Package: elrm (via r-universe)

October 20, 2026

Version 1.2.6

Date 2024-12-17

Type Package

Title Exact Logistic Regression via MCMC

Author David Zamar [aut, cre], Jinko Graham [aut], Brad McNeney [aut]

Maintainer David Zamar <zamar.david@gmail.com>

Depends R(>= 2.7.2), coda, graphics, stats

Description Implements a Markov Chain Monte Carlo algorithm to approximate exact conditional inference for logistic regression models. Exact conditional inference is based on the distribution of the sufficient statistics for the parameters of interest given the sufficient statistics for the remaining nuisance parameters. Using model formula notation, users specify a logistic model and model terms of interest for exact inference. See Zamar et al. (2007) <doi:10.18637/jss.v021.i03> for more details.

License GPL (>= 2)

NeedsCompilation yes

Repository https://kookoot15.r-universe.dev

Date/Publication 2024-12-18 08:00:07 UTC

RemoteUrl https://github.com/cran/elrm

RemoteRef HEAD

RemoteSha eedd5cbef0242361551661349be9b97db93985cb

Contents

crashDat 2
diabDat 2
drugDat 3
elrm 3
plot.elrm 6

summary.elrm 7
titanDat 8
update.elrm 9
utiDat 11

Index **12**

`crashDat` *Crash Dataset: Calibration of Crash Dummies in Automobile Safety Tests*

Description

Data from 58 simulated car crashes were analyzed. The relationship between the crash outcome (fatal, non-fatal) and 3 covariates was modeled.

Usage

```
data(drugDat)
```

Format

Matrix with columns:

[,1]	age	numeric	designed age of the crash dummy
[,2]	vel	numeric	velocity of the car at impact
[,3]	acl	numeric	acceleration of the car at impact
[,4]	y	numeric	outcome from test (1=fatal, 0=non-fatal)
[,5]	n	numeric	number of binomial trials. For this data, n is a vector of ones, thus each trial is modeled as a Bernoulli

---

`diabDat` *Simulated Diabetes Dataset*

Description

This dataset was simulated by sampling cases from an existing type 1 diabetes study (the original data could not be disclosed). The study investigates the relationship between concentration levels (low and high) of the islet antigen 2 antibody (IA2A) and several covariates of potential interest in type 1 diabetes patients (age, gender, and the number of copies (0,1 or 2) of the HLA-DQ2, HLA-DQ8 and HLA-DQ6.2 haplotypes).

Usage

```
data(diabDat)
```

Format

Matrix with columns:

[,1]	n	numeric	number of binomial trials.
[,2]	IA2A	numeric	number of patients with high concentration levels of the islet antigen 2 antibody
[,3]	gender	numeric	gender of patient
[,4]	age	numeric	age of patient
[,5]	nDQ2	numeric	number of copies of the HLA-DQ2 haplotype (0,1, or 2)
[,6]	nDQ8	numeric	number of copies of the HLA-DQ8 haplotype (0,1, or 2)
[,7]	nDQ6.2	numeric	number of copies of the HLA-DQ6.2 haplotype (0,1, or 2)

---

drugDat *Drug Dataset*

Description

Simulated data for a hypothetical drug experiment comparing control versus treatment.

Usage

```
data(drugDat)
```

Format

Matrix with columns:

[,1]	sex	numeric	gender (1=male, 0=female)
[,2]	treatment	numeric	treatment type (1=treatment, 0=control)
[,3]	recovered	numeric	number of subjects that recovered
[,4]	n	numeric	number of binomial trials

---

elrm *elrm: exact-like inference in logistic regression models*

Description

`elrm` implements a modification of the Markov Chain Monte Carlo algorithm proposed by Forster et al. (2003) to approximate exact conditional inference for logistic regression models. The modifications can handle larger datasets than the original algorithm (Zamar 2006). Exact conditional inference is based on the distribution of the sufficient statistics for the parameters of interest given the sufficient statistics for the remaining nuisance parameters. Using model formula notation, users specify a logistic model and model terms of interest for exact inference.

