Package ‘gp’

October 23, 2023

Type Package

Title Maximum Likelihood Estimation of the Generalized Poisson Distribution

Version 1.1

Date 2023-10-23

Author Michail Tsagris [aut, cre]

Maintainer Michail Tsagris <mtsagris@uoc.gr>

Depends R (>= 4.0)

Imports Rfast, RNGforGPD, stats

Description Functions to estimate the parameters of the generalized Poisson distribution with or without covariates using maximum likelihood. The references include Nikoloulopoulos A.K. & Karlis D. (2008). ``On modeling count data: a comparison of some well-known discrete distributions''. Journal of Statistical Computation and Simulation, 78(3): 437--457, <doi:10.1080/10629360601010760> and Consul P.C. & Famoye F. (1992). ``Generalized Poisson regression model''. Communications in Statistics - Theory and Methods, 21(1): 89--109, <doi:10.1080/03610929208830766>.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2023-10-23 15:50:16 UTC

R topics documented:

`gp-package` *Maximum Likelihood Estimation of the Generalized Poisson Distribution.*

Description

Maximum likelihood estimation of the generalized Poisson distribution. Regression modelling is also supported.

Details

Package:	gp
Type:	Package
Version:	1.1
Date:	2023-10-23
License:	GPL-2

Maintainers

Michail Tsagris <mtsagris@uoc.gr>.

Author(s)

Michail Tsagris <mtsagris@uoc.gr>.

References

Nikoloulopoulos A.K. & Karlis D. (2008). On modeling count data: a comparison of some well-known discrete distributions. Journal of Statistical Computation and Simulation, 78(3): 437–457.

Consul P.C. & Famoye F. (1992). Generalized poisson regression model. Communications in Statistics - Theory and Methods, 21(1): 89–109.

`Cumulative probability mass of the generalized Poisson distribution` *Cumulative probability mass of the generalized Poisson distribution*

Description

Cumulative probability mass of the generalized Poisson distribution.

Usage

```
pgp(y, theta, lambda)
```

Arguments

`y`	A vector with non negative integer values.
`theta`	The value of the θ parameter.
`lambda`	The value of the λ parameter.

Details

The cumulative probability mass of the generealized Poisson distribution is computed.

Value

A vector with the probabilities.

Author(s)

Michail Tsagris.

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

References

Nikoloulopoulos A.K. & Karlis D. (2008). On modeling count data: a comparison of some well-known discrete distributions. Journal of Statistical Computation and Simulation, 78(3): 437–457.

Demirtas H. (2017). On accurate and precise generation of generalized Poisson variates. Communications in Statistics - Simulation and Computation, 46(1): 489–499.

See Also

`dgp`, `rgp`

Examples

```
y <-  rgp(1000, 10, 0.5, method = "Inversion")
a <- gp.mle(y)
pgp(y[1:10], a[1], a[2])
```

Density computation of the generalized Poisson distribution

Density computation of the generalized Poisson distribution

Description

Density computation of the generalized Poisson distribution.

Usage

```
dgp(y, theta, lambda, logged = TRUE)
```

Arguments

`y`	A vector with non negative integer values.
`theta`	The value of the θ parameter.
`lambda`	The value of the λ parameter.
`logged`	Should the logarithm of the density values be computed? The default value is TRUE.

Details

The density of the generealized 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

Nikoloulopoulos A.K. & Karlis D. (2008). On modeling count data: a comparison of some well-known discrete distributions. Journal of Statistical Computation and Simulation, 78(3): 437–457.

Demirtas H. (2017). On accurate and precise generation of generalized Poisson variates. Communications in Statistics - Simulation and Computation, 46(1): 489–499.

See Also

rgp, pgp

Examples

```
y <-  rgp(1000, 10, 0.5, method = "Inversion")
a <- gp.mle(y)
f <- dgp(y, a[1], a[2])
sum(f)
```

`gp.mle` *Maximum likelihood estimation of the generalized Poisson distribution*

Description

Maximum likelihood estimation of the generalized Poisson distribution.

Usage

```
gp.mle(y)
```

Arguments

y	A vector with non negative integer values.

Details

The probability density function of the generalized Poisson distribution is the following (Nikoloulopoulos & Karlis, 2008):

$$P(Y = y|\theta, \lambda) = \theta(\theta + \lambda y)^{y-1}\frac{e^{-\theta-\lambda y}}{y!}, \;\; y = 0, 1 ... \;\; \theta > 0, \;\; 0 \leq \lambda \leq 1.$$

To ensure that θ is positive we use the "log" link and for λ to lie within 0 and 1 we use the "logit" link within the `optim` function.

Value

A vector with three numbers, the θ and λ parameters and the value of the log-likelihood.

Author(s)

Michail Tsagris.

