Package ‘bivpois’

May 27, 2005

Title Bivariate Poisson Models Using The EM Algorithm

Version 0.50-1

Date 2005-05-17

Depends R (>= 2.0.1)

Author Dimitris Karlis and Ioannis Ntzoufras

Description Functions for fitting Bivariate Poisson Models using the EM algorithm. Details can be found in Karlis and Ntzoufras (2003, RSS D & 2004,AUEB Technical Report)

Maintainer Ioannis Ntzoufras <ntzoufras@aueb.gr>

URL http://www.stat-athens.aueb.gr/~jbn/papers/paper14.htm

R topics documented:

`bivpois.table` *Probability of Bivariate Poisson Using Recursive relations*

Description

Returns the probability of the bivariate Poisson distribution using recursive relations.

Usage

```
bivpois.table(x, y, lambda = c(1, 1, 1))
```

Arguments

`x, y`	single values containing which values should evaluated (x and y should be at least 1)
`lambda`	Vector (of length 3) containing values of the parameters lambda1, lambda2 and lambda3 of the bivariate poisson distribution.

Details

In order to calculate bivpoison probability values we use recursive relationships. This function is much slower than pbivpois

Value

A matrix with dimension (x+1) X (y+1) is returned. Cell ij contains the probability P(X=i-1, Y=j-1).

Author(s)

1. Dimitris Karlis, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨karlis@aueb.gr⟩ .

2. Ioannis Ntzoufras, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨ntzoufras@aueb.gr⟩ .

References

1. Karlis, D. and Ntzoufras, I. (2004). Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in S. (submitted). Technical Report, Department of Statistics, Athens University of Economics and Business, Athens, Greece.

2. Karlis, D. and Ntzoufras, I. (2003). Analysis of Sports Data Using Bivariate Poisson Models. Journal of the Royal Statistical Society, D, (Statistician), 52, 381 - 393.

See Also

`pbivpois`. `simple.bp`. `lm.bp`. `lm.dibp`.

`ex1.sim` *Bivpois Example 1 Dataset: Simulated Data*

Description

The data has one pair (x, y) of bivariate Poisson variables and five variables $(z_1, \ldots, z_5)$ generated from $N(0, 0.01)$ distribution. Hence

$$Xi, Yi \sim BP(\lambda_{1i}, \lambda_{2i}, \lambda_{3i}) \text{ with}$$
$$\log \lambda_{1i} = 1.8 + 2Z_{1i} + 3Z_{3i}$$
$$\log \lambda_{2i} = 0.7 - Z_{1i} - 3Z_{3i} + 3Z_{5i}$$
$$\log \lambda_{3i} = 1.7 + Z_{1i} - 2Z_{2i} + 2Z_{3i} - 2Z_{4i}.$$

Usage

```
data(ex1.sim)
```

Format

A data frame with 100 observations on the following 7 variables.

x,y Simulated Bivariate Poisson Variables used as response

z1,z2,z3,z4,z5 Simulated N(0,0.01) explanatory variables

Details

This data is used as example one in Karlis and Ntzoufras (2004).

Source

Karlis, D. and Ntzoufras, I. (2004). Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in S. (submitted). Technical Report, Department of Statistics, Athens University of Economics and Business, Athens, Greece.

References

Karlis, D. and Ntzoufras, I. (2003). Analysis of Sports Data Using Bivariate Poisson Models. Journal of the Royal Statistical Society, D, (Statistician), 52, 381 - 393.

Examples

```
library(bivpois)       # load bivpois library
data(ex1.sim)          # load data of example 1
# ----------------------------------------------------------------------------
# Simple Bivariate Poisson Model
ex1.simple<-simple.bp( ex1.sim$x, ex1.sim$y ) # fit simple model of section 4.1.1
names(ex1.simple)      # monitor output variables
ex1.simple$lambda      # view lambda1
ex1.simple$BIC         # view BIC
ex1.simple             # view all results of the model
#
# ----------------------------------------------------------------------------
# Fit Double and Bivariate Poisson models ()
```

```
#
# Model 2: DblPoisson(l1, l2)
ex1.m2<-lm.bp(x~1 , y~1 , data=ex1.sim, zeroL3=TRUE)
# Model 3: BivPoisson(l1, l2, l3); same as simple.bp(ex1.sim$x, ex1.sim$y)
ex1.m3<-lm.bp(x~1 , y~1 , data=ex1.sim)
# Model 4: DblPoisson (l1=Full, l2=Full)
ex1.m4<-lm.bp(x~. , y~. , data=ex1.sim, zeroL3=TRUE)
# Model 5: BivPoisson(l1=full, l2=full, l3=constant)
ex1.m5<-lm.bp(x~. , y~. , data=ex1.sim)
# Model 6: DblPois(l1,l2)
ex1.m6<-lm.bp(x~z1 , y~z1+z5 , l1l2=~z3, data=ex1.sim, zeroL3=TRUE)
# Model 7: BivPois(l1,l2,l3=constant)
ex1.m7<-lm.bp(x~z1 , y~z1+z5 , l1l2=~z3, data=ex1.sim)
# Model 8: BivPoisson(l1=full, l2=full, l3=full)
ex1.m8<-lm.bp(x~. , y~. , l3=~., data=ex1.sim)
# Model 9: BivPoisson(l1=full, l2=full, l3=z1+z2+z3+z4)
ex1.m9<-lm.bp(x~. , y~. , l3=~.-z5, data=ex1.sim)
# Model 10: BivPoisson(l1, l2, l3=full)
ex1.m10<-lm.bp(x~z1 , y~z1+z5 , l1l2=~z3, l3=~., data=ex1.sim)
# Model 11: BivPoisson(l1, l2, l3= z1+z2+z3+z4)
ex1.m11<-lm.bp(x~z1 , y~z1+z5 , l1l2=~z3, l3=~.-z5, data=ex1.sim)
#
ex1.m11$coef  # monitor all beta parameters of model 11
#
ex1.m11$beta1 # monitor all beta parameters of lambda1 of model 11
ex1.m11$beta2 # monitor all beta parameters of lambda2 of model 11
ex1.m11$beta3 # monitor all beta parameters of lambda3 of model 11
```

