

Package ‘phylopairs’

June 12, 2024

Title What the Package Does (One Line, Title Case)

Version 0.0.0.9000

Description What the package does (one paragraph).

License `use_mit_license()`, `use_gpl3_license()` or friends to pick a license

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.1

Biarch true

Depends R (>= 3.5.0)

Imports methods, Rcpp (>= 0.12.0), RcppParallel (>= 5.0.1), rstan (>= 2.18.1), rstantools (>= 2.4.0)

LinkingTo BH (>= 1.66.0), Rcpp (>= 0.12.0), RcppEigen (>= 0.3.3.3.0), RcppParallel (>= 5.0.1), rstan (>= 2.18.1), StanHeaders (>= 2.18.0)

SystemRequirements GNU make

NeedsCompilation yes

Author First Last [aut, cre] (YOUR-ORCID-ID)

Maintainer First Last <first.last@example.com>

Contents

phylopairs-package	2
betareg_stan	2
data1	4
data2	5
data3	5
data4	6
data5	6
data6	7
data7	7
data8	8
sim.cov.pairs	8

Index	9
--------------	----------

phylopairs-package	<i>The 'phylopairs' package.</i>
--------------------	----------------------------------

Description

A DESCRIPTION OF THE PACKAGE

Author(s)

Maintainer: First Last <first.last@example.com> (**ORCID**)

References

Stan Development Team (NA). RStan: the R interface to Stan. R package version 2.32.6. <https://mc-stan.org>

betareg_stan	<i>betareg_stan</i>
--------------	---------------------

Description

The function `betareg_stan` fits one of two beta regression models to a dataset in the Stan Bayesian modeling framework via the 'rstan' package. Users can choose to fit either a standard model of beta regression or a beta-regression mixed model in which there are covarying residuals in the linear predictor. For the latter, users must supply a covariance matrix. In both models, users can choose one of four link functions.

Usage

```
betareg_stan(
  des,
  y,
  link = "logit",
  covmat = NULL,
  itnum = 6000,
  chains = 4,
  coef.u = 0,
  coef.sd = 10,
  phi.shape = 0.01,
  phi.rate = 0.01,
  scale.u = -1,
  scale.sd = 1,
  cores = 4,
  ...
)
```

Arguments

<code>des</code>	A vector of predictor variable observations OR, in the case of multiple predictors, a matrix in which each column is a vector of observations of a given predictor. 'betareg.stan' adds a column of 1s to make this a design matrix whose first column corresponds to the model intercept (unless such a column already exists).
<code>y</code>	A vector of response variable observations.
<code>itnum</code>	Number of iterations to run on each chain; defaults to 6000.
<code>chains</code>	Number of chains to run; defaults to 4.
<code>coef.u</code>	Mean of the Gaussian prior for each predictor variable coefficient; defaults to 0.
<code>coef.sd</code>	SD of the Gaussian prior for each predictor variable coefficient; defaults to 10.
<code>phi.shape</code>	Shape parameter for gamma prior of beta distribution's phi parameter; defaults to 0.01.
<code>phi.rate</code>	Rate parameter for gamma prior of beta distribution's phi parameter; defaults to 0.01.
<code>scale.u</code>	Mean of the lognormal prior for the scale of the residual covariance; defaults to -1.
<code>scale.sd</code>	SD of the lognormal prior for the scale of the residual covariance; defaults to 1.
<code>cores</code>	Number of cores to be used; defaults to 4 (one chain per core).
<code>...</code>	additional arguments passed to <code>rstan::sampling</code> .
<code>cov</code>	Covariance matrix for model residuals (a Cp matrix if analyzing lineage-pair data or a phylogenetic vcv matrix if analyzing bounded species data).

Value

A list containing two elements: (1) the posterior distribution of beta model parameters, and (2) the log-likelihood of the posteriors for use in downstream analyses (e.g. the calculation of model fitting metrics like loo or waic)

