Package: BayesDecon (via r-universe)

October 13, 2026

Type Package

Title Density Deconvolution Using Bayesian Semiparametric Methods

Version 0.1.7

Maintainer Mainak Manna <mainakmanna29@utexas.edu>

Description Estimates the density of a variable in a measurement error setup, potentially with an excess of zero values. For more details see Sarkar (2021). <doi:10.1080/01621459.2020.1782220>.

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 7.3.3

LinkingTo Rcpp, RcppArmadillo

Imports Rcpp, msm, corpcor, ks, mvtnorm

Suggests ggplot2, gridExtra, parallel, foreach, doParallel, RColorBrewer

Depends R (>= 4.1.0)

NeedsCompilation yes

Author Blake Moya [aut], Mainak Manna [cre, aut], Abhra Sarkar [aut], The University of Texas at Austin [cph, fnd]

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

Date/Publication 2026-07-27 12:38:40 UTC

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

RemoteRef HEAD

RemoteSha b736c2ffd8c3b130ab7b94421c54625a3529c77f

Contents

autoplot.bdeconv_result 2
BayesDecon_data_simulated 2

bdeconv 3
dens_x 5
dist_x 6
mean.bdeconv_result 7
plot.bdeconv_result 7
print.bdeconv_result 8
subset_result 9

Index **10**

`autoplot.bdeconv_result` *Visualization of* `bdeconv_result` *object using* `ggplot2`

Description

Visualization of `bdeconv_result` object using `ggplot2`

Usage

```
autoplot.bdeconv_result(object)
```

Arguments

`object`	A `bdeconv_result` object.

Value

A collection of plots illustrating various densities and related functions made using `ggplot2`.

`BayesDecon_data_simulated` *Simulated dataset for BayesDecon*

Description

A small example dataset used to demonstrate bayesdecon functions. There are 1000 observations each with 3 replicates for 3 variables V1, V2, V3. The first column represents the observations.

Usage

```
BayesDecon_data_simulated
```

Format

A data frame with 3000 rows and 4 variables.

Source

Generated for package examples.

---

bdeconv *Obtaining MCMC samples from the posterior*

Description

Obtaining MCMC samples from the posterior

Usage

```
bdeconv(
  obj,
  scaling = TRUE,
  lwr = 0,
  upr = NULL,
  wgts = NULL,
  regular_var = NULL,
  res_type = c("final", "mean", "all"),
  err_type = c("flex", "norm", "lapl"),
  var_type = c("heteroscedastic", "homoscedastic"),
  na_rm = FALSE,
  control = list(),
  progress = FALSE,
  core_num = 1,
  parallel = FALSE
)
```

Arguments

obj	A data frame where the first column is subject labels.
scaling	Weather or not scaling is required for the variables. Default is TRUE.
lwr	If scaling = TRUE, unused. If scaling = FALSE, lower bound for the deconvolved estimates. For zero inflated variables, automatically taken to be 0. Default is 0.
upr	If scaling = TRUE, the ratio of an upper bound for the deconvolved estimates to the upper bound of the surrogates. Default is 0.5. If scaling = FALSE, upper bound for the deconvolved estimates. Default is 0.5 * upper bound of the surrogates. It can be either a scalar or a vector with a length that matches the number of numeric columns in obj.
wgts	Unused.
regular_var	Specifies which variables to consider as regular; all others are treated as zero inflated.
res_type	The type of result returned from the MCMC sampler.

err_type	The shape of the error distribution used for deconvolution.
var_type	whether to use homogeneous or heterogeneous error distribution.
na_rm	A logical value indicating whether to ignore rows which contain NA values in numeric columns.
control	A list of arguments specifying prior hyper-parameters and parameters of the MCMC sampler.
progress	Whether a progress bar for MCMC will be shown or not. Default is FALSE.
core_num	If parallelized, number of cores to be used.
parallel	For multivariate data. Whether the pre-processing of each variable will be parallelized or not. Default is FALSE.

Details

Performs univariate and multivariate Bayesian semiparametric density deconvolution when the true latent distribution of main interest and the measurement error distribution are both unknown, the errors may be conditionally heteroscedastic, and replicated proxies are available for each individual in the sample.

Notably, the support of the univariate densities of interest for zero inflated variables must be non-negative; so negative proxies are not allowed for zero inflated variables. For regular variables, if negative proxies are observed, scaling must be FALSE and the true range of the support must be provided through lwr and upr. All hyper-parameters of the model and variables necessary for the MCMC sampler can be specified through a list via the control argument; some of the most important ones are listed.

