

Package ‘BayesPocket’

May 13, 2026

Type Package

Title Bayesian Causal Inference for Periodontal Diseases in Longitudinal Studies

Version 0.1.0

Description Implements the Mixed Treatment-State Causal Model (MTSCM), a Bayesian framework for estimating causal effects of clinical interventions on bounded continuous outcomes in longitudinal observational studies with irregular visits. The methodology is specifically designed for periodontal disease research, where discrete treatments and continuous disease states (e.g., proportion of periodontal pockets exceeding 3 mm) reciprocally influence one another under dynamic feedback. The package integrates a double-censored Tobit likelihood to handle boundary mass at zero and one, subject-specific random effects to capture within-subject correlation, and flexible tree-based ensemble priors (standard BART and Soft BART) to model complex nonlinear interactions without parametric restrictions. Causal identification is established under the potential outcomes framework via the G-computation formula, with key estimands including the Mixed Average Potential Outcome (MAPO) and the Mixed Probability of Disease Resolution (MPDR). The package provides functions for model fitting, posterior inference, and causal estimand estimation.

License GPL (>= 3)

Encoding UTF-8

RoxygenNote 7.3.2

Imports stats (>= 4.4.2), GIGrvg (>= 0.8), truncnorm (>= 1.0-9), progress (>= 1.2.3), stochtree (>= 0.1.1), SoftBart (>= 1.0.3), parallel (>= 4.4.2), pbmcapply (>= 1.5.1)

Depends R (>= 3.5)

NeedsCompilation no

Author Qingyang Liu [aut, cre] (ORCID: <https://orcid.org/0000-0003-3265-6330>),
Debdeep Pati [aut],
Yang Ni [aut],
Dipankar Bandyopadhyay [aut]

Maintainer Qingyang Liu <rh8liuqy@gmail.com>

Repository CRAN

Date/Publication 2026-05-13 08:10:22 UTC

Contents

BayesPocket	2
causal_estimand_inference	3
causal_inference_model	5
data_generation	8

Index **10**

BayesPocket *The 'BayesPocket' package.*

Description

Implements a Bayesian double-censored model for causal inference in longitudinal studies of periodontal disease progression. The package provides tools for estimating causal effects of treatments on disease outcomes, accounting for time-varying confounders and left- and right-censored outcomes. It uses a Tobit regression model with extended Bayesian additive regression trees (XBART) for flexible modeling of complex relationships. The methodology is designed for observational dental data where treatments are assigned adaptively over time. Includes functions for model fitting and posterior inference of causal estimands.

Value

This is the summary page. No return value.

Author(s)

Maintainer: Qingyang Liu <rh8liuqy@gmail.com> ([ORCID](#))

Authors:

- Debdeep Pati <dpati2@wisc.edu>
- Yang Ni <yang.ni@austin.utexas.edu>
- Dipankar Bandyopadhyay <dbandyop@vcu.edu>

`causal_estimand_inference`*Iterate Causal Estimand Calculations Over a Grid*

Description

Wrapper to iterate causal estimand calculations over a grid of previous status values.

Usage

```
causal_estimand_inference(  
  outcome_model_results,  
  df,  
  continuous_name,  
  categorical_name,  
  treatment_name,  
  treatment_value,  
  previous_status_name,  
  credible_interval_level = 0.95,  
  num_of_grids = 50  
)
```

Arguments

<code>outcome_model_results</code>	the results from the outcome model. The datatype is <code>list</code> .
<code>df</code>	the input dataframe. The datatype is <code>data.frame</code> .
<code>continuous_name</code>	the name of continuous predictors. The datatype is <code>character</code> .
<code>categorical_name</code>	the name of categorical predictors. The datatype is <code>character</code> .
<code>treatment_name</code>	the name of the treatment predictor. The datatype is <code>character</code> .
<code>treatment_value</code>	the value of the treatment variable. The datatype is <code>factor</code> .
<code>previous_status_name</code>	the name of the variable that represents previous status. The datatype is <code>character</code> .
<code>credible_interval_level</code>	the nominal level of credible intervals of causal estimand. The datatype is <code>double</code> .
<code>num_of_grids</code>	the number of grid points in <code>[0,1]</code> to evaluate previous status on. The datatype is <code>integer</code> .

