Package: poLCAParallel (via r-universe)

May 16, 2026

Type Package

Title Polytomous Variable Latent Class Analysis Parallel

Version 1.2.7

Depends R (>= 4.1.2)

Imports Rcpp (>= 1.0.7), scatterplot3d, MASS, parallel, poLCA

LinkingTo Rcpp, RcppArmadillo (>= 0.12.4.1.0)

Description A 'C++' reimplementation of 'poLCA' - latent class analysis and latent class regression models for polytomous outcome variables, also known as latent structure analysis. It attempts to reproduce results and be as similar as possible to the original code, while running faster, especially with multiple repetitions, by utilising multiple threads. Further reading is available on the Queen Mary, University of London, IT Services Research blog <https://blog.hpc.qmul.ac.uk/speeding_up_r_packages/>.

License GPL-2

URL https://github.com/QMUL/poLCAParallel

LazyLoad yes

NeedsCompilation yes

RoxygenNote 7.3.3

Encoding UTF-8

Suggests testthat (>= 3.0.0), roxygen2, usethis

Roxygen list(markdown = TRUE)

Config/testthat/edition 3

SystemRequirements C++20

Repository https://qmul.r-universe.dev

Date/Publication 2026-03-12 10:40:42 UTC

RemoteUrl https://github.com/qmul/polcaparallel

RemoteRef HEAD

RemoteSha 2425fa131753ed608bffd4d579060662b060a936

Contents

`carcinoma` *Diagnoses of carcinoma (sample data)*

Description

Dichotomous ratings by seven pathologists of 118 slides for the presence or absence of carcinoma in the uterine cervix. Pathologists are labeled `A` through `G`. There were 20 different observed response patterns. This data set appears in Agresti (2002, p. 542) as Table 13.1.

Usage

```
data(carcinoma)
```

Format

A data frame with 118 observations on 7 variables representing pathologist ratings with 1 denoting "no" and 2 denoting "yes".

Source

Agresti, Alan. 2002. *Categorical Data Analysis, second edition*. Hoboken: John Wiley & Sons.

Examples

```
##
## Replication of latent class models in Agresti (2002, p. 543),
## Table 13.2 and Table 13.3.
##
data(carcinoma)
f <- cbind(A,B,C,D,E,F,G)~1
lca2 <- poLCA(f,carcinoma,nclass=2) # log-likelihood: -317.2568
lca3 <- poLCA(f,carcinoma,nclass=3) # log-likelihood: -293.705
lca4 <- poLCA(f,carcinoma,nclass=4,nrep=10,maxiter=5000) # log-likelihood: -289.2858
```

cheating *GPA and chronic cheating (sample data)*

Description

Dichotomous responses by 319 undergraduates to four questions about cheating behavior, and each student's academic GPA.

Students responded either (1) no or (2) yes as to whether they had ever lied to avoid taking an exam (LIEEXAM), lied to avoid handing a term paper in on time (LIEPAPER), purchased a term paper to hand in as their own or had obtained a copy of an exam prior to taking the exam (FRAUD), or copied answers during an exam from someone sitting near to them (COPYEXAM).

The GPA variable is partitioned into five groups: (1) 2.99 or less; (2) 3.00-3.25; (3) 3.26-3.50; (4) 3.51-3.75; (5) 3.76-4.00.

This data set appears in Dayton (1998, pp. 33 and 85) as Tables 3.4 and 7.1.

Usage

```
data(cheating)
```

Format

A data frame with 319 observations on 5 variables. Note: GPA data were not available for four students who reported never cheating.

Source

Dayton, C. Mitchell. 1998. *Latent Class Scaling Analysis*. Thousand Oaks, CA: SAGE Publications.

