Package 'TwoStepCLogit'

May 7, 2026

Version 1.2.6

Date 2025-03-26

Title Conditional Logistic Regression: A Two-Step Estimation Method

Description Conditional logistic regression with longitudinal follow up and individual-level random coefficients: A stable and efficient two-step estimation method.

License GPL-2

Imports survival, stats

Suggests testthat

LazyData true

RoxygenNote 7.3.2

NeedsCompilation no

Author Radu V. Craiu [aut],
Thierry Duchesne [aut, cre],
Daniel Fortin [aut],
Sophie Baillargeon [aut]

Maintainer Thierry Duchesne <thierry.duchesne@mat.ulaval.ca>

Repository CRAN

Date/Publication 2025-03-27 12:00:06 UTC

Contents

TwoStepCLogit-package *Conditional Logistic Regression: A Two-Step Estimation Method*

Description

Conditional logistic regression with longitudinal follow up and individual-level random coefficients: A stable and efficient two-step estimation method (see `Ts.estim`).

Details

Package:	TwoStepCLogit
Type:	Package
Version:	1.2.6
Date:	2025-03-26
License:	GPL-2

Author(s)

Radu V. Craiu, Thierry Duchesne, Daniel Fortin and Sophie Baillargeon

Maintainer: Thierry Duchesne <thierry.duchesne@mat.ulaval.ca>

References

Craiu, R.V., Duchesne, T., Fortin, D. and Baillargeon, S. (2011), Conditional Logistic Regression with Longitudinal Follow-up and Individual-Level Random Coefficients: A Stable and Efficient Two-Step Estimation Method, *Journal of Computational and Graphical Statistics*. **20**(3), 767-784.

bison *Bison Dataset*

Description

Bison data collected in Prince Albert National Park, Saskatchewan, Canada (Craiu et al. 2011).

Format

A data frame with 16818 observations on the following 10 variables.

Cluster pair of animals (dyad) ID

Strata stratum ID

Y response variable: 1 for visited locations, 0 otherwise

water land cover indicator covariate: 1 for water, 0 otherwise

agric land cover indicator covariate: 1 for agricultural locations , 0 otherwise

forest land cover indicator covariate: 1 for forests, 0 otherwise

meadow land cover indicator covariate: 1 for meadows, 0 otherwise

biomass continuous covariate: above-ground vegetation biomass index measured (in kg/m^2) only at locations within meadows, 0 otherwise

pmeadow continuous covariate: the proportion of meadow in a circular plot (700 m in radius) centered at the bison's location

Details

This data set was collected in order to study habitat selection by groups of free-ranging bison. For each observed group, two individuals (dyad) equipped with GPS radio-collars were followed simultaneously. A cluster is defined here as a pair of bison. This data set contains 20 clusters. The number of strata per cluster varies between 13 and 345 for a total of 1410 strata. A stratum is composed of two visited GPS locations (one for each individual) gathered at the same time, together with 10 random locations (five drawn within 700 m of each of the two focal bison). Therefore, there are 12 observations per stratum, with 2 cases (Y=1) and 10 controls (Y=0). However, due to problems in the data collection, 17 of the 1410 strata have only 6 observations (1 case and 5 controls).

References

Craiu, R.V., Duchesne, T., Fortin, D. and Baillargeon, S. (2011), Conditional Logistic Regression with Longitudinal Follow-up and Individual-Level Random Coefficients: A Stable and Efficient Two-Step Estimation Method, *Journal of Computational and Graphical Statistics*. **20**(3), 767-784.

Examples

```
# Some descriptive statistics about the data set:
ddim(formula = Y ~ strata(Strata) + cluster(Cluster), data = bison)

# Model 1: covariates meadow, biomass and biomass^2
# Random effects in front of biomass and biomass^2
# Main diagonal covariance structure for D
Fit1 <- Ts.estim(formula = Y ~ meadow + biomass + I(biomass^2) +
        strata(Strata) + cluster(Cluster), data = bison,
        random = ~ biomass + I(biomass^2), all.m.1=FALSE, D="UN(1)")

Fit1
```

```
# Model 2: only covariates biomass and biomass^2
# Random effects in front of biomass and biomass^2
# Main diagonal covariance structure for D
Fit2 <- Ts.estim(formula = Y ~ biomass + I(biomass^2) + strata(Strata) +
        cluster(Cluster), data = bison, all.m.1=FALSE, D="UN(1)")
Fit2
# Results reported in Table 2 of Craiu et al. (2011).
```

---

ddim *Data Dimension Statistics*

Description

Function that computes dimension statistics for a data set with clusters and strata and its `print` method.