Usage

```
elrm(formula, interest, r = 4, iter = 1000, dataset, burnIn = 0, alpha = 0.05)
```

Arguments

formula	a formula object that contains a symbolic description of the logistic regression model of interest in the usual R formula format. One exception is that the binomial response should be specified as *success/trials*, where *success* gives the number of successes and *trials* gives the number of binomial trials for each row of *dataset*.
interest	a formula object that contains a symbolic description of the model terms for which exact conditional inference is of interest.
r	a parameter of the MCMC algorithm that influences how the Markov chain moves around the state space. Small values of r cause the chain to take small, relatively frequent steps through the state space; larger values cause larger, less frequent steps. The value of r must be an even integer less than or equal to the length of the response vector. Typical values are 4, 6 or 8; default=4.
iter	an integer representing the number of Markov chain iterations to make (must be larger than or equal to 1000); default=1000.
dataset	a data.frame object where the data are stored.
burnIn	the burn-in period to use when conducting inference. Values of the Markov chain in the burn-in period are discarded; default=0.
alpha	determines the level used for confidence intervals; default=0.05.

Details

The function `summary()` (i.e., `summary.elrm`) can be used to obtain or print a summary of the results.

Each estimated exact p-value is based on the conditional probabilities test.

The Monte Carlo standard error of each p-value is computed by the batch-means method (Geyer C.J. 1992).

Inference on each parameter must be based on a Markov chain of at least 1000 iterations, otherwise `NA` is returned.

If the observed value of the sufficient statistic for a parameter is either the maximum or the minimum value sampled, the MUE of the parameter is given instead of the CMLE. In such cases, the resulting confidence interval is open-ended on one side.

Apart from the documentation files accompanying this package, the **elrm** package vignette may be downloaded from https://www.jstatsoft.org/article/view/v021i03. The vignette is also distributed with the code.

Value

`coeffs`	a vector containing the parameter estimates.
`coeffs.ci`	a list containing (1-*alpha*)*100% confidence intervals for each parameter of interest.
`p.values`	a vector containing the estimated p-value for jointly testing that the parameters of interest are simultaneously equal to zero, and the full conditional p-values from separately testing each parameter equal to zero.
`p.values.se`	a vector containing the Monte Carlo standard errors of the estimated p-values of each term of interest.
`mc`	an `mcmc` object containing the Markov chain of sampled values of the sufficient statistics for the parameters of interest. Columns correspond to parameters; rows to Monte Carlo iterations.
`mc.size`	a vector containing the lengths of the extracted Markov chains used in testing each parameter. The length of the Markov chain used for the joint test (i.e., *iter*) is also included as the first element.
`obs.suff.stat`	a vector containing the observed value of the sufficient statistic for each parameter of interest.
`distribution`	a list containing distribution tables for the sampled values of the sufficient statistic of the parameters of interest conditional on all the rest.
`call.history`	a list composed of the matched call and the history of calls to `update()`.
`dataset`	the data.frame object that was passed to `elrm()` as an argument.
`last`	the last response vector sampled by the Markov chain.
`r`	the value of *r* passed to `elrm()` as an argument.
`ci.level`	the level used when constructing the confidence intervals for the parameters of interest. The level is calculated as (1-*alpha*)*100%.

Warning

The labels of the terms in the in the *interest* model should match those found in the *formula* model. Thus, the *term.labels* attribute of `terms.formula(`*interest*`)` should match those found in `terms.formula(`*formula*`)`. Please see the Examples section for more details.

Author(s)

David Zamar, Jinko Graham, Brad McNeney

References

Forster J.J., McDonald J.W. and Smith P.W.F. Markov Chain Monte Carlo Exact Inference for Binomial and Multinomial Logistic Regression Models. Statistics and Computing, 13:169-177, 2003.

Geyer C.J. Practical Markov chain Monte Carlo. Statistical Science, 7:473-511, 1992.

Zamar David. Monte Carlo Markov Chain Exact Inference for Binomial Regression Models. Master's thesis, Statistics and Actuarial Sciences, Simon Fraser University, 2006.

Zamar D, McNeney B and Graham J. elrm: Software Implementing Exact-like Inference for Logistic Regression Models. Journal of Statistical Software 2007, 21(3).

See Also

update.elrm, summary.elrm, plot.elrm.

Examples

```
# Drug dataset example with sex and treatment as the variables of interest
data(drugDat);
drug.elrm = elrm(formula=recovered/n~sex+treatment, interest=~sex+treatment,
 r=4,iter=40000, burnIn=1000, dataset=drugDat);


# crash dataset example where the terms of interest are age and
# the interaction of age and velocity.
data(crashDat);
crash.elrm = elrm(formula=y/n~vel+age+vel:age, interest=~vel:age, r=4, iter=10000,
 dataset=crashDat, burnIn=100);

# Urinary tract dataset example with dia as the variable of interest
data(utiDat);
uti.elrm = elrm(uti/n~age+current+dia+oc+pastyr+vi+vic+vicl+vis, interest=~dia,r=4,
 iter=20000,burnIn=1000, dataset=utiDat);
```

---

plot.elrm *Plot Diagnostics for an elrm Object*

Description

Produces both a trace plot and histogram of the sampled values of each sufficient statistic of interest. Sampled values within the burn-in period are also plotted.