---

ex2.sim *Bivpois Example 2 Dataset: Simulated Data*

Description

The data has one pair (x, y) of diagonal inflated bivariate Poisson variables and five variables $(z_1, \ldots, z_5)$ generated from $N(0, 0.12)$ distribution. Hence

$$X_i, Y_i \sim DIBP(\lambda_{1i}, \lambda_{2i}, \lambda_{3i}, p = 0.30, Poisson(2)) \text{ with}$$
$$\log \lambda_{1i} = 1.8 + 2Z_{1i} + 3Z_{3i}$$
$$\log \lambda_{2i} = 0.7 - Z_{1i} - 3Z_{3i} + 3Z_{5i}$$
$$\log \lambda_{3i} = 1.7 + Z_{1i} - 2Z_{2i} + 2Z_{3i} - 2Z_{4i}.$$

Usage

```
data(ex2.sim)
```

Format

A data frame with 100 observations on the following 7 variables.

x,y Simulated Bivariate Poisson Variables used as response

z1,z2,z3,z4,z5 Simulated N(0,0.01) explanatory variables

Details

This data is used as example one in Karlis and Ntzoufras (2004).

Source

Karlis, D. and Ntzoufras, I. (2004). Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in S. (submitted). Technical Report, Department of Statistics, Athens University of Economics and Business, Athens, Greece.

References

Karlis, D. and Ntzoufras, I. (2003). Analysis of Sports Data Using Bivariate Poisson Models. Journal of the Royal Statistical Society, D, (Statistician), 52, 381 - 393.

Examples

```
library(bivpois) # load bivpois library
data(ex2.sim)    # load ex2.sim data from bivpois library
#
# Model 1: BivPois
ex2.m1<-lm.bp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim )
# Model 2: Zero Inflated BivPois
ex2.m2<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=0)
# Model 3: Diagonal Inflated BivPois with DISCRETE(1) diagonal  distribution
ex2.m3<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=1)
# Model 4: Diagonal Inflated BivPois with DISCRETE(2) diagonal  distribution
ex2.m4<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=2)
# Model 5: Diagonal Inflated BivPois with DISCRETE(3) diagonal  distribution
ex2.m5<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=3)
# Model 6: Diagonal Inflated BivPois with DISCRETE(4) diagonal  distribution
ex2.m6<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=4)
# Model 7: Diagonal Inflated BivPois with DISCRETE(5) diagonal  distribution
ex2.m7<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=5)
# Model 8: Diagonal Inflated BivPois with DISCRETE(6) diagonal  distribution
ex2.m8<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=6)
# Model 9: Diagonal Inflated BivPois with POISSON diagonal distribution
ex2.m9<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim ,
                 distribution="poisson")
# Model 10: Diagonal Inflated BivPois with GEOMETRIC diagonal distribution
ex2.m10<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim ,
                  distribution="geometric")
#
# printing parameters of model 7
ex2.m7$beta1
ex2.m7$beta2
ex2.m7$beta3
ex2.m7$p
ex2.m7$theta
#
# printing parameters of model 9
ex2.m9$beta1
ex2.m9$beta2
ex2.m9$beta3
ex2.m9$p
ex2.m9$theta
```

`ex3.health` *Bivpois Example 3 Dataset: Health Care Data*

Description

Demand for health care in Australia data (Cameron and Trivedi, 1986). The data refer to the Australian Health survey for 1977-1978 with sample size equal to 5190.

Usage

```
data(ex3.health)
```

Format

A data frame with 5190 observations on the following 20 variables.

doctorco Number of consultations with a doctor or specialist in the past 2 weeks

prescrib Total number of prescribed medications used in past 2 days

sex 1 if female, 0 if male

age Age in years divided by 100 (measured as mid-point of 10 age groups from 15-19 years to 65-69 with 70 or more coded treated as 72)

agesq AGE squared

income Annual income in Australian dollars divided by 1000 (measured as mid-point of coded ranges Nil, <200, 200-1000, 1001-, 2001-, 3001-, 4001-, 5001-, 6001-, 7001-, 8001-10000, 10001-12000, 12001-14000, with 14001- treated as 15000).