Examples

```
##
## Replication of latent class models in Dayton (1998)
##
## Example 1. Two-class LCA. (Table 3.3, p. 32)
##
data(cheating)
f <- cbind(LIEEXAM,LIEPAPER,FRAUD,COPYEXAM)~1
ch2 <- poLCA(f,cheating,nclass=2) # log-likelihood: -440.0271

##
## Example 2. Two-class latent class regression using
## GPA as a covariate to predict class membership as
## "cheaters" vs. "non-cheaters".
## (Table 7.1, p. 85, and Figure 7.1, p. 86)
##
f2 <- cbind(LIEEXAM,LIEPAPER,FRAUD,COPYEXAM)~GPA
ch2c <- poLCA(f2,cheating,nclass=2) # log-likelihood: -429.6384
```

```
GPAmat <- cbind(1,c(1:5))
exb <- exp(GPAmat %*% ch2c$coeff)
matplot(c(1:5),cbind(1/(1+exb),exb/(1+exb)),type="l",lwd=2,
        main="GPA as a predictor of persistent cheating",
        xlab="GPA category, low to high",
        ylab="Probability of latent class membership")
text(1.7,0.3,"Cheaters")
text(1.7,0.7,"Non-cheaters")

##
## Compare results from Example 1 to Example 2.
## Non-simultaneous estimation of effect of GPA on latent class
## membership biases the estimated effect in Example 1.
##
cheatcl <- which.min(ch2$P)
predcc <- sapply(c(1:5),function(v) mean(ch2$posterior[cheating$GPA==v,cheatcl],na.rm=TRUE))
## Having run Ex.2, add to plot:
matplot(c(1:5),cbind(1-predcc,predcc),type="l",lwd=2,add=TRUE)
text(4,0.14,"Cheaters\n (non-simul. estimate)")
text(4,0.87,"Non-cheaters\n (non-simul. estimate)")
```

`election` *2000 National Election Studies survey (sample data)*

Description

Survey data from the 2000 American National Election Study. Two sets of six questions with four responses each, asking respondents' opinions of how well various traits (moral, caring, knowledgable, good leader, dishonest, intelligent) describe presidential candidates Al Gore and George W. Bush. The responses are (1) Extremely well; (2) Quite well; (3) Not too well; (4) Not well at all. Many respondents have varying numbers of missing values on these variables.

The data set also includes potential covariates `VOTE3`, the respondent's 2000 vote choice (when asked); `AGE`, the respondent's age; `EDUC`, the respondent's level of education; `GENDER`, the respondent's gender; and `PARTY`, the respondent's Democratic-Republican partisan identification.

`VOTE3` is coded as (1) Gore; (2) Bush; (3) Other.

`EDUC` is coded as (1) 8 grades or less; (2) 9-11 grades, no further schooling; (3) High school diploma or equivalency; (4) More than 12 years of schooling, no higher degree; (5) Junior or community college level degree; (6) BA level degrees, no advanced degree; (7) Advanced degree.

`GENDER` is coded as (1) Male; (2) Female.

`PARTY` is coded as (1) Strong Democrat; (2) Weak Democrat; (3) Independent-Democrat; (4) Independent-Independent; (5) Independent-Republican; (6) Weak Republican; (7) Strong Republican.

Usage

```
data(election)
```

Format

A data frame with 1785 observations on 17 survey variables. Of these, 1311 individuals provided responses on all twelve candidate evaluations.

Source

The National Election Studies (https://www.electionstudies.org/). THE 2000 NATIONAL ELECTION STUDY [dataset]. Ann Arbor, MI: University of Michigan, Center for Political Studies [producer and distributor].

