

Package: nbhdmodel (via r-universe)

October 26, 2024

Title Neighborhood Modeling and Analysis

Version 0.1.0.9000

Description Functionality for fitting neighborhood models of McCartan, Brown, and Imai <[arXiv:2110.14014](https://arxiv.org/abs/2110.14014)>. The core methodology is described in the paper and can be implemented with any tool that can fit generalized linear mixed models (GLMMs). However, some of the preprocessing necessary to set up the GLMM can be onerous. In addition to providing a specialized GLMM routine, this package provides several preprocessing functions that, while not completely general, should be useful for others performing these kinds of analyses.

Depends R (>= 4.0.0)

Imports rlang, cli, methods, dplyr (>= 1.0.0), fastmatch, posterior, stringr, Rcpp (>= 0.12.0), RcppParallel (>= 5.0.1), lme4, rstan (>= 2.18.1), s2, sf

LinkingTo BH (>= 1.66.0), cpp11, Rcpp (>= 0.12.0), RcppEigen (>= 0.3.3.3.0), RcppParallel (>= 5.0.1), rstan (>= 2.18.1), StanHeaders (>= 2.18.0)

Suggests ggplot2, readr

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

SystemRequirements GNU make

URL <https://github.com/CoryMcCartan/nbhdmodel>

Repository <https://corymccartan.r-universe.dev>

RemoteUrl <https://github.com/CoryMcCartan/nbhdmodel>

RemoteRef HEAD

RemoteSha 6e7d39e648997ca57c9bd14d833337617c2211d6

Contents

binned_resid	2
calc_indiv_frame	2
eff_dist	3
fastAUC	4
local_area	4
nbhd_fit	5
neighborhood_model	5
plot.nbhd_fit	7
post_incl	7
simulate_neighborhood	8

Index	10
--------------	-----------

binned_resid	<i>Binned Residual Plot</i>
--------------	-----------------------------

Description

Binned Residual Plot

Usage

```
binned_resid(model, bins = 16)
```

Arguments

model	the model object, which should have fitted, resid, etc. methods.
bins	the number of bins

Value

A ggplot

calc_indiv_frame	<i>Create a model data frame for an individual respondent</i>
------------------	---

Description

This function combines individual, neighborhood, and geographic information to produce a data frame suitable for use in fitting a `neighborhood_model()`. Can be applied in a loop over respondents to generate a model frame for an entire sample.

Usage

```
calc_indiv_frame(row, nbhd, block_d, block_gr)
```

Arguments

row	A single data frame row containing relevant individual covariates for a respondent.
nbhd	The respondent's neighborhood as a vector of indices indexing the blocks in block_gr.
block_d	A data frame of census blocks, including a column centroid with an s2 point geography of each block's centroid.
block_gr	An adjacency graph object: a list containing an element graph with the adjacency list, and blocks a character vector of GEOIDs or codes corresponding to the indices in graph.

Value

A tibble that can be used inside a modeling function. Will contain the entries in row, plus the relevant entries in block_d for each block, plus columns:

- ring containing the "ring" indicator around the residence: 0 indicates the respondent's block, 1 for blocks touching the residential block, 2 for blocks touching those, etc.
- incl a binary indicator for whether the block is in the neighborhood
- dist the distance to the respondent's block
- frac_con the fraction of nearer blocks in the neighborhood this block is connected to.

 eff_dist

Get Posterior Mean of Effective Block Distance

Description

Not exported. Assumes block_d has column fips which is used inside the neighborhood column of new_resp..

Usage

```
eff_dist(fit, new_resp, block_d, proc_fn = function(x) x)
```

Arguments

fit	the model fit, from neighborhood_model()
new_resp	a single-row respondent data frame to make predictions from
block_d	the block data frame
proc_fn	a processing function that is used to prepare raw model data for fitting

Value

a numeric vector of effective distances

fastAUC	<i>Calculate AUC</i>
---------	----------------------

Description

Calculate AUC

Usage

```
fastAUC(x, y)
```

Arguments

x	the predictor
y	a binary indicator

Value

the scalar AUC

local_area	<i>Filter block data to a radius of a FIPS code</i>
------------	---

Description

Filter block data to a radius of a FIPS code

Usage

```
local_area(fips, block_d, geom_d, dist = 0.5)
```

Arguments

fips	the FIPS code to center the area at
block_d	the block data, with a \$fips column matching fips argument.
geom_d	the block geometry data
dist	the radius of the area, in miles

Value

a filtered block_d

nbhd_fit	<i>Functions for working with neighborhood fits</i>
----------	---

Description

Functions for working with neighborhood fits

Usage

```
## S3 method for class 'nbhd_fit'  
summary(object, ...)  
  
## S3 method for class 'nbhd_fit'  
coef(object, ...)  
  
## S3 method for class 'nbhd_fit'  
fixef(object, ...)  
  
## S3 method for class 'nbhd_fit'  
ranef(object, ...)  
  
## S3 method for class 'nbhd_fit'  
fitted(object, ...)  
  
## S3 method for class 'nbhd_fit'  
residuals(object, ...)  
  
## S3 method for class 'nbhd_fit'  
as.matrix(x, ...)  
  
## S3 method for class 'nbhd_fit'  
as.data.frame(x, ...)
```

Arguments

object, x	a nbhd_fit object
...	Ignored.

neighborhood_model	<i>Fit the Neighborhood Model</i>
--------------------	-----------------------------------

Description

Fits the neighborhood GLMM using a provided formula and using a Bernoulli outcome with cloglog link.

