

Package ‘randomLCA’

April 19, 2009

Type Package

Title Random Effects Latent Class Analysis

Version 0.7-1

Date 2009-02-09

Author Ken Beath <ken@kjbeath.com.au>

Maintainer Ken Beath <ken@kjbeath.com.au>

Description Fits random effects latent class models, as well as standard latent class models.

Depends R(>= 2.6.0), lattice, nlme, boot

LazyLoad yes

LazyData yes

License GPL (>= 2)

Repository CRAN

Date/Publication 2009-04-14 17:33:06

R topics documented:

AIC	2
BIC	2
calc.cond.prob	3
calc.cond2.prob	4
calc.marg.prob	4
dentistry	5
fitted	6
genderrole	6
hivtests	8
logLik	9
myocardial	9

outcome.probs	10
plot	11
randomLCA	12
ranef	14
simulate	14
summary	15
symptoms	16
uterinecarcinoma	17

Index	19
--------------	-----------

AIC	<i>AIC for randomLCA object</i>
-----	---------------------------------

Description

Returns AIC for a randomLCA object.

Usage

```
## S3 method for class 'randomLCA':
AIC(object, ...)
```

Arguments

object	randomLCA object
...	additional argument; currently none is used.

Details

Not complete.

Author(s)

Ken Beath

 BIC

BIC for randomLCA object

Description

Returns BIC for a randomLCA object.

Usage

```
## S3 method for class 'randomLCA':
BIC(object, ...)
```

Arguments

<code>object</code>	randomLCA object
<code>...</code>	additional argument; currently none is used.

Details

Not complete.

Author(s)

Ken Beath

`calc.cond.prob`
Calculate Conditional Outcome Probabilities

Description

Calculates the conditional outcome probabilities for random effects models or for standard latent class returns the outcome probabilities. For random effects, the outcome probabilities may be calculated for various levels of the random effect.

Usage

```
calc.cond.prob(object, conditionalp = 0.5)
```

Arguments

<code>object</code>	RandomLCA object
<code>conditionalp</code>	the p value for the random effect

Value

Returns a data frame containing class, block, outcome, outcomep (outcome probability) and perc (p value for random effect) if conditionalp is specified.

Author(s)

Ken Beath <ken@kjbeath.com.au>

calc.cond2.prob *Calculate Conditional Outcome Probabilities for 2 Level Models*

Description

The conditional probabilities are obtained integrating over the period random effect.

Usage

```
calc.cond2.prob(object, conditionalp = 0.5)
```

Arguments

object RandomLCA object
conditionalp the p value for the subject random effect

Value

Returns a data frame containing perc (p value for random effect) , class, block, outcome and outcomep (outcome probability).

Author(s)

Ken Beath <ken@kjbeath.com.au>

calc.marg.prob *Calculates Marginal Outcome Probabilities*

Description

Calculates the marginal outcome probabilities for a random effects latent class model, by integrating the outcome probability over the random effect. This is performed using Gauss-Hermite quadrature with the number of quadrature points specified for the model fitting.

Usage

```
calc.marg.prob(object)
```

Arguments

object randomLCA object

Value

Returns a data frame containing class, block, outcome, outcomep (outcome probability).

Author(s)

Ken Beath

dentistry *Dental X-ray data*

Description

Six dentists evaluated dental x-rays for incipient caries in Handelman et al (1986), data consists of 5 of the dentists analysed by Espeland and Handelman (1989) using a latent class model. Further analysis incorporating a random effects latent class model by Qu et al (1996), and by Albert and Dodd (2004)

Usage

```
data(dentistry)
```

Format

A data frame with 32 observations on the following 6 variables.

v1 Dentist 1

v2 Dentist 2

v3 Dentist 3

v4 Dentist 4

v5 Dentist 5

freq Number of subjects

Source

Espeland and Handelman (1989)

References

Handelman, S.L., Leverett, D.H., Espeland, M.A. and Curzon, J.A. (1986) Clinical radiographic evaluation of sealed carious and sound tooth surfaces. *Journal of the American Dental Association*, **113**, 751–754.

Espeland, M.A. and Handelman, S.L. (1989) Using latent class models to characterize and assess relative error in discrete distributions. *Biometrics*, **45**, 587–599.

Qu, Y., Tan, M. and Kutner, M.H. (1996) Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics*, **52**, 797–810.

Albert P.S. and Dodd, L.E. (2004) A cautionary note on the robustness of Latent Class Models for estimating diagnostic error without a gold standard. *Biometrics*, **60**, 427–435.

