Package: bivpois (via r-universe)

September 13, 2024

Type Package

Title Bivariate Poisson Distribution

Version 1.0

Date 2023-10-18

Author Michail Tsagris [aut, cre]

Maintainer Michail Tsagris <mtsagris@uoc.gr>

Depends R (>= 4.0)

Imports Rfast, stats

Description Maximum likelihood estimation, random values generation, density computation and other functions for the bivariate Poisson distribution. References include: Kawamura K. (1984). ``Direct calculation of maximum likelihood estimator for the bivariate Poisson distribution''. Kodai Mathematical Journal, 7(2): 211--221. <doi:10.2996/kmj/1138036908>. Kocherlakota S. and Kocherlakota K. (1992). ``Bivariate discrete distributions''. CRC Press. <doi:10.1201/9781315138480>. Karlis D. and Ntzoufras I. (2003). ``Analysis of sports data by using bivariate Poisson models''. Journal of the Royal Statistical Society: Series D (The Statistician), 52(3): 381--393. <doi:10.1111/1467-9884.00366>.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2023-10-19 07:00:17 UTC

Contents

`bivpois-package` *Bivariate Poisson Distribution*

Description

The Bivariate Poisson Distribution.

Details

Package:	bivpois
Type:	Package
Version:	1.0
Date:	2023-10-18
License:	GPL-2

Maintainers

Michail Tsagris <mtsagris@uoc.gr>.

Author(s)

Michail Tsagris <mtsagris@uoc.gr>.

References

Kawamura K. (1984). Direct calculation of maximum likelihood estimator for the bivariate Poisson distribution. Kodai Mathematical Journal, 7(2): 211–221.

Kocherlakota S. and Kocherlakota K. (1998). Bivariate discrete distributions. Wiley Online Library.

Karlis D. and Ntzoufras I. (2003). Analysis of sports data by using bivariate poisson models. Journal of the Royal Statistical Society: Series D (The Statistician), 52(3): 381–393.

Density computation of the bivariate Poisson distribution
Density computation of the bivariate Poisson distribution

Description

Density computation of the bivariate Poisson distribution.

Usage

```
dbp(x1, x2 = NULL, lambda, logged = TRUE)
```

Arguments

x1	Either a numerical vector with the values of the first variable or a matrix with 2 columns containing both variables. In the latter case, x2 must be NULL.
x2	A numerical vector with the values of the second. If x1 is a matrix with 2 columns containing both variables, x2 must be NULL.
lambda	A vector with three numbers, the estimates of the λ_s.
logged	Should the logarithm of the density values be computed? The default value is TRUE.

Details

The density of the bivariate Poisson distribution is computed.

Value

A vector with the logged density values.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Kawamura K. (1984). Direct calculation of maximum likelihood estimator for the bivariate Poisson distribution. Kodai Mathematical Journal, 7(2): 211–221.

Kocherlakota S. and Kocherlakota K. (1992). Bivariate discrete distributions. CRC Press.

Karlis D. and Ntzoufras I. (2003). Analysis of sports data by using bivariate poisson models. Journal of the Royal Statistical Society: Series D (The Statistician), 52(3): 381–393.

See Also

rbp

Examples

```
x <- rbp( 300, c(3, 5, 2) )
a <- bp.mle(x)
f <- dbp(x, lambda = a$lambda)
sum(f)
```

Goodness of fit test for the bivariate Poisson distribution

Goodness of fit test for the bivariate Poisson distribution

Description

Goodness of fit test for the bivariate Poisson distribution.

Usage

```
bp.gof(x1, x2 = NULL, R = 999)
bp.gof2(x1, x2 = NULL, R = 999)
```

Arguments

x1	Either a numerical vector with the values of the first variable or a matrix with 2 columns containing both variables. In the latter case, x2 must be NULL.
x2	A numerical vector with the values of the second. If x1 is a matrix with 2 columns containing both variables, x2 must be NULL.
R	The number of Monte Carlo replicates to use.