Details

This function evaluates the causal estimands (such as MAPO and MPDR) across a specified grid of values for the previous disease state. For comprehensive details regarding the underlying framework, methodology, and the main model fitting procedure, please refer to [causal_inference_model](#).


```

previous_status_name = inference_output$previous_status_name,
credible_interval_level = 0.95,
num_of_grids = 2) # Example uses a small 2-point grid

# View the newly calculated closed-form causal estimands -----

# 1. Print results for the first grid point
cat("--- Results for Grid Point 1 ---\n")
print(inference_results[[1]]$mapo_summary)
print(inference_results[[1]]$mpdr_summary)

# 2. Print results for the second grid point
cat("\n--- Results for Grid Point 2 ---\n")
print(inference_results[[2]]$mapo_summary)
print(inference_results[[2]]$mpdr_summary)

```

causal_inference_model

Bayesian Mixed Treatment-State Causal Model (MTSCM)

Description

Fits a Bayesian Mixed Treatment-State Causal Model (MTSCM) tailored for longitudinal settings with irregular visits. This model is specifically designed for bounded continuous outcomes with mass at both boundaries, such as the proportion of periodontal pockets exceeding 3 mm.

Usage

```

causal_inference_model(
  df,
  y_name,
  continuous_name,
  categorical_name,
  treatment_name,
  previous_status_name,
  subjectID_name,
  num_warmup,
  num_samples,
  model_type,
  thin = 1,
  L = 50,
  alpha = 0.95,
  beta = 1.25,
  leaf_model_scale = 0.3/50,
  cutpoint_grid_size = 100,
  max_depth = 10,
  credible_interval_level = 0.95,
  print_progress = TRUE,

```

```

    random_seed = 100,
    calculate_causal_estimand = FALSE,
    previous_status_grid_size = 100
  )

```

Arguments

df the input dataframe. The datatype is `data.frame`.

y_name the name of the response variable. The datatype is character.

continuous_name the name of continuous predictors. The datatype is character.

categorical_name the name of categorical predictors. The datatype is character.

treatment_name the name of the treatment predictor. The datatype is character.

previous_status_name the name of variable that represents previous status of a subject. The datatype is character.

subjectID_name the name of variable that represents subjectID. The datatype is character.

num_warmup the number of warmup iterations. The datatype is integer.

num_samples the number of post-warmup iterations. The datatype is integer.

model_type the type of causal inference models. It must be one of "Tobit-XBART", "Tobit-SBART", "Tobit-LH", "N-XBART", "N-SBART", or "N-LH".

thin the period between saved samples. This should typically be left at its default (no thinning) unless memory is a problem. The datatype is integer.

L the number of trees. The datatype is integer.

alpha the tree prior parameters. $\alpha \times (1 + \text{depth})^{-\beta}$ represents the prior probability of splitting at one of the cutpoints. Check equation (4) and related descriptions from "Stochastic Tree Ensembles for Regularized Nonlinear Regression" for more details. The datatype is double.

beta the tree prior parameters. $\alpha \times (1 + \text{depth})^{-\beta}$ represents the prior probability of splitting at one of the cutpoints. Check equation (4) and related descriptions from "Stochastic Tree Ensembles for Regularized Nonlinear Regression" for more details. The datatype is double.

leaf_model_scale the prior variance on leaf mean equals to `leaf_model_scale/L`. The datatype is double.

cutpoint_grid_size the number of cutoff points in XBART. The datatype is integer.

max_depth the maximum depth of tree allowed. The datatype is integer.

credible_interval_level the nominal level of credible intervals of causal estimand. The datatype is double.

print_progress whether print progress bar or not. The datatype is boolean.

random_seed the random seed of the MCMC sampler. The datatype is integer.


```

continuous_name = c("previous_value",
                    "confounder"),
categorical_name = c("treatment"),
treatment_name = "treatment",
previous_status_name = "previous_value",
subjectID_name = "subjectID",
num_warmup = 2,
num_samples = 2,
model_type = "Tobit-XBART",
thin = 1,
L = 5,
alpha = 0.95,
beta = 1.25,
leaf_model_scale = 0.3/5,
cutpoint_grid_size = 100,
max_depth = 10,
credible_interval_level = 0.95,
random_seed = 100,
calculate_causal_estimand = FALSE,
previous_status_grid_size = 2)

```

data_generation

Data Generation Program

Description

Generates simulated data for evaluating the causal inference models.

Usage

```
data_generation(random_seed, N, sigma, sigma_u)
```

Arguments

random_seed	a single random seed for reproducibility. The datatype is integer.
N	the total number of subjects to simulate. The datatype is integer.
sigma	the global error standard deviation. The datatype is double.
sigma_u	the standard deviation of the subject-level random effects. The datatype is double.

Value

data_generation returns a simulated data.frame.

Examples

```
df1 <- data_generation(random_seed = 100,  
                       N = 100,  
                       sigma = 0.2,  
                       sigma_u = 0.1)  
  
print(head(df1))
```

Index

BayesPocket, [2](#)

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

causal_estimand_inference, [3](#)

causal_inference_model, [3](#), [4](#), [5](#)

data_generation, [8](#)