Examples

```
data(dentistry)
# fit LCR model from Qu et al (1996)
dentistry.lca <- randomLCA(dentistry[,1:5], freq=dentistry$freq, blocksize=5)
dentistry.lcarandom <- randomLCA(dentistry[,1:5], freq=dentistry$freq,
  initmodel=dentistry.lca, random=TRUE, blocksize=1, probit=TRUE)
dentistry.lcarandomunequal <- randomLCA(dentistry[,1:5], freq=dentistry$freq,
  initmodel=dentistry.lcarandom, random=TRUE, blocksize=5, probit=TRUE)
```

fitted

fitted values

Description

Extract fitted values for randomLCA object

Usage

```
## S3 method for class 'randomLCA':
fitted(object, ...)
```

Arguments

object	randomLCA object
...	additional argument; currently none is used.

Value

Vector of fitted values corresponding to the observed frequencies.

Author(s)

Ken Beath <ken@kjbeath.com.au>

genderrole

*Gender Role Opinion Items***Description**

Opinions collected on gender roles in a study by Felling et al (1987). This was originally published in Heinen (1996) and subsequently in Galindo Garre and Vermunt (2006).

Usage

```
data(genderrole)
```

Format

A data frame with 16 observations on the following 5 variables.

Q1 Women's liberation sets women against men.

Q2 It's better for a wife not to have a job because that always poses problems in the household, especially if there are children.

Q3 The most natural situation occurs when the man is the breadwinner and the woman runs the household and takes care of the children.

Q4 It isn't really as important for a girl to get a good education as it is for a boy.

Q5 A woman is better suited to raise small children than a man.

Freq Number of subjects

Source

Galindo Garre and Vermunt (2006)

References

Felling, A., Peters, J., and Schreuder, O. (1987) Religion in Dutch society 85: Documentation of a national survey on religious and secular attitudes in 1985. Amsterdam: Steinmetz Archive.

Galindo Garre, F. and Vermunt, J.K. (2006) Avoiding boundary estimates in latent class analysis by Bayesian posterior mode estimation. *Behaviormetrika*, **33**, 43–59.

Heinen, T. (1996) Latent Class and Discrete Latent Trait Models: Similarities and Differences.

Examples

```
data(genderrole)
# standard latent class
genderrole.lca1 <- randomLCA(genderrole[,1:5], freq=genderrole$Freq, nclass=1)
genderrole.lca2 <- randomLCA(genderrole[,1:5], freq=genderrole$Freq)
genderrole.lca3 <- randomLCA(genderrole[,1:5], freq=genderrole$Freq, nclass=3)
# repeat with random effect
genderrole.lca1random <- randomLCA(genderrole[,1:5], freq=genderrole$Freq, nclass=1, random=TRUE)
```

```

genderrole.lca2random <- randomLCA(genderrole[,1:5],freq=genderrole$Freq,random=TRUE)
genderrole.lca3random <- randomLCA(genderrole[,1:5],freq=genderrole$Freq,nclass=3,random=TRUE)
# improved BIC for 1 class random
print(c(BIC(genderrole.lca1),BIC(genderrole.lca2),BIC(genderrole.lca3)))
print(c(BIC(genderrole.lca1random),BIC(genderrole.lca2random),BIC(genderrole.lca3random)))
# can also repeat fits with blocksize=5 to give mixture of IRT models

```

hivtests

HIV testing data

Description

Serum samples are tested for HIV by 4 different biossays in Alvord et al (1988) and sensitivity and specificity determined using latent class analysis. Qu et al (1996) repeat the analysis using a model incorporating a random effect.

Usage

```
data(hivtests)
```

Format

A data frame with 16 observations on the following 5 variables.

v1 Test 1

v2 Test 2

v3 Test 3

v4 Test 4

freq Number of subjects

Source

Qu, Tan and Kutner (1989)

References

Alvord, W.G., Drummond, J.E., Arthur, L.O., Goedert, J.J., Levine, P.H., Murphy, E.L., Weiss, S.H., and Blattner, W.A. (1988) A method for predicting individual HIV infection status in teh absence of clinical information. *AIDS Research and Human Retroviruses*, **4**, 295–304.

Qu, Y., Tan, M. and Kutner, M.H. (1996) Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics*, **52**, 797–810.

