Package 'PMD'

August 5, 2021

Type Package

Title Computation of Poisson-Multinomial Distribtuions

Version 1.0

Date 2021-08-06

Maintainer Yili Hong <yilihong@vt.edu>

Description Applying exact, simulation and approximation methods for computing probability mass functions, cumulative distribution functions of Poisson-Multinomial distributions together with a random number generator to conduct Poisson-Multinomial distribution sampling. The exact method is based on fast Fourier transformation of the characteristic functions of Poisson-Multinomial distributions. Simulation method is built via naive sampling scheme of different multinomial distributions. Approximation method is designed by asymptotic distrbutions of Poisson-Multinomial distributions, which are normal.

License GPL (>= 2)

Encoding UTF-8

Imports mvtnorm, Rcpp

LinkingTo Rcpp, RcppArmadillo

SystemRequirements fftw3(>=3.3)

RoxygenNote 7.1.1

NeedsCompilation yes

Author Yili Hong [aut, ctb],
Zhengzhi Lin [aut, cre],
Yueyao Wang [aut, ctb]

R topics documented:

dpmd *Probability Mass Function of Poisson-Multinomial Distributions*

Description

Probability mass function of Poisson-Multinomial distributions specified by input matrix and computed through selected method. This function is capable for computation of the whole probability mass function as well as of one single probability mass point.

Usage

```
dpmd(pp, method = "DFT-CF", vec = c(0, 0, 0, 0, 0), B = 100)
```

Arguments

pp	A matrix of probabilities. Each row of pp should add up to 1.
method	Character string stands for the method selected by user to compute the probability mass. The method can only be one of the following four, "DFT-CF" "simulation" "NA by demands" "simulation by demands"
vec	Result vector(probability mass point) specified by user. Eg. pp is 4 by 3 matrix then a user might be interested in the probability of getting result: vec=c(0,0,1,2).
B	Simulation repeating time. Will be ignored if users do not choose "simulation" or "simulation by demands" as method.

Details

For the methods we applied in dpmd, "DFT-CF" is an exact method to calculate all probability mass points of Poisson-Multinomial Distributions via FFT algorithm. When users select "DFT-CF", dpmd will ignore vec and output the whole probability mass function.

"simulation" is a simulation method using naive simulation scheme to calculate the whole probability mass function, under this selection the input of vec will be ignore. Notice the accuracy and running time will be effected by user choice of B. Usually B=10^5 or 10^6 will be accurate enough. Increasing B to larger than 10^8 will heavily aggravate computation burden of a CPU or GPU.

Given pp with dimension $n \times m$, the number of total probability mass points is $(n+1)^{m-1}$. Thus when the dimension of pp increases and the users selected method is one of "DFT-CF" and "simulation", the computation burden of dpmd might challenge the capability of a computer because both of the methods calculate all probability mass points of Poisson-Multinomial distributions.

"NA by demands" specifies an approximation method using Normal approximation to compute the probability mass point of the vec vector input by user.

"simulation by demands" is as same as "simulation" except that it only computes a single probability mass point specified by vec.

Value

For a single probability mass point, dpmd returns a probability value.

For all probability mass points of a given pp, it returns a multi-dimensional array. For instance, for the pp matrix in the following example, the value of the array element $a_{1,2,1}$ = 0.90 means the value of probability mass point (0,1,0,2) is 0.90.

Examples

```
pp=matrix(c(.1, .1, .1, .7, .1, .3, .3, .3, .5, .2, .1, .2), nrow=3, byrow=TRUE)


dpmd(pp)
dpmd(pp,"simulation",B=10^3)
dpmd(pp,"NA by demands", vec = c(0,0,1,2))
dpmd(pp,"simulation by demands", vec = c(0,0,1,2), B=10^3)
```

---

ppmd *Cumulative Distribution Function of Poisson-Multinomial Distribution*

Description

This function computes cumulative distribution function of Poisson-Multinomial distributions that specified by input probability matrix via given method.

Usage

```
ppmd(pp, x, method = "DFT-CF", B = 1000)
```

Arguments

pp	A matrix of probabilities. Each row of pp should add up to 1.
x	Vector $x = (x_1, x_2, \ldots)$ for computing $P(X_1 \leq x_1, X_2 \leq x_2, \ldots)$.
method	Character string stands for the method selected by user to compute the probability mass. The method can only be one of the following three, "DFT-CF" "simulation" "NA"
B	Simulation repeating time. Will be ignored if users do not choose "simulation" as method.

Details

Three methods are same as listed in the details of dpmd but "NA" stands for normal approximation.

ppmd computes the cumulative distribution function by adding all probability mass points within hyper-dimensional space limited by x.

Value

The value of $P(X_1 \leq x_1, X_2 \leq x_2, \ldots)$ of given $x = (x_1, x_2, \ldots)$.

Examples

```
pp=matrix(c(.1, .1, .1, .7, .1, .3, .3, .3, .5, .2, .1, .2), nrow=3, byrow=TRUE)


ppmd(pp, x = c(3,2,1,3))
ppmd(pp, x = c(3,2,1,3), method = "simulation", B = 10^3)
ppmd(pp, x = c(3,2,1,3), method = "NA")
```

rpmd *Poisson-Multinomial Distribution Random Number Generator*

Description

Generating random samples of given a Poisson-Multinomial distribution.

Usage

```
rpmd(pp, n)
```

Arguments

pp	A matrix of probabilities. Each row of pp should add up to 1.
n	Number of samples to be generated.

Value

A matrix of samples, each row stands for one sample.

Examples

```
pp=matrix(c(.1, .1, .1, .7, .1, .3, .3, .3, .5, .2, .1, .2), nrow=3, byrow=TRUE)

n=5

rpmd(pp,5)
```

Index