Usage

```
neighborhood_model(
  formula,
  data,
  prior_coef_scale = 2.5,
  draws = 1000,
  imp_samp = TRUE,
  init = 0,
  ...,
  hessian = TRUE,
  verbose = FALSE
)
```

Arguments

formula	a one-sided model formula. The actual formula used to fit the GLMM will be generated from this one: it will contain a term for <code>log(dist)</code> (where <code>dist</code> must be the column that encodes distance), and will have random effects based on the <code>id</code> column. The column <code>incl</code> will be used as the left-hand-side variable indicating that a block is in the neighborhood.
data	the model data frame. Should have an <code>incl</code> column for block inclusion, an <code>id</code> column with respondent IDs, and a <code>dist</code> column with distances.
prior_coef_scale	the scale of the prior on the standardized predictors.
draws	the number of approximate posterior draws to generate
imp_samp	whether to perform importance resampling on the approximate draws.
init	initial values for the model fitting function <code>rstan::optimizing()</code> .
...	other arguments to to the model fitting function <code>rstan::optimizing()</code> .
hessian	whether to compute the Hessian. Required for full inference.
verbose	if TRUE, show verbose optimization output.

Value

a fitted model object of class `nbhd_fit`, which is a list which includes the following elements:

- `map`, the MAP estimates for the parameters.
- `vcov`, the covariance matrix for the MAP estimates, calculated from the Hessian of the log posterior.
- `raw_ids` the vector of ids
- `X` the design matrix
- `y` the outcome value (`1 - incl`)
- `post`, the approximate posterior samples, from `posterior::draws_rvars`
- `lp`, the log posterior probability of each sample
- `lp_norm`, the log probability of Normal approx. to posterior for each sample

plot.nbhd_fit	<i>Plot coefficient estimates</i>
---------------	-----------------------------------

Description

50% and 90% credible intervals plotted by default.

Usage

```
## S3 method for class 'nbhd_fit'
plot(x, y = NULL, intercept = FALSE, inner_prob = 0.5, outer_prob = 0.9, ...)
```

Arguments

x	a nbhd_fit object from <code>neighborhood_model()</code> .
y	ignored
intercept	if FALSE, don't plot the intercept estimate.
inner_prob	the inner credible interval probability
outer_prob	the inner credible interval probability
...	Ignored.

Value

A ggplot.

post_incl	<i>Get Posterior Block Inclusion Probabilities</i>
-----------	--

Description

Get Posterior Block Inclusion Probabilities

Usage

```
post_incl(
  fit,
  new_resp,
  resp_id = NULL,
  block_d,
  proc_fn = function(x) x,
  use_distance = TRUE
)
```

Arguments

fit	the model fit, from <code>neighborhood_model()</code>
new_resp	a single-row respondent data frame to make predictions from
resp_id	the respondent ID, used to get the random effects. If NULL a new random effect will be simulated for each draw. If NA, the random effect will be set to zero.
block_d	the block data frame
proc_fn	a processing function that is used to prepare raw model data for fitting
use_distance	if FALSE, remove block-to-home distance from the linear predictor.

Value

a numeric vector of inclusion probabilities

`simulate_neighborhood` *Simulate a neighborhood for a respondent*

Description

Simulate a neighborhood for a respondent

Usage

```
simulate_neighborhood(
  fit,
  new_resp,
  draws = 1,
  resp_id = NULL,
  block_d,
  block_gr,
  max_ring = 10L,
  proc_fn = function(x) x
)
```

Arguments

fit	the model fit, from <code>neighborhood_model()</code>
new_resp	a single-row respondent data frame to make predictions from. Should have a neighborhood column with codes matching <code>block_gr\$blocks</code> .
draws	the number of simulated neighborhoods per respondent
resp_id	the respondent ID, used to get the random effects. If NULL a new random effect will be simulated for each draw. If NA, the random effect will be set to zero. If a numeric, the random effect will be set to this value.
block_d	the block data frame. Should have a centroid geography column.

<code>block_gr</code>	An adjacency graph object: a list containing an element graph with the adjacency list, and <code>blocks</code> a character vector of GEOIDs or codes corresponding to the indices in graph.
<code>max_ring</code>	the maximum graph distance from the starting block to allow.
<code>proc_fn</code>	a processing function that is used to prepare raw model data for fitting

Value

A list of draws integer vectors containing the indices of the blocks (in `block_d`) making up each simulated neighborhood.

Index

`as.data.frame.nbhd_fit (nbhd_fit)`, 5

`as.matrix.nbhd_fit (nbhd_fit)`, 5

`binned_resid`, 2

`calc_indiv_frame`, 2

`coef.nbhd_fit (nbhd_fit)`, 5

`eff_dist`, 3

`fastAUC`, 4

`fitted.nbhd_fit (nbhd_fit)`, 5

`fixef.nbhd_fit (nbhd_fit)`, 5

`local_area`, 4

`nbhd_fit`, 5, 6

`neighborhood_model`, 5

`neighborhood_model()`, 2, 3, 7, 8

`plot.nbhd_fit`, 7

`post_incl`, 7

`posterior::draws_rvars`, 6

`ranef.nbhd_fit (nbhd_fit)`, 5

`residuals.nbhd_fit (nbhd_fit)`, 5

`rstan::optimizing()`, 6

`simulate_neighborhood`, 8

`summary.nbhd_fit (nbhd_fit)`, 5