levyplus 1 if covered by private health insurance fund for private patient in public hospital (with doctor of choice), 0 otherwise

freepoor 1 if covered by government because low income, recent immigrant, unemployed, 0 otherwise

freepera 1 if covered by government because low income, recent immigrant, unemployed, 0 otherwise

illness Number of illnesses in past 2 weeks with 5 or more coded as 5

actdays Number of days of reduced activity in past two weeks due to illness or injury

hscore General health questionnaire score using Goldberg's method. High score indicates bad health.

chcond1 1 if chronic condition(s) but not limited in activity, 0 otherwise

chcond2 1 if chronic condition(s) and limited in activity, 0 otherwise

nondocco Number of consultations with non-doctor health professionals (chemist, optician, physiotherapist, social worker, district community nurse, chiropodist or chiropractor) in the past 2 weeks

hospadmi Number of admissions to a hospital, psychiatric hospital, nursing or convalescent home in the past 12 months (up to 5 or more admissions which is coded as 5)

hospdays Number of nights in a hospital, etc. during most recent admission: taken, where appropriate, as the mid-point of the intervals 1, 2, 3, 4, 5, 6, 7, 8-14, 15-30, 31-60, 61-79 with 80 or more admissions coded as 80. If no admission in past 12 months then equals zero.

medicine Total number of prescribed and nonprescribed medications used in past 2 days

nonpresc Total number of nonprescribed medications used in past 2 days

constant Constant term

Details

Details can be found in Cameron and Trivedi (1986). This data is used as example three in Karlis and Ntzoufras (2004). In this illustration two variables are used as response: the number of consultations with a doctor or a specialist and the total number of prescribed medications used in past 2 days (doctorco, prescrib). Three variables have been used as covariates: the gender (1 female, 0 male), the age in years divided by 100 (measured as midpoints of age groups) and the annual income in Australian dollars divided by 1000 (measured as midpoint of coded ranges) sex, age, income.

Source

Cameron, A.C. and Trivedi, P.K. (1986). Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests, Journal of Applied Econometrics, 1, 29 - 54.

References

1. Cameron, A.C., Trivedi, P.K., Milne, F. and Piggott, J. (1988). A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia, Review of Economic Studies, 55, 85 - 106.

2. Cameron, A.C. and Trivedi, P.K. (1993). Tests of Independence in Parametric Models with Applications and Illustrations, Journal of Business & Economics Statistics, 11, 29 - 43.

3. Karlis, D. and Ntzoufras, I. (2004). Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in S. (submitted). Technical Report, Athens University of Economics and Business, Athens, Greece.

4. Karlis, D. and Ntzoufras, I. (2003). Analysis of Sports Data Using Bivariate Poisson Models. Journal of the Royal Statistical Society, D, (Statistician), 52, 381 - 393.

Examples

```
library(bivpois)
data(ex3.health)
# Bivariate Poisson models
#ex3.model.a<-lm.bp(doctorco~sex+age+income, prescrib~sex+age+income,
#                   data=ex3.health)
#ex3.model.b<-lm.bp(doctorco~sex+age+income, prescrib~sex+age+income, l3=~sex,
#                   data=ex3.health)
# Double Poisson model
#ex3.model.c<-lm.bp(doctorco~sex+age+income, prescrib~sex+age+income,
#                   data=ex3.health, zeroL3=TRUE)
#
# diagonal inflated models
#ex3.dibp.a<-lm.dibp(doctorco~sex+age+income, prescrib~sex+age+income,
#                    data=ex3.health, jmax=0)                         # model (a)
#ex3.dibp.b<-lm.dibp(doctorco~sex+age+income, prescrib~sex+age+income,l3=~sex,
#                    data=ex3.health, jmax=0)                         # model (b)
```

ex4.ita91 *Bivpois Example 4 Dataset: Italian Series A Football Scores for Season 1991-92*

Description

Italian Serie A football scores for season 1991-92.

Usage

```
data(ex4.ita91)
```

Format

A data frame with 306 observations on the following 4 variables.

g1 Goals scored by the home team

g2 Goals scored by the away team

team1 a factor indicating the home team with levels `Ascoli` `Atalanta` `Bari` `Cagliari` `Cremonese` `Fiorentina` `Foggia` `Genoa` `Inter` `Juventus` `Lazio` `Milan` `Napoli` `Parma` `Roma` `Sampdoria` `Torino` `Verona`

team2 a factor indicating the away team with levels `Ascoli` `Atalanta` `Bari` `Cagliari` `Cremonese` `Fiorentina` `Foggia` `Genoa` `Inter` `Juventus` `Lazio` `Milan` `Napoli` `Parma` `Roma` `Sampdoria` `Torino` `Verona`

Details

Data were originally used in Karlis and Ntzoufras (2003). The data consist of pairs of counts indicating the number of goals scored by each of the two competing teams. As covariates we have used dummy variables to model the team strength. In modelling outcomes of football games, it has been observed an excess of draws and small over-dispersion. Introducing diagonal inflated models we correct for both the over-dispersion and the excess of draws.

Source

Karlis, D. and Ntzoufras, I. (2003). Analysis of Sports Data Using Bivariate Poisson Models. Journal of the Royal Statistical Society, D, (Statistician), 52, 381 - 393.

References

Karlis, D. and Ntzoufras, I. (2004). Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in S. (submitted). Technical Report, Department of Statistics, Athens University of Economics and Business, Athens, Greece.