Usage

```
ddim(formula, data)

## S3 method for class 'ddim'
print(x, ...)
```

Arguments

`formula`	A formula object, with the response on the left of a `~` operator and, on the right hand side, a `strata` and a `cluster` term (ex. `formula = Y ~ strata(var_strata) + cluster(var_cluster)`). The `strata` and `cluster` functions (from the package survival) are used to identify the stratification and the cluster variables, respectively.
`data`	A data frame (or object coercible by as.data.frame to a data frame) containing the variables in the model.
`x`	An object, produced by the `ddim` function, to print.
`...`	Further arguments to be passed to `print.default`.

Value

`Sc`	The number of strata in each cluster.
`Ystat`	A data.frame with `n`, the numbers of observations per stratum (n_s^c), and `m`, the sum of the responses per stratum (m_s^c).

See Also

`Ts.estim`

Examples

```
dimstat <- ddim(formula = Y ~ strata(Strata) + cluster(Cluster), data = bison)
dimstat
```

`Ts.estim` *Two-Step Estimator*

Description

Function that computes the two-step estimator proposed in Craiu et al. (2011) and its `print` method.

Usage

```
Ts.estim(
  formula,
  data,
  random,
  all.m.1 = FALSE,
  D = "UN(1)",
  itermax = 2000,
  tole = 1e-06
)

## S3 method for class 'Ts.estim'
print(x, ...)
```

Arguments

`formula`	A formula object, with the response on the left of a ~ operator, and the covariates on the right. The right hand side of the model must also include two special terms: a `strata` and a `cluster` term (ex. `formula = Y ~ X1 + X2 + X3 + strata(var_strata) + cluster(var_cluster)`). The `strata` and `cluster` functions (from the package **survival**) are used to identify the stratification and the cluster variables, respectively.
`data`	A data frame (or object coercible by as.data.frame to a data frame) containing the variables in the model.
`random`	A formula object, with a blank on the left of a ~ operator, and, on the right, the covariates with random coefficients among the covariate listed in the model `formula` (ex. `random = ~ X1 + X3`). The default is to add random coefficients for every covariates listed in the model `formula`.
`all.m.1`	`TRUE` if sum of Y's in all strata is 1, `FALSE` otherwise (the default). When in doubt use `FALSE` (always works, but slower than necessary if all stratum sums are 1).
`D`	The form of the between-cluster variance-covariance matrix of the regression coefficients (matrix D) : either `"UN"` for unstructured matrix D or `"UN(1)"` (the default) for diagonal matrix D.

itermax	maximal number of EM iterations (default = 2000)
tole	maximal distance between successive EM iterations tolerated before declaring convergence (default = 0.000001)
x	An object, produced by the Ts.estim function, to print.
...	Further arguments to be passed to print.default.

Details

Calls coxph from the package **survival**.

Value

beta	A vector: the regression coefficients.
se	A vector: the regression coefficients' standard errors.
vcov	A matrix: the variance-covariance matrix of the regression coefficients.
D	A matrix: estimate of the between-cluster variance-covariance matrix of the regression coefficients (matrix D).
r.effect	The random effect estimates.
coxph.warn	A list of character string vectors. If the coxph function generates one or more warnings when fitting the Cox model to a cluster, a copy of these warnings are stored in coxph.warn$Cluster_name where Cluster_name is the identification value for the cluster in the data set. A NULL list element means that coxph did not produce any warnings for that cluster.
Call	The function call.

Author(s)

Radu V. Craiu, Thierry Duchesne, Daniel Fortin and Sophie Baillargeon

References

Craiu, R.V., Duchesne, T., Fortin, D. and Baillargeon, S. (2011), Conditional Logistic Regression with Longitudinal Follow-up and Individual-Level Random Coefficients: A Stable and Efficient Two-Step Estimation Method, *Journal of Computational and Graphical Statistics*. **20**(3), 767-784.

See Also

ddim

Examples

```
# Two ways for specifying the same model
# Data: bison
# Model: covariates forest, biomass and pmeadow
# Random effects in front of forest and biomass
# Main diagonal covariance structure for D (the default)
way1 <- Ts.estim(formula = Y ~ forest  + biomass + pmeadow +
        strata(Strata) + cluster(Cluster), data = bison,
```

```
        random = ~ forest + biomass)
way1
way2 <- Ts.estim(formula = bison[,3] ~ as.matrix(bison[,c(6,8:9)]) +
        strata(bison[,2]) + cluster(bison[,1]), data = bison,
        random = ~ as.matrix(bison[,c(6,8)]))
way2

# Unstructured covariance for D
Fit <- Ts.estim(formula = Y ~ forest  + biomass + pmeadow +
        strata(Strata) + cluster(Cluster), data = bison,
        random = ~ forest + biomass, D="UN")
Fit
```

Index