X <- cBern1dep(5, c(0.4,0.5,0.6), c(0.2,0.3))
```

cBernDCP *Generate binary data with decaying-product correlated structure*

Description

Equivalent to cBern(n, p, rho, type="DCP")

Usage

```
cBernDCP(n, p, rho)
```

Arguments

n	number of observations
p	the vector of marginal probabilities with dimension m
rho	either a non-negative value indecating the shared correlation coefficient or and m-1 vector indicating the correlation coefficients between adjacent variables.

Value

an n*p matrix of binary data

Examples

```
X <- cBernDCP(10, rep(0.5,3), c(0.2,0.2))
```

cBernEx *Generate binary data with exchangeable correlated structure*

Description

Equivalent to cBern(n, p, rho, type="exchange")

Usage

```
cBernEx(n, p, rho)
```

Arguments

n	number of observations
p	the vector of marginal probabilities with dimension m
rho	a non-negative value indecating the shared correlation coefficient

Value

an n*p matrix of binary data

Examples

```
X <- cBernEx(10, rep(0.5,3), 0.5)
```

`rhoMax1dep` *To calculate the maximal allowed correlations max for using cBern1dep to generate binary data with 1-dependent structure*

Description

To calculate the maximal allowed correlations max for using cBern1dep to generate binary data with 1-dependent structure

Usage

```
rhoMax1dep(p)
```

Arguments

`p`	the vector of marginal probabilities with dimension m

Value

an (m-1)-dimensional vector rho, which is the maximum the correlation between the adjacent variables

---

`rhoMaxDCP` *For calculating the maximal allowed correlations max for binary data with decaying-product structure.*

Description

For calculating the maximal allowed correlations max for binary data with decaying-product structure.

Usage

```
rhoMaxDCP(p)
```

Arguments

`p`	marginal probabilities

Value

an (m-1)-dimensional vector rho, which is the maximum the correlation between the adjacent variables

rhoMaxEx *For calculating the maximal allowed correlation coefficient for binary data with exchangeable structure.*

Description

For calculating the maximal allowed correlation coefficient for binary data with exchangeable structure.

Usage

```
rhoMaxEx(p)
```

Arguments

p	the vector of marginal probabilities with dimension m

Value

the maximal allowed correlation coefficient

Index