Examples

```
# Latent class models with one (loglinear independence) to three classes
data(election)
f <- cbind(MORALG,CARESG,KNOWG,LEADG,DISHONG,INTELG,
           MORALB,CARESB,KNOWB,LEADB,DISHONB,INTELB)~1
nes1 <- poLCA(f,election,nclass=1)  # log-likelihood: -18647.31
nes2 <- poLCA(f,election,nclass=2)  # log-likelihood: -17344.92
nes3 <- poLCA(f,election,nclass=3)  # log-likelihood: -16714.66

# Three-class model with a single covariate (party)
f2a <- cbind(MORALG,CARESG,KNOWG,LEADG,DISHONG,INTELG,
             MORALB,CARESB,KNOWB,LEADB,DISHONB,INTELB)~PARTY
nes2a <- poLCA(f2a,election,nclass=3,nrep=5)    # log-likelihood: -16222.32
pidmat <- cbind(1,c(1:7))
exb <- exp(pidmat %*% nes2a$coeff)
matplot(c(1:7),(cbind(1,exb)/(1+rowSums(exb))),ylim=c(0,1),type="l",
    main="Party ID as a predictor of candidate affinity class",
    xlab="Party ID: strong Democratic (1) to strong Republican (7)",
    ylab="Probability of latent class membership",lwd=2,col=1)
text(5.9,0.35,"Other")
text(5.4,0.7,"Bush affinity")
text(1.8,0.6,"Gore affinity")
```

---

gss82 *1982 General Social Survey (sample data)*

Description

Attitudes towards survey taking across two dichotomous and two trichotomous items among 1202 white respondents to the 1982 General Social Survey. Respondents give their opinion of the purpose of surveys (PURPOSE; good/depends/waste of time and money), the accuracy of surveys (ACCURACY; mostly true/not true), their understanding of survey questions (UNDERSTA; good/fair, poor), and how well they cooperated with the interviewer (COOPERAT; interested/cooperative/impatient, hostile). This data set appears in McCutcheon (1987, p. 30) as Table 3.1.

Usage

```
data(gss82)
```

Format

A data frame with 1202 observations on 4 survey variables.

Source

McCutcheon, A.L. 1987. *Latent class analysis*. Newbury Park: SAGE Publications.

Examples

```
data(gss82)
f <- cbind(PURPOSE,ACCURACY,UNDERSTA,COOPERAT)~1
gss.lc2 <- poLCA(f,gss82,nclass=2) # log-likelihood = -2783.268

# Could also try:
# gss.lc3 <- poLCA(f,gss82,nclass=3,maxiter=3000,nrep=10) # log-likelihood = -2754.545
# gss.lc4 <- poLCA(f,gss82,nclass=4,maxiter=15000,nrep=10,tol=1e-7) # log-likelihood = -2746.621
```

---

values *Universalistic vs. particularistic values (sample data)*

Description

Dichotomous survey responses from 216 respondents to four questions (`A`, `B`, `C`, `D`) measuring tendencies towards "universalistic" or "particularistic" values. This data set appears in Goodman (2002, p. 14) as Table 4, and previously appeared in Goodman (1974) and Stouffer and Toby (1951).

Usage

```
data(values)
```

Format

A data frame with 216 observations on 4 variables representing survey responses to dichotomous questions, with 1 denoting the "particularistic" values response and 2 denoting the "universalistic" values response.

Source

Stouffer, S.A. and J. Toby. 1951. "Role conflict and personality." *American Journal of Sociology*. 56: 395:406.

Goodman, Leo A. 1974. "Exploratory Latent-Structure Analysis Using Both Identifiable and Unidentifiable Models." *Biometrika*. 61(2): 215-231.

Goodman, Leo A. 2002. "Latent Class Analysis; The Empirical Study of Latent Types, Latent Variables, and Latent Structures." in Jacques A. Hagenaars and Allan L. McCutcheon, eds. *Applied Latent Class Analysis*. Cambridge: Cambridge University Press.

Examples

```
##
## Replication of latent class models in Goodman (2002),
## Tables 5b, 5c, and 6.
##
data(values)
f <- cbind(A,B,C,D)~1
M0 <- poLCA(f,values,nclass=1) # log-likelihood: -543.6498
M1 <- poLCA(f,values,nclass=2) # log-likelihood: -504.4677
M2 <- poLCA(f,values,nclass=3,maxiter=8000) # log-likelihood: -503.3011
```

Index