Examples

```

data(hivtests)
# fit 2LC model from Qu et al (1996)
hivtests.lca <- randomLCA(hivtests[,1:4],freq=hivtests$freq)
# fit model with random effect
# this is different to teh Qu paper which only applies the RE to tests 2 and 3
hivtests.lcarandom <- randomLCA(hivtests[,1:4],freq=hivtests$freq,
                               initmodel=hivtests.lca,random=TRUE,blocksize=1,probit=TRUE)

```

logLik

log Likelihood for randomLCA object

Description

Returns log Likelihood for a randomLCA object.

Usage

```

## S3 method for class 'randomLCA':
logLik(object, ...)

```

Arguments

object	randomLCA object
...	additional argument; currently none is used.

Details

Not complete.

Author(s)

Ken Beath

myocardial

Myocardial Infarction (MI)

Description

Four tests were performed on hospital patients to determine if a myocardial infarction had occurred.

Usage

```

data(myocardial)

```

Format

A data frame with 32 observations on the following 6 variables.

Q.wave result from ECG test

History clinical history

LDH flipped, enzyme related to tissue breakdown

CPK high, creatine kinase or creatine phosphokinase, related to muscle damage

freq Number of subjects

Source

Rindskopf and Rindskopf (1986)

References

Galen, R.S. and Gambino, S.R. (1975) Beyond Normality: The Predictive Value and Efficiency of Medical Diagnosis. Wiley:New York.

Rindskopf, D. and Rindskopf, W. (1986) The Value of Latent Class Analysis in Medical Diagnosis. *Statistics in Medicine*, **5**, 21–27.

Examples

```
data(myocardial)
# fit 2 class model from Rindskopf and Rindskopf (1986)
myocardial.lca2 <- randomLCA(myocardial[,1:4], freq=myocardial$freq)
```

```
outcome.probs      Extract outcome probabilities for randomLCA object
```

Description

Extract outcome probabilities and confidence intervals for a randomLCA object.

Usage

```
## S3 method for class 'randomLCA':
outcome.probs(object, level = 0.95, boot=FALSE, type="perc", R=ifelse(type=="perc",
```

Arguments

object	randomLCA object
level	confidence interval
boot	use parametric bootstrap to obtain confidence interval
type	type of bootstrap confidence intervals to use "perc", "normal" and "basic" are valid, see boot.ci for description. There is a risk of excessively large confidence limits when using other than "perc"
R	replications for parametric bootstrap
...	additional argument; currently none is used.

Details

Confidence intervals are calculated based on asymptotic normality of the estimates transformed by either the inverse of the probit or logistic, or using parametric bootstrap. The asymptotic confidence intervals are currently only available for models without random effects.

Author(s)

Ken Beath

Examples

```
# dentist data
data(dentistry)
# standard latent class with 2 classes
dentistry.lca2 <- randomLCA(dentistry[,1:5],freq=dentistry$freq,nclass=2)
print(outcome.probs(dentistry.lca2))
```

plot

Plot a randomLCA object

Description

Plots the outcome probabilities for a randomLCA object, for random effects objects this can be either marginal or conditional or both. For a 2 level random effects model conditional2 will condition on the subject random effect and integrate over the period random effects.

Usage

```
## S3 method for class 'randomLCA':
plot(x, ... ,graphtype = c("conditional", "marginal",
    "conditional2","both"), conditionalp = 0.5, classshorizontal = TRUE)
```

Arguments

x	randomLCA object
graphtype	Type of graph
conditionalp	For a conditional graph the percentile corresponding to the random effect at which the outcome probability is to be calculated
classshorizontal	classes to be plotted across the page
...	additional parameters to xyplot

Author(s)

Ken Beath <ken@kjbeath.com.au>

See Also

[calc.cond.prob,calc.cond.prob](#)

Examples

```
data(uterinecarcinoma)
# standard latent class with 2 classes
uterinecarcinoma.lca2 <- randomLCA(uterinecarcinoma[,1:7],freq=uterinecarcinoma$freq)
plot(uterinecarcinoma.lca2)
```

randomLCA

Fits a Latent Class Model including a Random Effect

Description

Fit latent class models, which may include a random effect.