Usage

```
## S3 method for class 'elrm'
plot(x, p = 1, breaks = "Sturges", ask=FALSE, ...)
```

Arguments

x	an object of class elrm, resulting from a call to elrm() or a previous call to update().
p	the sampling fraction of points to be plotted. A random sample consisting of p*100% of all the observations in the Markov chain is plotted; default=1.
breaks	a vector giving the number of cells to use for the histogram of each sufficient statistic of interest or a single number giving the number of cells for each histogram or the character string naming an algorithm to compute the number of cells.
ask	the graphics parameter *ask*: see par for details. If set to TRUE will ask between plots corresponding to each sufficient statistic; default=FALSE.
...	additional arguments to the plot function (currently unused).

Details

The default for *breaks* is "Sturges": see `nclass.Sturges`. Other names for which algorithms are supplied are "Scott" and "FD".

Author(s)

David Zamar, Jinko Graham, Brad McNeney

References

Zamar David. Monte Carlo Markov Chain Exact Inference for Binomial Regression Models. Master's thesis, Statistics and Actuarial Sciences, Simon Fraser University, 2006.

Zamar D, McNeney B and Graham J. elrm: Software Implementing Exact-like Inference for Logistic Regression Models. Journal of Statistical Software 2007, 21(3).

See Also

`update.elrm`, `summary.elrm`, `elrm`.

Examples

```
# Drug dataset example with sex and treatment as the variables of interest
data(drugDat);
drug.elrm = elrm(formula=recovered/n~sex+treatment, interest=~sex+treatment, r=4,
 iter=40000, burnIn=1000, dataset=drugDat);

# Plot the sampled values of the sufficient statistic for each parameter of
# interest (sex and treatment)
plot(drug.elrm,p=0.10,ask=TRUE);
```

`summary.elrm` *Summarize an elrm Object*

Description

Summary method for class `elrm` that formats and prints out the results of an `elrm` object.

Usage

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

Arguments

`object`	an object of class `elrm`, resulting from a call to `elrm()` or a previous call to `update()`.
`...`	additional arguments to the summary function (currently unused).

Details

The following results are formatted and printed to the screen: the matched call, coefficient estimates and confidence intervals for each model term of interest, estimated p-value for jointly testing that the parameters of interest are simultaneously equal to zero, full conditional p-values from separately testing each parameter equal to zero, length of the Markov chain that inference was based on, and the Monte Carlo standard error of each reported p-value.

Author(s)

David Zamar, Jinko Graham, Brad McNeney

References

Zamar David. Monte Carlo Markov Chain Exact Inference for Binomial Regression Models. Master's thesis, Statistics and Actuarial Sciences, Simon Fraser University, 2006.

Zamar D, McNeney B and Graham J. elrm: Software Implementing Exact-like Inference for Logistic Regression Models. Journal of Statistical Software 2007, 21(3).

See Also

update.elrm, plot.elrm, elrm.

Examples

```
# Drug dataset example with both sex and treatment as the variables of interest
data(drugDat);
drug.elrm = elrm(formula=recovered/n~sex+treatment, interest=~sex+treatment, r=4,
 iter=50000, burnIn=1000, dataset=drugDat);

# Summarize the results:
summary(drug.elrm);


# Urinary tract dataset example with dia as the variable of interst
data(utiDat);
uti.elrm = elrm(uti/n~age+current+dia+oc+pastyr+vi+vic+vicl+vis, interest=~dia, r=4,
 iter=30000, burnIn=100, dataset=utiDat);

# Summarize the results:
summary(uti.elrm);
```

---

titanDat *Titanic Dataset*

Description

Relationship between survival and passenger class on the Titanic. The records of the sinking of the Titanic were studied to establish the relationship between survival and passenger class on the ship. For each person on board the ocean liner, this dataset records sex, age (child/adult), class (crew, 1st, 2nd, 3rd class) and whether or not the person survived.