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

References

Nikoloulopoulos A.K. & Karlis D. (2008). On modeling count data: a comparison of some well-known discrete distributions. Journal of Statistical Computation and Simulation, 78(3): 437–457.

See Also

`gp.reg`, `rgp`

Examples

```
y <-  rgp(1000, 10, 0.5, method = "Inversion")
gp.mle(y)
```

gp.reg *Generalized Poisson regression*

Description

Generalized Poisson regression.

Usage

```
gp.reg(y, x, tol = 1e-7)
gp.reg2(y, x, tol = 1e-7)
```

Arguments

`y`	The response variable, a vector with non negative integer values.
`x`	A data.frame or a matrix with the independent variables.
`tol`	The tolerance value to terminate the optimization.

Details

The loglikelihood of the generalised Poisson distribution when covariates are present is the following (Consul & Famoye, 1992):

$$\ell(\beta,\phi) = \sum_{i=1}^n \log(\mu_i) + \sum_{i=1}^n (y_i - 1)\log\left[\mu_i + (\phi-1)y_i\right] - \log\phi \sum_{i=1}^n y_i - \frac{1}{\phi}\sum_{i=1}^n \left[\mu_i + (\phi-1)y_i\right] - \sum_{i=1}^n \log\left(y_i\right),$$

where $\mu_i = e^{\sum_{j=0}^k X_{ij}\beta_j}$, n denotes the sample size, k is the number of β coefficients, and $\phi > 0$.

Breslow (1984) suggested the (moment) estimation of a dispersion parameter by equating the chi-square statistic to its degrees of freedom. For the generalised Poisson regression model, this leads to $\sum_{i=1}^n \frac{(y_i-\mu_i)^2}{\mu_i\phi^2} = n - k$ and we solve this for ϕ.

According to Consul and Famoye (1992) we begin by fitting a Poisson regression model and obtain initial values for β_s and ϕ. If $\hat{\phi} \approx 1$, it implies that the Poisson regression model is appropriate and no further estimation needs to be done. However, if $\hat{\phi} \neq 1$, this is used to obtain new values of the estimated β_s by maximizing the log-likelihood. This process is iterated until we obtain a stable solution.

The function as seen below returns the log-likelihood of the initial Poisson regression as well. This is useful if one wants to test, via the log-likelihood ratio test as 1 degree of freedom, if the generalized Poisson regression is to be preferred over the Poisson regression.

gp.reg() estimates the $beta$ coefficients using Newton-Raphson, whereas gp.reg2() uses the `optim` function. For some reason these two do not always agree. One might yield higher log-likelihood than the other and this is why I offer both ways.

Value

A list including:

`pois.loglik`	The initial Poisson regression log-likelihood.
`gp.loglik`	The generalized Poisson regression log-likelihood.
`be`	The estimated $beta$ coefficients.
`phi`	The estimated ϕ parameter.

Author(s)

Michail Tsagris.

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

References

Consul P.C. & Famoye F. (1992). Generalized poisson regression model. Communications in Statistics - Theory and Methods, 21(1): 89–109.

Breslow N. E. (1984). Extra-Poisson variation in log-linear models. Journal of the Royal Statistical Society: Series C (Applied Statistics), 33(1): 38–44.

See Also

`gp.mle`

Examples

```
n <- 500
x <- matrix (rnorm(n * 2), nrow = n, ncol = 2)
be <- c(1, 1)
mi <- x[, 1] * be[1] + x[, 2] * be[2] + 1
mi <- exp(mi)
y <- numeric(n)
for (i in 1:n)  y[i] <- rgp(2, mi[i], 0.5, method = "Inversion")[1]
gp.reg(y, x)
gp.reg2(y, x)
```

`rgp` *Random values simulation from the generalized Poisson distribution*

Description

Random values simulation from the generalized Poisson distribution.

Usage

```
rgp(n, theta, lambda, method)
```

Arguments

n	The number of random values to generate.
theta	The value of the θ parameter.
lambda	The value of the λ parameter.
method	The simulation method to use. The available options are: "Inversion", "Branching", "Normal-Approximation", "Build-Up" and "Chop-Down".

Details

The values θ and λ affect the method of simulation to use. See Li et al. (2020) for more information.

Value

A vector with values from the generalized Poisson distribution.

Author(s)

Michail Tsagris.

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

References

Li H., Demirtas H. & Chen R. (2020). RNGforGPD: An R Package for Generation of Univariate and Multivariate Generalized Poisson Data. R JOURNAL, 12(2): 173–188.

Demirtas H. (2017). On accurate and precise generation of generalized Poisson variates. Communications in Statistics - Simulation and Computation, 46(1): 489–499.

See Also

gp.mle

Examples

```
y <-  rgp(1000, 10, 0.5, method = "Inversion")
gp.mle(y)
```

Index