Usage

```
randomLCA(patterns, freq, nclass=2, calcSE=TRUE, initmodel=NULL, blocksize=1, notrials=20,
           random=FALSE, byclass=FALSE, quadpoints=21, level2=FALSE, probit=FALSE,
           verbose=FALSE, seed = as.integer(runif(1, 0, .Machine$integer.max)))
```

Arguments

patterns	Matrix of 0 and 1 defining the outcome patterns
freq	Frequency for each outcome pattern, if missing this is calculated from the patterns
nclass	Number of classes to be fitted
calcSE	Calculate standard errors for parameters
initmodel	Initial model of class randomLCA
blocksize	Size of blocks of outcomes
notrials	For a standard latent class model, the number of random starting values used
random	Random effect
byclass	Random effect by class
quadpoints	Number of quadrature points for adaptive quadrature
level2	Fit 2 level random effects model (further details to follow)
probit	Probit model for random effect
seed	Initial random seed for generating starting values
verbose	Prints verbose of fit progress

Details

The structure of the patterns is assumed to be a number of blocks of different outcomes each of blocksize, allowing outcomes to be repeated. Each outcome is assumed to have it's own loading. An example is the width of the patterns is n and the blocksize is n , resulting in n outcomes and therefore n loadings. Alternatively if the blocksize is 1, then there are n repeats of the same outcome (but with different probabilities) with the same loading. In practice they may not be the same type of outcome, but usually will be.

The algorithm used is EM for the standard latent class and adaptive (in the sense of moving the location of the quadrature points) Gauss-Hermite quadrature for the random effects models. The number of quadrature points defaults to 21.

There are some improvements that will be made to improve the speed. The first of these will be to allow Bayesian estimation or penalised likelihood using a Dirichlet prior for the outcome probabilities. This should allow a considerable improvement in convergence. The second will be to recode the major part of the likelihood calculations in C. Switching to a GEM algorithm to improve convergence may also be a possibility.

Value

randomLCA object This contains

outcomep	Outcome probability
classp	Class probabilities
lambdacoef	Loadings

Author(s)

Ken Beath

Examples

```
# dentist data
data(dentistry)
# standard latent class with 2 classes
dentistry.lca2 <- randomLCA(dentistry[,1:5], freq=dentistry$freq, nclass=2)
# random effects model with constant random effect loading
dentistry.lca2random <- randomLCA(dentistry[,1:5], freq=dentistry$freq,
  initmodel=dentistry.lca2, nclass=2, random=TRUE, probit=TRUE)
# allow loading to vary by dentist
# this is the 2LCR model from Qu et al (1996)
dentistry.lca2random1 <- randomLCA(dentistry[,1:5], freq=dentistry$freq,
  initmodel=dentistry.lca2random, nclass=2, random=TRUE, probit=TRUE, blocksize=5)
```

ranef	<i>Extract random effects from a randomLCA object</i>
-------	---

Description

Extracts the Empirical Bayes estimates of the random effects.

Usage

```
## S3 method for class 'randomLCA':
ranef(object, ...)
```

Arguments

object	randomLCA object with a random effect
...	additional argument; currently none is used.

Value

A matrix with the first column containing the random effects and the second column the standard error of the random effects.

Author(s)

Ken Beath

simulate	<i>Simulate</i>
----------	-----------------

Description

Simulate data from a fitted randomLCA model

Usage

```
## S3 method for class 'randomLCA':
simulate(object, nsim, seed, ...)
```

Arguments

object	randomLCA object
nsim	number of data sets to be simulated
seed	random seed
...	additional optional arguments.

Author(s)

Ken Beath

summary

Summary for randomLCA object

Description

Summarises the fit of a randomLCA object.

Usage

```
## S3 method for class 'randomLCA':
summary(object, ...)
```

Arguments

object	randomLCA object
...	additional argument; currently none is used.

Details

Not complete.

Value

logLik	Log Likelihood
AIC	AIC
BIC	BIC
nclass	no of classes
probit	link is probit
classp	class probabilities
outcomep	outcome probabilities (conditional)
margoutcomep	outcome probabilities (marginal), if model contains random effects
random	model includes random effects
level2	model has 2 level hierarchy
byclass	lambda and tau vary by class
lambdacoef	lambda coefficients
taucoef	tau coefficients

Author(s)

Ken Beath

 symptoms

Symptoms data

Description

Allergy and respiratory symptoms for infants 0 to 2 years in six month periods. Outcome is presence or absence of symptom in the six months. Original data was collected at Visits 1-7 over the 2 year period which were summarised to six month periods.

This is data relating to a paper which will hopefully be published shortly. I will be updating this entry then.

Note that these models can be slow to fit, with the "symptoms.lca2random2" model taking about 3 hours. I have a couple more possibilities to speed things up. These are to code the algorithms completely in C (they are currently a mix of C and R), use a Bayesian methodology to prevent the estimates moving towards the boundaries (this can also be described as penalized likelihood) and to look at other optimisation functions.

Thanks to the investigators of the CAPS study for making the data available.

Usage

```
data(symptoms)
```

Format

A data frame with 444 observations on the following 17 variables.