Examples

```
library(bivpois) # loading of bivpois library
data(ex4.ita91)  # loading ex4.ita91 data from bivpois library
#
# formula for modeling of lambda1 and lambda2
form1 <- ~c(team1,team2)+c(team2,team1)
```

```
#
# Model 1: Double Poisson
ex4.m1<-lm.bp( g1~1, g2~1, l1l2=form1, zeroL3=TRUE, data=ex4.ita91)
#
# Models 2-5: bivariate Poisson models
ex4.m2<-lm.bp(g1~1,g2~1, l1l2=form1, data=ex4.ita91)
ex4.m3<-lm.bp(g1~1,g2~1, l1l2=form1, l3=~team1, data=ex4.ita91)
ex4.m4<-lm.bp(g1~1,g2~1, l1l2=form1, l3=~team2, data=ex4.ita91)
ex4.m5<-lm.bp(g1~1,g2~1, l1l2=form1, l3=~team1+team2, data=ex4.ita91)
#
# Model 6: Zero Inflated Model
ex4.m6 <-lm.dibp(g1~1,g2~1, l1l2=form1, data=ex4.ita91, jmax=0)
#
# Models 7-11: Diagonal Inflated Bivariate Poisson Models
ex4.m7 <-lm.dibp(g1~1,g2~1, l1l2=form1, data=ex4.ita91, distribution="geometric" )
ex4.m8 <-lm.dibp(g1~1,g2~1, l1l2=form1, data=ex4.ita91, jmax=1)
ex4.m9 <-lm.dibp(g1~1,g2~1, l1l2=form1, data=ex4.ita91, jmax=2)
ex4.m10<-lm.dibp(g1~1,g2~1, l1l2=form1, data=ex4.ita91, jmax=3)
ex4.m11<-lm.dibp(g1~1,g2~1, l1l2=form1, data=ex4.ita91, distribution="poisson" )
#
# Models 12: Diagonal Inflated Double Poisson Model
ex4.m12 <- lm.dibp( g1~1,g2~1, l1l2=form1, data=ex4.ita91, distribution="poisson",
                    zeroL3=TRUE )
# ------------------------------------------------------------------------

#
# ------------------------------------------------------------------------
# monitoring parameters for model 1: Dbl Poisson
ex4.m1$coef     # all parameters
ex4.m1$beta1    # model parameters for lambda1
ex4.m1$beta2    # model parameters for lambda2.
                # All are the same as in beta1 except the intercept
ex4.m1$beta2[1] # Intercpept for lambda2.
ex4.m1$beta2[1]-ex4.m1$beta2[2] # estimated home effect

# estimating the effect for 18th level of attack (team1..team2) [Verona]
-sum(ex4.m1$coef[ 2:18])
# estimating the effect for 18th level of defence(team2..team1) [Verona]
-sum(ex4.m1$coef[19:35])
#
# ------------------------------------------------------------------------
# monitoring parameters for model 2: BivPoisson(lamdba1,lambda2,constant lamdba3)
#
#
# monitoring parameters for model 1: Dbl Poisson
ex4.m2$beta1      # model parameters for lambda1
ex4.m2$beta2      # model parameters for lambda2.
                  # All are the same as in beta1 except the intercept
ex4.m2$beta3      # model parameters for lambda3 (Here only the intercept)
ex4.m2$beta2[1]   # Intercpept for lambda2.
ex4.m2$beta2[1]-ex4.m2$beta2[2] # estimated home effect

# estimating the effect for 18th level of attack (team1..team2) [Verona]
-sum(ex4.m2$coef[ 2:18])
# estimating the effect for 18th level of defence(team2..team1) [Verona]
-sum(ex4.m2$coef[19:35])
#
```

```
# ---------------------------------------------------------------------------

# ---------------------------------------------------------------------------
# monitoring parameters for model 8: Biv.Poisson with Dis(1) diagonal distribution
#
#
# monitoring parameters for model 1: Dbl Poisson
ex4.m8$beta1      # model parameters for lambda1
ex4.m8$beta2      # model parameters for lambda2.
                  # All are the same as in beta1 except the intercept
ex4.m8$beta3      # model parameters for lambda3. Here beta3 has only the intercept
ex4.m8$beta2[1]   # Intercpept for lambda2.
ex4.m8$beta2[1]-ex4.m8$beta2[2] # estimated home effect

# estimating the effect for 18th level of attack (team1..team2) [Verona]
-sum(ex4.m8$coef[ 2:18])
# estimating the effect for 18th level of defence(team2..team1) [Verona]
-sum(ex4.m8$coef[19:35])

ex4.m8$beta3                  # parameters for lambda3 (here the intercept)
exp(ex4.m8$beta3)             # lambda3 (here constant)
ex4.m8$diagonal.distribution # printing details for the diagonal distribution
ex4.m8$p                      # mixing proportion
ex4.m8$theta                  # printing theta parameters
```

`lm.bp` *General Bivariate Poisson Model*

Description

Produces a "list" object which gives details regarding the fit of a bivariate Poisson regression model of the form