Details

Kocherlakota and Kocherlakota (1992) mention the following a goodness of fit test for the bivariate Poisson distribution, the index of dispersion test. They mention that Loukas and Kemp (1986) developed this test as an extension of the univariate dispersion test. They test for departures from the bivariate Poisson againsta alternatives which involve an increase in the generalised variance, the determinant of the covariance matrix of the two variables.

Rayner, Thas and Best (2009) mentions a revised version of this test whose test statistic is now given by

$$I_{B^*} = \frac{n}{1-r^2}\left(\frac{S_1^2}{\bar{x}_1} - 2r^2\sqrt{\frac{S_1^2}{\bar{x}_1}\frac{S_2^2}{\bar{x}_2}} + \frac{S_2^2}{\bar{x}_2}\right),$$

where n is the sample size, r is the sample Pearson correlation coefficient, S_1^2 and S_2^2 are the two sample variances and $\bar{x}_1$ and $\bar{x}_2$ are the two sample means. Under the null hypothesis the I_{B^*} follows asymptotically a χ^2 with $2n-3$ degrees of freedom. However, I did some simulations and I saw that it does not perform very well in terms of the type I error. If you see the simulations in their book (page 132) you will see this. For this reason, the function calculates the p-value of the I_{B^*} using Monte Carlo (or parametric bootstrap).

The second function, bp.gof2(), is a vectorised version of the first, and much faster. I put both of them here to show how one can vectorize a function and make it faster.

Value

A list including:

`runtime`	The duration of the algorithm.
`tab`	The contingency table of the two variables.
`pvalue`	The Monte-Carlo based estimated p-value.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Kocherlakota S. and Kocherlakota K. (1992). Bivariate discrete distributions. CRC Press.

Loukas S. and Kemp C. (1986). The index of dispersion test for the bivariate Poisson distribution. Biometrics, 42(4): 941–948.

Rayner J. C., Thas O. and Best D. J. (2009). Smooth Tests of Goodness of Fit: Using R. John Wiley & Sons.

See Also

`bp.mle`

Examples

```
x <- rbp( 300, c(3, 5, 2) )
bp.gof(x)
bp.gof2(x)
```

`Maximum likelihood estimation of the bivariate Poisson distribution`
Maximum likelihood estimation of the bivariate Poisson distribution

---

Description

Maximum likelihood estimation of the bivariate Poisson distribution.

Usage

```
bp.mle(x1, x2 = NULL)
bp.mle2(x1, x2 = NULL)
```

Arguments

`x1` Either a numerical vector with the values of the first variable or a matrix with 2 columns containing both variables. In the latter case, x2 must be NULL.

`x2` A numerical vector with the values of the second. If x1 is a matrix with 2 columns containing both variables, x2 must be NULL.

Details

Using the addition method (see function rbp) to simulate random values from the bivariate Poisson, its representation is given by

$$P\left(X=x,Y=y\right)e^{-\left(\lambda_1+\lambda_2+\lambda_3\right)}\frac{\lambda_1^x}{x!}\frac{\lambda_2^y}{y!}\sum_{k=0}^{\min\left(x,y\right)}\binom{x}{k}\binom{y}{k}k!\left(\frac{\lambda_3}{\lambda_1\lambda_2}\right)^k.$$

The above form is found in Karlis and Ntzoufras (2003). This bivariate distribution allows for dependence between the two random variables. Marginally each random variable follows a Poisson distribution with $E\left(X\right)=\lambda_1+\lambda_3$ and $E\left(Y\right)=\lambda_2+\lambda_3$. In addition, $Cov\left(X,Y\right)=\lambda_3$. If $\lambda_3=0$, the above expression becomes a product of two Poisson distributions. Hence, λ_3 is a measure of dependence between the two random variables.

The function bp.mle() returns a lot of information and is slower than bp.mle2(), which returns fewer information, but is faster.

Value

For the function bp.mle() a list including:

`lambda` A vector with the estimated values of (λ_1, λ_2) and λ_3. Note that $\hat{\lambda}_1=\bar{x}_1-\lambda_3$ and $\hat{\lambda}_1=\bar{x}_1-\lambda_3$, where $\bar{x}_1$ and $\bar{x}_2$ are the two sample means.