Nightcough.13 Night cough in visits 1-3
Wheeze.13 Wheeze in visits 1-3
Itchyrash.13 Itchy rash in visits 1-3
FlexDerma.13 Flexural Dermatitis in visits 1-3
Nightcough.45 Night cough in visits 1-3
Wheeze.45 Wheeze in visits 4-5
Itchyrash.45 Itchy rash in visits 4-5
FlexDerma.45 Flexural Dermatitis in visits 4-5
Nightcough.6 Night cough in visit 6
Wheeze.6 Wheeze in visit 6
Itchyrash.6 Itchy rash in visit 6
FlexDerma.6 Flexural Dermatitis in visits 1-3
Nightcough.7 Night cough in visit 7
Wheeze.7 Wheeze in visit 7
Itchyrash.7 Itchy rash in visit 7
FlexDerma.7 Flexural Dermatitis in visit 7
Freq Number of subjects

Source

Mihrshai et al (2001)

References

Mihrshahi, S., Peat, J.K., Webb, K., Tovey, R.E., Marks, G.B., Mellis, C.M. and Leeder S.R. (2001) The Childhood Asthma Prevention Study (CAPS): Design and research protocol of a randomized trial for the primary prevention of asthma. *Control led Clinical Trials*, **22**:333–354.

Examples

```
data(symptoms)
## Not run:
symptoms.lca2 <- randomLCA(symptoms[,1:16],freq=symptoms$Freq,nclass=2)
symptoms.lca2random <- randomLCA(symptoms[,1:16],freq=symptoms$Freq,initmodel=symptoms.lca2,
symptoms.lca2random2 <- randomLCA(symptoms[,1:16],freq=symptoms$Freq,initmodel=symptoms.lca2)
## End(Not run)
```

uterinecarcinoma *Uterine Carcinoma Data*

Description

Classification of 118 histology samples by 118 pathologists. Original classification in Holmquist et al (1967) was to one of five categories, this has been reduced to two. Analysed by a number of authors, with a random effects in Qu et al (1996).

Usage

```
data(uterinecarcinoma)
```

Format

A data frame with 20 observations on the following 8 variables.

v1 Pathologist 1

v2 Pathologist 2

v3 Pathologist 3

v4 Pathologist 4

v5 Pathologist 5

v6 Pathologist 6

v7 Pathologist 7

freq Number of observed pattern

Source

Qu et al (1996)

References

Holmquist, N.D., McMahan, C.A., and Williams, O.D. (1967) Variability in classification of carcinoma in situ of the uterine cervix. *Archives of Pathology*, **84**, 344–345. Qu, Y., Tan, M. and Kutner, M.H. (1996) Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics*, **52**, 797–810.

Examples

```
data(uterinecarcinoma)
uterinecarcinoma.lca2 <- randomLCA(uterinecarcinoma[,1:7],freq=uterinecarcinoma$freq)

uterinecarcinoma.lcarandom2 <- randomLCA(uterinecarcinoma[,1:7],freq=uterinecarcinoma$freq,
    initmodel=uterinecarcinoma.lca2,random=TRUE,probit=TRUE,quadpoints=41)
# LCR1 model of Que et al. This is fairly unstable and requires starting values from the sim
# it is also slow and doesn't improve the model fit
## Not run:
uterinecarcinoma.lcarandom2by <- randomLCA(uterinecarcinoma[,1:7],freq=uterinecarcinoma$freq,
    initmodel=uterinecarcinoma.lcarandom2,byclass=TRUE,random=TRUE,probit=TRUE,quadpoint
## End(Not run)
```

Index

*Topic **datasets**

- dentistry, 4
- genderrole, 6
- hivtests, 7
- myocardial, 9
- symptoms, 15
- uterinecarcinoma, 17

*Topic **methods**

- AIC, 2
- BIC, 2
- calc.cond.prob, 3
- calc.cond2.prob, 3
- calc.marg.prob, 4
- fitted, 5
- logLik, 8
- outcome.probs, 10
- plot, 11
- ranef, 13
- simulate, 14
- summary, 14

*Topic **multivariate**

- randomLCA, 12

AIC, 2

BIC, 2

calc.cond.prob, 3, 11

calc.cond2.prob, 3

calc.marg.prob, 4

dentistry, 4

fitted, 5

genderrole, 6

hivtests, 7

logLik, 8

myocardial, 9

outcome.probs, 10

plot, 11

randomLCA, 12

ranef, 13

simulate, 14

summary, 14

symptoms, 15

uterinecarcinoma, 17