Usage

```
data(titanDat)
```

Format

Matrix with columns:

[,1]	surv	numeric	number of survivors
[,2]	n	numeric	total number of people
[,3]	class	numeric	passenger class (0 = crew, 1 = first, 2 = second, 3 = third)
[,4]	age	numeric	age group (1 = adult, 0 = child))
[,5]	sex	numeric	gender (1 = male, 0 = female)

References

Mehta CR, Patel NR. Logxact for Windows. Cytel Software Corporation, Cambridge, USA, 1999.

---

update.elrm *Update Method for Objects of Class elrm.*

Description

An update method for objects created by `elrm()`. Extends the Markov chain of an `elrm` object by a specified number of iterations.

Usage

```
## S3 method for class 'elrm'
update(object, iter, burnIn = 0, alpha = 0.05, ...)
```

Arguments

`object`	an object of class `elrm`, resulting from a call to `elrm()` or a previous call to `update()`.
`iter`	an integer representing the number of Markov chain iterations to make.
`burnIn`	the burn-in period to use when conducting inference. Values of the Markov chain in the burn-in period are discarded; default=0.
`alpha`	determines the level used for confidence intervals; default=0.05.
`...`	additional arguments to the update function (currently unused).

Details

Extends the Markov chain of an `elrm` object by creating a new Markov chain of the specified length using the last sampled value as the starting point. The newly created chain is then appended to the original. Subsequent inference is based on the extended Markov chain.

Value

An object of class `elrm`.

Author(s)

David Zamar, Jinko Graham, Brad McNeney

References

Zamar David. Monte Carlo Markov Chain Exact Inference for Binomial Regression Models. Master's thesis, Statistics and Actuarial Sciences, Simon Fraser University, 2006.

Zamar D, McNeney B and Graham J. elrm: Software Implementing Exact-like Inference for Logistic Regression Models. Journal of Statistical Software 2007, 21(3).

See Also

`summary.elrm`, `plot.elrm`, `elrm`.

Examples

```
# Drug dataset example with sex as the variables of interest
data(drugDat);
drug.elrm = elrm(formula=recovered/n~sex, interest=~sex, r=2,
 iter=1000, burnIn=0, dataset=drugDat);

# Summarize the results
summary(drug.elrm);

# Call update and extend the chain by 500 iterations and set the burn-in
# period to 100 iterations
drug.elrm = update(drug.elrm, iter=500, burnIn=100);

# Summarize the results
summary(drug.elrm);

# Now change the burn-in to 200
drug.elrm = update(drug.elrm, iter=0, burnIn=200);

# Summarize the results
summary(drug.elrm);
```

utiDat *Urinary Tract Infection and Contraceptive Use*

Description

How is the development of first-time urinary tract infection (UTI) related to contraceptive use? A study of sexually active college women with UTI was conducted, and their use of various contraceptives was surveyed.

Usage

```
data(utiDat)
```

Format

Matrix with columns:

[,1]	uti	numeric	infection status (1=yes, 0=no)
[,2]	n	numeric	number of binomial trials
[,3]	age	numeric	age of the person
[,4]	current	numeric	no regular current sex partner (1=yes,0=no)
[,5]	dia	numeric	diaphragm use (1=yes, 0=no)
[,6]	oc	numeric	oral contraceptive (1=yes, 0=no)
[,7]	pastyr	numeric	no regular partner with relationship < 1yr (1=yes, 0=no)
[,8]	vi	numeric	vaginal intercourse (1=yes, 0=no)
[,9]	vic	numeric	vaginal intercourse with condom (1=yes, 0=no)
[,10]	vicl	numeric	vaginal intercourse with lubricated condom (1=yes, 0=no)
[,11]	vis	numeric	vaginal intercourse with spermicide (1=yes, 0=no)

References

Mehta CR, Patel NR. Logxact for Windows. Cytel Software Corporation, Cambridge, USA, 1999.

Index

∗ **datasets**
crashDat, 2
diabDat, 2
drugDat, 3
titanDat, 8
utiDat, 11
∗ **methods**
elrm, 3
plot.elrm, 6
summary.elrm, 7
update.elrm, 9

crashDat, 2

diabDat, 2
drugDat, 3

elrm, 3, *7, 8*, *10*

mcmc, *5*

nclass.Sturges, *7*

par, *6*
plot.elrm, 6, *6*, *8*, *10*

summary.elrm, *4*, *6*, 7, *7*, *10*

titanDat, 8

update.elrm, *6–8*, 9
utiDat, 11