`rho` The estimated correlation coefficient, that is: $\frac{\hat{\lambda}_3}{\sqrt{\left(\hat{\lambda}_1+\hat{\lambda}_3\right)\left(\hat{\lambda}_2+\hat{\lambda}_3\right)}}$.

`ci` The 95% Confidence intervals using the observed and the asymptotic information matrix.

`loglik` The log-likelihood values assuming independence ($\lambda_3=0$) and assuming the bivariate Poisson distribution.

`pvalue` Three p-values for testing $\lambda_3=0$. These are based on the log-likelihood ratio and two Wald tests using the observed and the asymptotic information matrix.

For the function bp.mle2() a list including:

`lambda` A vector with the estimated values of (λ_1, λ_2) and λ_3. Note that $\hat{\lambda}_1=\bar{x}_1-\lambda_3$ and $\hat{\lambda}_1=\bar{x}_1-\lambda_3$, where $\bar{x}_1$ and $\bar{x}_2$ are the two sample means.

`loglik` The log-likelihood values assuming independence ($\lambda_3=0$) and assuming the bivariate Poisson distribution.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Kawamura K. (1984). Direct calculation of maximum likelihood estimator for the bivariate Poisson distribution. Kodai Mathematical Journal, 7(2): 211–221.

Kocherlakota S. and Kocherlakota K. (1992). Bivariate discrete distributions. CRC Press.

Karlis D. and Ntzoufras I. (2003). Analysis of sports data by using bivariate poisson models. Journal of the Royal Statistical Society: Series D (The Statistician), 52(3): 381–393.

See Also

rbp

Examples

```
x <- rbp( 300, c(3, 5, 2) )
bp.mle(x)
```

Profile likelihood of the covariance parameter of the bivariate Poisson distribution

Profile likelihood of the covariance parameter (λ_3) of the bivariate Poisson distribution

---

Description

Profile likelihood of the covariance parameter (λ_3) of the bivariate Poisson distribution.

Usage

```
lambda3.profile(x1, x2 = NULL)
```

Arguments

x1	Either a numerical vector with the values of the first variable or a matrix with 2 columns containing both variables. In the latter case, x2 must be NULL.
x2	A numerical vector with the values of the second. If x1 is a matrix with 2 columns containing both variables, x2 must be NULL.

Details

The function plots the profile log-likelihood of λ_3 and computes the relevant 95% confidence interval for the parameter.

Value

A plot with the profile log-likelihood of λ_3 and a vector with the 95% confidence interval for the parameter.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Kawamura K. (1984). Direct calculation of maximum likelihood estimator for the bivariate Poisson distribution. Kodai Mathematical Journal, 7(2): 211–221.

Kocherlakota S. and Kocherlakota K. (1992). Bivariate discrete distributions. CRC Press.

Karlis D. and Ntzoufras I. (2003). Analysis of sports data by using bivariate poisson models. Journal of the Royal Statistical Society: Series D (The Statistician), 52(3): 381–393.

See Also

bp.mle

Examples

```
x <- rbp( 300, c(3, 5, 2) )
lambda3.profile(x)
```

Random values generation from the bivariate Poisson distribution

Random values generation from the bivariate Poisson distribution

---

Description

Random values generation from the bivariate Poisson distribution.

Usage

```
rbp(n, lambda)
```

Arguments

n	The sample size.
lambda	A vector with the three parameters, $(\lambda_1, \lambda_2, \lambda_3)$ of the Poisson distribution.

Details

In order to generate values from this distribution one needs three independent Poisson variables, $X_1 \sim \text{Po}(\lambda_1)$, $X_2 \sim \text{Po}(\lambda_2)$ and $X_3 \sim \text{Po}(\lambda_3)$. Then, $(X, Y) = (X_1 + X3, X_2 + X_3) \sim \text{BP}(\lambda_1, \lambda_2, \lambda_3)$.

Value

A matrix with n rows and 2 columns.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Kocherlakota S. and Kocherlakota K. (1992). Bivariate discrete distributions. CRC PRess.

See Also

bp.mle

Examples

```
x <- rbp( 300, c(3, 5, 2) )
bp.mle(x)
```

Index