$X_i, Y_i \sim BP(\lambda_{1i}, \lambda_{2i}, \lambda_{3i})$ with

$\log \underline{\lambda}_1 = \mathbf{w}_1 \underline{\beta}_1$, $\log \underline{\lambda}_2 = \mathbf{w}_2 \underline{\beta}_2$ and $\log \underline{\lambda}_3 = \mathbf{w}_3 \underline{\beta}_3$ for $i = 1, 2, \ldots, n$;

where

*** n is the sample size

*** $\underline{\lambda}_k = (\lambda_{k1}, \lambda_{k2}, \ldots, \lambda_{kn})^T$ for $k = 1, 2, 3$ are vectors of length n with the estimated lambda for each observation

*** $\mathbf{w_1}, \mathbf{w_2}$ are $n \times p$ data matrices containing explanatory variables for λ_1 and λ_2.

*** $\mathbf{w_3}$ is a $n \times p_2$ data matrix containing explanatory variables for λ_3.

*** $\underline{\beta}_1, \underline{\beta}_2, \underline{\beta}_3$ are parameter vectors used in the linear predictors of λ_1, λ_2 and λ_3.

Usage

```
lm.bp( l1, l2, l1l2=NULL, l3=~1, data, common.intercept=FALSE,
       zeroL3=FALSE, maxit=300, pres=1e-8, verbose=getOption("verbose") )
```

Arguments

`l1` Formula of the form "$x \sim X_1 + \ldots + X_p$" for parameters of $\log \lambda_1$.

`l2` Formula of the form "$y \sim X_1 + \ldots + X_p$" for parameters of $\log \lambda_2$.

`l1l2` Formula of the form "$\sim X_1 + \ldots + X_p$" for the common parameters of $\log \lambda_1$ and $\log \lambda_2$. If the explanatory variable is also found on `l1` and/or `l2` then a model using interaction type parameters is fitted (one parameter common for both predictors [main effect] and differences from this for the other predictor [interaction type effect]). Special terms of the form "`c(X1,X2)`" can be also used here. These terms imply common parameters of λ_1 and λ_2 on different variables. For example if `c(x1,x2)` is used then use the same beta for the effect of x_1 on $\log \lambda_1$ and the effect of x_2 on $\log \lambda_2$. For details see example 4 - dataset `ex4.ita91`.

`l3` Formula of the form "$\sim X_1 + \ldots + X_p$" for the parameters of $\log \lambda_3$.

`data` Data frame containing the variables in the model.

`common.intercept`
Logical function specifying whether a common intercept on $\log \lambda_1$ and $\log \lambda_2$ should be used. The default value is `FALSE`.

`zeroL3` Logical argument controlling whether λ_3 should be set equal to zero (therefore fits a double Poisson model).

`maxit` Maximum number of EM steps. Default value is 300 iterations.

`pres` Precision used in stopping the EM algorithm. The algorithm stops when the relative log-likelihood difference is lower than the value of pres.

`verbose` Logical argument controlling whether beta parameters will we printed while EM runs. Default value is taken equal to the value of `options()$verbose`. If `verbose=FALSE` then only the iteration number, the loglikelihood and its relative difference from the previous iteration are printed. If `verbose=TRUE` then the model parameters β_1, β_2 and β_3 are additionally printed

Value

A list object returned with the following variables.

`coefficients` Estimates of the model parameters for β_1, β_2 and β_3. When a factor is used then its default set of constraints is used.

`fitted.values` Data frame with n lines and 2 columns containing the fitted values for x and y. For the bivariate Poisson model the fitted values are given by $\lambda_1 + \lambda_3$ and $\lambda_2 + \lambda_3$ respectively.

`residuals` Data frame with n lines and 2 columns containing the residuals of the model for x and y. For the bivariate Poisson model the residual values are given by $x - \lambda_1 - \lambda_3$ and $y - \lambda_2 - \lambda_3$ respectively.

`beta1,beta2, beta3`
Vectors β_1, β_2 and β_3 containing the coefficients involved in the linear predictors of λ_1, λ_2 and λ_3 respectively. When `zeroL3=TRUE` then `beta3` is not calculated.

`lambda1, lambda2`
Vectors of length n containing the estimated λ_1 and λ_2 for each observation

`lambda3`	vector containing the values of λ_3. If `zeroL3=TRUE` then `lambda3` is equal to zero and is not provided.
`loglikelihood`	Maximized log-likelihood of the fitted model. This is given in a vector form (one value per iteration). Using this vector we can monitor the log-likelihood evolution in each EM step.
`AIC, BIC`	AIC and BIC of the model. Values are also provided for the double Poisson model and the saturated model.
`parameters`	Number of parameters.
`iterations`	Number of iterations.
`call`	Argument providing the exact calling details of the `lm.bp` function.

Author(s)

1. Dimitris Karlis, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨karlis@aueb.gr⟩ .

2. Ioannis Ntzoufras, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨ntzoufras@aueb.gr⟩ .

References

1. Karlis, D. and Ntzoufras, I. (2004). Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in S. (submitted). Technical Report, Department of Statistics, Athens University of Economics and Business, Athens, Greece.

2. Karlis, D. and Ntzoufras, I. (2003). Analysis of Sports Data Using Bivariate Poisson Models. Journal of the Royal Statistical Society, D, (Statistician), 52, 381 - 393.

See Also

`pbivpois`, `simple.bp`, `lm.dibp`.