Examples

```
## Example 1: Fit beta regression models with different link functions to independent data
# Load a data simulated with a logit link function
data(data5)
# Run the betareg function
result1 = betareg_stan(des=data5[,1], y=data5[,2], itnum=1000)
# Observe posterior parameter estimates
result1[[1]]
# Fit the model again but this time without the covariance matrix
result2 = betareg_stan(des=data5[,1], y=data5[,2], link="probit", itnum=1000)
# Observe posterior parameter estimates
result2[[1]]
# Compare the fit of the two models via loo and waic
loo1 = loo::loo(result1[[2]])
loo2 = loo::loo(result2[[2]])
waic1 = loo::waic(result1[[2]])
waic2 = loo::waic(result2[[2]])
loo1
loo2
waic1
```

```

waic2
loo::loo_compare(loo1, loo2)
loo::loo_compare(waic1, waic2)
## Example 2: Fit beta regression models to a simulated dataset in which the data are non-independent
## Not run:
# Also load the lineage-pair covariance matrix that arose from those simulations
data(data7)
data(sim.cov.pairs)
# Run the betareg function
result1 = betareg_stan(des=data7[,1], y=data7[,2], cov=sim.cov.pairs, itnum=1000)
# Observe posterior parameter estimates
result1[[1]]
# Fit the model again but this time without the covariance matrix
result2 = betareg_stan(des=data7[,1], y=data7[,2], itnum=1000)
# Observe posterior parameter estimates
result2[[1]]
# Fit the model again with the covariance matrix but now with a probit link function
result3 = betareg_stan(des=data7[,1], y=data7[,2], cov=sim.cov.pairs, link="probit", itnum=1000)
# Observe posterior parameter estimates
result3[[1]]
# Compare the fit of the three models via loo
loo1 = suppressWarnings(loo::loo(result1[[2]]))
loo2 = suppressWarnings(loo::loo(result2[[2]]))
loo3 = suppressWarnings(loo::loo(result3[[2]]))
loo_compare(loo1, loo2, loo3)

## End(Not run)

```

data1

Simulated Dataset 1

Description

Unbounded response, linear rshp between response and predictor, no covariance in residuals

Usage

```
data1
```

Format

A data frame with 190 rows and 2 variables:

x1 Numeric predictor variable

y Numeric response variable

Source

Simulated data generated with the script provided in the inst directory.

data2	<i>Simulated Dataset 2</i>
-------	----------------------------

Description

Unbounded response, no rshp between response and predictor, no covariance in residuals

Usage

data2

Format

A data frame with 190 rows and 2 variables:

x1 Numeric predictor variable

y Numeric response variable

Source

Simulated data generated with the script provided in the inst directory.

data3	<i>Simulated Dataset 3</i>
-------	----------------------------

Description

Unbounded response, linear rshp between response and predictor, covariance in residuals

Usage

data3

Format

A data frame with 190 rows and 2 variables:

x1 Numeric predictor variable

y Numeric response variable

Source

Simulated data generated with the script provided in the inst directory.

data4

Simulated Dataset 4

Description

Unbounded response, no rshp between response and predictor, covariance in residuals

Usage

data4

Format

A data frame with 190 rows and 2 variables:

x1 Numeric predictor variable

y Numeric response variable

Source

Simulated data generated with the script provided in the inst directory.

data5

Simulated Dataset 5

Description

Bounded response, linear rshp between response and predictor on link scale, no covariance in residuals

Usage

data5

Format

A data frame with 190 rows and 2 variables:

x1 Numeric predictor variable

y Numeric response variable

Source

Simulated data generated with the script provided in the inst directory.

data6*Simulated Dataset 6*

Description

Bounded response, no rshp between response and predictor on link scale, no covariance in residuals

Usage

data6

Format

A data frame with 190 rows and 2 variables:

x1 Numeric predictor variable

y Numeric response variable

Source

Simulated data generated with the script provided in the inst directory.

data7*Simulated Dataset 7*

Description

Bounded response, linear rshp between response and predictor on link scale, covariance in residuals

Usage

data7

Format

A data frame with 190 rows and 2 variables:

x1 Numeric predictor variable

y Numeric response variable

Source

Simulated data generated with the script provided in the inst directory.

data8	<i>Simulated Dataset 8</i>
-------	----------------------------

Description

Bounded response, no rshp between response and predictor on link scale, covariance in residuals

Usage

```
data8
```

Format

A data frame with 190 rows and 2 variables:

x1 Numeric predictor variable

y Numeric response variable

Source

Simulated data generated with the script provided in the inst directory.

sim.cov.pairs	<i>Simulated lineage-pair covariance matrix</i>
---------------	---

Description

A 190*190 covariance matrix used in simulating the example datasets in this package.

Usage

```
sim.cov.pairs
```

Format

A matrix of 190 rows and 190 columns

Source

Simulated data generated with the script provided in the inst directory.

Index

* datasets

- data1, [4](#)
- data2, [5](#)
- data3, [5](#)
- data4, [6](#)
- data5, [6](#)
- data6, [7](#)
- data7, [7](#)
- data8, [8](#)
- sim.cov.pairs, [8](#)

betareg_stan, [2](#)

- data1, [4](#)
- data2, [5](#)
- data3, [5](#)
- data4, [6](#)
- data5, [6](#)
- data6, [7](#)
- data7, [7](#)
- data8, [8](#)

phylopairs (phylopairs-package), [2](#)
phylopairs-package, [2](#)

sim.cov.pairs, [8](#)