Package 'crimCV'

September 9, 2023

Version 1.0.0

Date 2023-09-06

Title Group-Based Modelling of Longitudinal Data

Author Jason D. Nielsen <drjdnielsen@gmail.com>

Maintainer Jason D. Nielsen <jdn@math.carleton.ca>

Depends R (>= 3.5.0), splines

Description A finite mixture of Zero-Inflated Poisson (ZIP) models for analyzing criminal trajectories.

License GPL (>= 2)

NeedsCompilation yes

Repository CRAN

Date/Publication 2023-09-09 19:30:02 UTC

R topics documented:

---

crimCV   *Fits finite mixtures of Zero-inflated Poisson models*

Description

This software fits finite mixtures of ZIP models to longitudinal count data.

Usage

```
crimCV(Dat,ng,dpolyp=3,dpolyl=3,model="ZIPt",rcv=FALSE,init=20,Risk=NULL)
```

Arguments

Dat	A matrix of the number of criminal offenses. Each row contains the number of criminal offenses committed by a particular individual for all time intervals observed (columns). A negative number is interpreted as a missing value.
ng	The number of groups to use in the analysis.
dpolyp	The degree of the polynomial in the log-linear component. Defaults to 3.
dpolyl	The degree of the polynomial in the logistic component (ZIP model only). Defaults to 3.
model	Either "ZIP" for the full ZIP or "ZIPt" for the ZIP(tau) sub-group model. Defaults to "ZIPt".
rcv	Set to TRUE to compute the CVE. Defaults to FALSE.
init	This controls how rigorously the initialization procedure searches for good starting values of the parameters. Larger values increase the odds of finding the true global solution but can dramatically increase the time required to fit the model. The default is set to 20 and from experimentation this seems to be a reasonable value for ng<=5. Note: As the number of groups (ng) gets larger finding the optimal global maximum of the likelihood becomes progressively more challenging.
Risk	A matrix of the same dimension as Dat with the time-at-risk correction. Defaults to 1 for all elements (i.e. at risk for the full time period).

Value

A list of class "dmZIP" or "dmZIPt".

Author(s)

Jason D. Nielsen

References

J.D. Nielsen, J.S. Rosenthal, Y. Sun, D.M. Day, I. Bevc, and T. Duchesne (2011). Group-based Criminal Trajectory Analysis using Cross-Validation Criteria. A draft of the manuscript is available at *http://www.probability.ca/jeff/research.html*.

Examples

```
# Loads crimCV into the interpreter
library(crimCV)

# Load the "divide-and-round" TO1 dataset
data(TO1adj)
# Fit a 2 component ZIP(tau) model of degree 2.  Here the CVE is not
```

```
# calculated and only ~1/4 of the data is used so that the code will
# run quickly enough to satisfy CRAN's package policies.  To compute
# CVE run as:
# out1<-crimCV(TO1adj,2,dpolyp=2,rcv=TRUE)
subTO1adj<-TO1adj[1:100,]
out1<-crimCV(subTO1adj,2,dpolyp=2,init=5)
# Plot the component trajectories
plot(out1)
# Print out some useful output
summary(out1)
```

TO1   *Adjudicated Toronto Youth Data (Sample 1)*

Description

Number of criminal unique court contacts for 378 individuals in Toronto, Ontario, Canada.

Usage

```
data(TO1)
```

Format

A 378 by 31 matrix where each row contains the number of unique court contacts per year for an individual from the age of 8 to 38.

TO1adj   *Adjusted Adjudicated Toronto Youth Data (Sample 1)*

Description

Number of unique court contacts for 378 individuals in Toronto, Ontario, Canada adjusted for time-at-risk by the "divide and round" approach.

Usage

```
data(TO1adj)
```

Format

A 378 by 31 matrix where each row contains the number of unique court contacts per year corrected for time-at-risk for an individual from the age of 8 to 38.

TO1Risk   *Time-at-risk for the Adjudicated Toronto Youth Data (Sample 1)*

Description

Time-at-risk per year for 378 individuals in the Toronto, Ontario, Canada.

Usage

```
data(TO1Risk)
```

Format

A 378 by 31 matrix where each row contains the time-at-risk per year for an individual from the age of 8 to 38.

TO2   *Adjudicated Toronto Youth Data (Sample 2)*

Description

Number of unique court contacts for 386 individuals in Toronto, Ontario, Canada.

Usage

```
data(TO2)
```

Format

A 386 by 30 matrix where each row contains the number of unique court contacts per year for an individual from the age of 9 to 38.

TO2adj   *Adjusted Adjudicated Toronto Youth Data (Sample 2)*

Description

Number of unique court contacts for 386 individuals in Toronto, Ontario, Canada adjusted for time-at-risk by the "divide and round" approach.

Usage

```
data(TO2adj)
```

Format

A 386 by 30 matrix where each row contains the number of unique court contacts per year corrected for time-at-risk for an individual from the age of 9 to 38.

TO2Risk   *Time-at-risk for the Adjudicated Toronto Youth Data (Sample 2)*

Description

Time-at-risk for 386 individuals in the Toronto, Ontario, Canada.

Usage

```
data(TO2Risk)
```

Format

A 386 by 30 matrix where each row contains the number of criminal offenses per year for an individual from the age of 9 to 38.

Index