Examples

```
data(ex1.sim)
# Fit Double and Bivariate Poisson models ()
#
# Model 2: DblPoisson(l1, l2)
ex1.m2<-lm.bp(x~1 , y~1 , data=ex1.sim, zeroL3=TRUE)
# Model 3: BivPoisson(l1, l2, l3); same as simple.bp(ex1.sim$x, ex1.sim$y)
ex1.m3<-lm.bp(x~1 , y~1 , data=ex1.sim)
# Model 4: DblPoisson (l1=Full, l2=Full)
ex1.m4<-lm.bp(x~. , y~. , data=ex1.sim, zeroL3=TRUE)
# Model 5: BivPoisson(l1=full, l2=full, l3=constant)
ex1.m5<-lm.bp(x~. , y~. , data=ex1.sim)
# Model 6: DblPois(l1,l2)
ex1.m6<-lm.bp(x~z1 , y~z1+z5 , l1l2=~z3, data=ex1.sim, zeroL3=TRUE)
# Model 7: BivPois(l1,l2,l3=constant)
ex1.m7<-lm.bp(x~z1 , y~z1+z5 , l1l2=~z3, data=ex1.sim)
# Model 8: BivPoisson(l1=full, l2=full, l3=full)
ex1.m8<-lm.bp(x~. , y~. , l3=~., data=ex1.sim)
# Model 9: BivPoisson(l1=full, l2=full, l3=z1+z2+z3+z4)
ex1.m9<-lm.bp(x~. , y~. , l3=~.-z5, data=ex1.sim)
# Model 10: BivPoisson(l1, l2, l3=full)
ex1.m10<-lm.bp(x~z1 , y~z1+z5 , l1l2=~z3, l3=~., data=ex1.sim)
```

```
# Model 11: BivPoisson(l1, l2, l3= z1+z2+z3+z4)
ex1.m11<-lm.bp(x~z1 , y~z1+z5 , l1l2=~z3, l3=~.-z5, data=ex1.sim)
```

`lm.dibp` *General Diagonal Inflated Bivariate Poisson Model*

Description

Produces a "list" object which gives details regarding the fit of a bivariate diagonal inflated Poisson regression model of the form

$(X_i, Y_i) \sim DIBP(\lambda_{1i}, \lambda_{2i}, \lambda_{3i}, D(\theta))$ which is equivalent to

$(Xi, Yi) \sim (1-p)BP(x_i, y_i|\lambda_{1i}, \lambda_{2i}, \lambda_{3i})$ if $x_i \neq y_i$

$(Xi, Yi) \sim (1-p)BP(x_i, y_i|\lambda_{1i}, \lambda_{2i}, \lambda_{3i}) + pD(x_i|\theta)$ if $x_i = y_i$ with

$\log \underline{\lambda}_1 = \mathbf{w}_1\underline{\beta}$, $\log \underline{\lambda}_2 = \mathbf{w}_2\underline{\beta}$ and $\log \underline{\lambda}_3 = \mathbf{w}_3\underline{\beta}_3$ for $i = 1, 2, \ldots, n$;

where

*** n is the sample size,

*** $\underline{\lambda}_k = (\lambda_{k1}, \lambda_{k2}, \ldots, \lambda_{kn})^T$ for $k = 1, 2, 3$ are vectors of length n containing the estimated lambda for each observation,

*** $\mathbf{w}_1$, $\mathbf{w}_2$ are $n \times p$ data matrices containing explanatory variables for λ_1 and λ_2,

*** $\mathbf{w}_3$ are $n \times p_2$ data matrix containing explanatory variables for λ_3.

*** $\underline{\beta}$ is a vector of length p which is common for λ_1 and λ_2 in order to allow for common effects.

*** $\underline{\beta}_3$ vector of length p_2.

*** $D(\theta)$ is a discrete distribution with parameter vector θ used to inflate the diagonal.

*** p is the mixing proportion.

Usage

```
lm.dibp( l1, l2, l1l2=NULL, l3=~1, data, common.intercept=FALSE,
         zeroL3 = FALSE, distribution = "discrete", jmax = 2, maxit = 300,
         pres = 1e-08, verbose=getOption("verbose") )
```

Arguments

`l1`	Formula of the form "$x \sim X_1 + \ldots + X_p$" for parameters of $\log \lambda_1$.
`l2`	Formula of the form "$y \sim X_1 + \ldots + X_p$" for parameters of $\log \lambda_2$.
`l1l2`	Formula of the form "$\sim X_1 + \ldots + X_p$" for the common parameters of $\log \lambda_1$ and $\log \lambda_2$. If the explanatory variable is also found on `l1` and/or `l2` then a model using interaction type parameters is fitted (one parameter common for both predictors [main effect] and differences from this for the other predictor [interaction type effect]). Special terms of the form "`c(X1,X2)`" can be also used here. These terms imply common parameters of λ_1 and λ_2 on different variables. For example if `c(x1,x2)` is used then use the same beta for the effect of x_1 on $\log \lambda_1$ and the effect of x_2 on $\log \lambda_2$. For details see example 4 - dataset `ex4.ita91`.

`l3` Formula of the form "$\sim X_1 + \ldots + X_p$" for the parameters of $\log \lambda_3$.

`data` Data frame containing the variables in the model.

`common.intercept`
: Logical function specifying whether a common intercept on $\log \lambda_1$ and $\log \lambda_2$ should be used. The default value is `FALSE`.

`zeroL3` Logical argument controlling whether λ_3 should be set equal to zero (therefore fits a double Poisson model).

`distribution` Specifies the type of inflated distribution; ="`discrete`": Discrete(J=jmax), ="`poisson`" : Poisson(θ) ="`geometric`": Geometric(θ).

`jmax` Number of parameters used in *Discrete* distribution. This argument is not used for the Poisson or the Geometric distributions are used as for the inflation of the diagonal.

`maxit` Maximum number of EM steps. Default value is 300 iterations.

`pres` Precision used in stopping the EM algorithm. The algorithm stops when the relative log-likelihood difference is lower than the value of pres.

`verbose` Logical argument controlling whether beta parameters will we printed while EM runs. Default value is taken equal to the value of `options()$verbose`. If `verbose=FALSE` then only the iteration number, the loglikelihood and its relative difference from the previous iteration are printed. If `verbose=TRUE` then the model parameters β_1, β_2 and β_3 are additionally printed

Value

A list object returned with the following variables.

`coefficients` Estimates of the model parameters for β_1, β_2 and β_3. When a factor is used then its default set of constraints is used.

`fitted.values` Data frame with n lines and 2 columns containing the fitted values for x and y.

`residuals` Data frame with n lines and 2 columns containing the residuals of the model for x and y.

`beta1,beta2, beta3`
: Vectors β_1, β_2 and β_3 containing the coefficients involved in the linear predictors of λ_1, λ_2 and λ_3 respectively. When `zeroL3=TRUE` then `beta3` is not calculated.

`lambda1, lambda2`
: Vectors of length n containing the estimated λ_1 and λ_2 for each observation

`lambda3` vector containing the values of λ_3. If `zeroL3=TRUE` then `lambda3` is equal to zero and is not provided.

`loglikelihood` Maximized log-likelihood of the fitted model. This is given in a vector form (one value per iteration). Using this vector we can monitor the log-likelihood evolution in each EM step.

`AIC, BIC` AIC and BIC of the model. Values are also provided for the double Poisson model and the saturated model.

`diagonal.distribution`
: label for the diagonal inflated distribution used.

`p` mixing proportion.

`theta`	Parameter vector of the diagonal distribution. For `discrete` distribution theta has length equal to jmax with θ_i =`theta[i]` and $\theta_0 = 1 - \sum_{i=1}^{JMAX} \theta_i$; for the Poisson distribution theta is the mean; for the Geometric distribution theta is the success probability.
`parameters`	Number of parameters.
`iterations`	Number of iterations.
`call`	Argument providing the exact calling details of the `lm.dibp` function.

Author(s)

1. Dimitris Karlis, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨karlis@aueb.gr⟩ .

2. Ioannis Ntzoufras, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨ntzoufras@aueb.gr⟩ .

References

1. Karlis, D. and Ntzoufras, I. (2004). Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in S. (submitted). Technical Report, Department of Statistics, Athens University of Economics and Business, Athens, Greece.

2. Karlis, D. and Ntzoufras, I. (2003). Analysis of Sports Data Using Bivariate Poisson Models. Journal of the Royal Statistical Society, D, (Statistician), 52, 381 - 393.

See Also

`pbivpois`, `simple.bp`, `lm.bp`.

Examples

```
data(ex2.sim)
#
# Model 1: BivPois
ex2.m1<-lm.bp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim )
# Model 2: Zero Inflated BivPois
ex2.m2<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=0)
# Model 3: Diagonal Inflated BivPois with DISCRETE(1) diagonal distribution
ex2.m3<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=1)
# Model 4: Diagonal Inflated BivPois with DISCRETE(2) diagonal distribution
ex2.m4<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=2)
# Model 5: Diagonal Inflated BivPois with DISCRETE(3) diagonal distribution
ex2.m5<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=3)
# Model 6: Diagonal Inflated BivPois with DISCRETE(4) diagonal distribution
ex2.m6<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=4)
# Model 7: Diagonal Inflated BivPois with DISCRETE(5) diagonal distribution
ex2.m7<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=5)
# Model 8: Diagonal Inflated BivPois with DISCRETE(6) diagonal distribution
ex2.m8<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim , jmax=6)
# Model 9: Diagonal Inflated BivPois with POISSON diagonal distribution
ex2.m9<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim ,
                 distribution="poisson")
# Model 10: Diagonal Inflated BivPois with GEOMETRIC diagonal distribution
ex2.m10<-lm.dibp( x~z1 , y~z1+z5, l1l2=~z3, l3=~.-z5, data=ex2.sim ,
                  distribution="geometric")
#
```

```
# printing parameters of model 7
ex2.m7$beta1
ex2.m7$beta2
ex2.m7$beta3
ex2.m7$p
ex2.m7$theta
#
# printing parameters of model 9
ex2.m9$beta1
ex2.m9$beta2
ex2.m9$beta3
ex2.m9$p
ex2.m9$theta
```

---

newnamesbeta *Internal Function of BIVPOIS Package*

Description

This function was made only for internal use in Bivpois package and it should not be called separately

Author(s)

1. Dimitris Karlis, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨karlis@aueb.gr⟩ .

2. Ioannis Ntzoufras, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨ntzoufras@aueb.gr⟩ .

---

pbivpois *Probability Function of the Bivariate Poisson Distribution*

Description

Returns the probability the probability (or the log) of the bivariate poisson distribution for values x and y.

Usage

```
pbivpois(x, y=NULL, lambda = c(1, 1, 1), log = FALSE)
```

Arguments

`x`	Matrix or Vector containing the data. If `x` is a matrix then we consider as x the first column and y the second column. Additional columns and `y` are ignored.
`y`	Vector containing the data of y. It is used only if `x` is also a vector. Vectors `x` and `y` should be of equal length.
`lambda`	Vector (of length 3) containing values of the parameters λ_1, λ_2 and λ_3 of the bivariate Poisson distribution.
`log`	Logical argument for calculating the log probability or the probability function. The default value is `FALSE`.

Details

This function evaluates the probability function (or the log) of the bivariate Poisson distribution with parameters λ_1, λ_2 and λ_3. Much faster than `bivpois.table` since it avoid 'for-loops' and does not use recursive relations.

Value

A vector of values of the probabilities of $PD(\lambda_1, \lambda_2, \lambda_3)$ evaluated at (x, y) when `log=FALSE` or the log-probabilities of $PD(\lambda_1, \lambda_2, \lambda_3)$ evaluated at (x, y) when `log=TRUE`.

Author(s)

1. Dimitris Karlis, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨karlis@aueb.gr⟩ .

2. Ioannis Ntzoufras, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨ntzoufras@aueb.gr⟩ .

References

1. Karlis, D. and Ntzoufras, I. (2004). Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in S. (submitted). Technical Report, Department of Statistics, Athens University of Economics and Business, Athens, Greece.

2. Karlis, D. and Ntzoufras, I. (2003). Analysis of Sports Data Using Bivariate Poisson Models. Journal of the Royal Statistical Society, D, (Statistician), 52, 381 - 393.

See Also

`bivpois.table`, `simple.bp`, `lm.bp`, `lm.dibp`.

Examples

```
# probability function of (x,y)=(3,1) for lambda_1=1, lambda_2=1, lambda_3=1
pbivpois(3, 1)
# probability function of (x,y)=(3,1) for lambda_1=3, lambda_2=1, lambda_3=1
pbivpois(3, 1, lambda=c(3,1,1))
# log-probability function of (x,y)=(3,1) for lambda_1=1, lambda_2=1, lambda_3=1
pbivpois(3, 1, lambda=c(3,1,1), log=TRUE)
#
# evaluates f(1,1), f(1,3) and f(3,1) for PD(2,1,0.2)
pbivpois( c(1,3,1), c(1,1,3), c( 2,1,0.2 ) )
# same as above
pbivpois( cbind(c(1,3,1), c(1,1,3)), lambda=c( 2,1,0.2 ) )
```

`simple.bp` *Simple Bivariate Poisson Model*

Description

Produces a "list" object which gives details regarding the fit of a simple bivariate Poisson model of the form $(X, Y) \sim BP(\lambda_1, \lambda_2, \lambda_3)$.

Usage

```
simple.bp(x, y, ini3 = 1, maxit = 300, pres = 1e-08)
```

Arguments

`x, y`	vectors containing the data.
`ini3`	Initial value for λ_3.
`maxit`	Maximum number of EM steps.
`pres`	Precision used in log-likelihood improvement.

Details

During the run of the algorithm the following details are printed: the iteration number, lambda1, lambda2, lambda3, the log-likelihood and the relative difference of the log-likelihood.

Value

A list object returned with the following variables.

`lambda`	Vector with parameter values λ_1 , λ_2, λ_3
`loglikelihood`	laximized log-likelihood of the fitted model. This is given in a vector form (one value per iteration).Using this we may monitor the log-likelihood improvement and how EM algorithm works.
`AIC, BIC`	AIC and BIC of the model. Values are also given for the double Poisson model and the saturated model.
`parameters`	Number of parameters.
`iterations`	Number of iterations.

Author(s)

1. Dimitris Karlis, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨karlis@aueb.gr⟩ .

2. Ioannis Ntzoufras, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨ntzoufras@aueb.gr⟩ .

References

1. Karlis, D. and Ntzoufras, I. (2004). Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in S. (submitted). Technical Report, Department of Statistics, Athens University of Economics and Business, Athens, Greece.

2. Karlis, D. and Ntzoufras, I. (2003). Analysis of Sports Data Using Bivariate Poisson Models. Journal of the Royal Statistical Society, D, (Statistician), 52, 381 - 393.

See Also

`pbivpois`, `lm.bp`, `lm.dibp`

Examples

```
#
# Generation of BP(1,2,3) data
x3<-rpois(100, 3)
x1<-rpois(100, 1)+x3
x2<-rpois(100, 2)+x3
#
# fits the model
x<-simple.bp(x1, x2)
#
# Monitors parameters
x$lambda1
x$lambda2
x$lambda3
```

`splitbeta` *Internal Function of BIVPOIS Package*

Description

This function was made only for internal use in Bivpois package and it should not be called separately

Author(s)

1. Dimitris Karlis, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨karlis@aueb.gr⟩ .

2. Ioannis Ntzoufras, Department of Statistics, Athens University of Economics and Business, Athens, Greece, ⟨ntzoufras@aueb.gr⟩ .

Index