- nsamp: Number of finally retained MCMC samples after burn-in and thinning. Default is 1000.
- nburn: Burn-in period. Default is 2000.
- nthin: Thinning interval. Default is 5.
- niter_uni: For multivariate data, number of iteration for pre-processing of each variable. Default is 500.
- K_t: Number of knot points for bspline. Default is 10.
- K_x: Number of mixture components in the model for the density of X when the surrogates are strictly continuous. Not applicable when the surrogates are zero-inflated in which case the density is modeled using normalized bsplines. Default is 10.
- K_epsilon: Number of mixture components in the model for the density of scaled errors when applicable. Default is 10.
- sigmasq_epsilon: variance parameter for the normal prior of the scaled error mixture components' location parameter. Default is 4.
- a_epsilon: shape parameter for the inverse gamma prior of the variances of scaled error mixture components. Default is 3.
- b_epsilon: rate parameter for the inverse gamma prior of the variances of scaled error mixture components. Default is 1.
- a_vartheta: shape parameter for the inverse gamma prior of the variances of bspline coefficients corresponding to the variance function. Default is 100.

- b_vartheta: rate parameter for the inverse gamma prior of the variances of bspline coefficients corresponding to the variance function. Default is 1.
- a_xi: shape parameter for the inverse gamma prior of the variances of bspline coefficients corresponding to the densities of episodic components. Default is 100.
- b_xi: rate parameter for the inverse gamma prior of the variances of bspline coefficients corresponding to the densities of episodic components. Default is 10.
- a_beta: shape parameter for the inverse gamma prior of the variances of bspline coefficients corresponding to the probability of consumption function. Default is 100.
- b_beta: rate parameter for the inverse gamma prior of the variances of bspline coefficients corresponding to the probability of consumption function. Default is 1.
- x_grd_res: Specifies the number of grid points used over the range of X for numerical approximation.

Value

A bdeconv_result object, containing all the posterior samples, suitable for further analysis and visualization.

Examples

```
data(BayesDecon_data_simulated)
set.seed(123)
### Selecting first 300 rows for demonstration.
### For best results, using the whole data is recommended.
result_uni <- bdeconv(BayesDecon_data_simulated[1:300,c(1,2)],
 upr = 1, res_type = "all",
 control = list(nsamp = 500, nburn = 0, nthin = 1),
 progress = TRUE)
plot(result_uni)


set.seed(123)
result_mult <- bdeconv(BayesDecon_data_simulated, upr = 1, res_type = "all", progress = TRUE)
plot(result_mult)
plot(result_mult[,c(1,2)])
value = rbind(c(1,1,1),c(2,2,2),c(3,3,3),c(4,4,4))
colMeans(dist_x(result_mult, vals = value))
colMeans(dens_x(result_mult, vals = value))
print(result_mult)
```

---

dens_x *Estimated density function for the true variable(s)*

Description

Estimated density function for the true variable(s)

Usage

```
dens_x(obj, vals, expand = FALSE)
```

Arguments

obj	A bdeconv_result object.
vals	A numeric vector (for univariate) or matrix (for multivariate) on which to evaluate the density.
expand	Used only for mv_res objects. A logical value indicating whether or not to expand the matrix of values.

Value

A numeric matrix or array of density values for which each row represents the estimated density corresponding to each MCMC iteration.

---

dist_x *Estimated cumulative distribution function for the true variable(s)*

Description

Estimated cumulative distribution function for the true variable(s)

Usage

```
dist_x(obj, vals, expand = FALSE)
```

Arguments

obj	A bdeconv_result object.
vals	A numeric vector (for univariate) or matrix (for multivariate) on which to evaluate the CDF.
expand	Used only for mv_res objects. A logical value indicating whether or not to expand the matrix of values.

Value

A numeric matrix or array of cumulative distribution values for which each row represents the estimated CDF corresponding to each MCMC iteration.

mean.bdeconv_result *Computing mean of MCMC samples*

Description

Computing mean of MCMC samples

Usage

```
## S3 method for class 'bdeconv_result'
mean(x, ...)
```

Arguments

x	A bdeconv_result object.
...	Unused.

Details

Where a direct point-wise mean is not appropriate (i.e. for mixture density results), mixtures are expanded to included every component from every sampled density with weights divided by n, the number of samples from which the mean is being taken.

Value

A bdeconv_result object with one sample constituting the mean

plot.bdeconv_result *Visualization of* bdeconv_result *object*

Description

Visualization of bdeconv_result object

Usage

```
## S3 method for class 'bdeconv_result'
plot(x, ..., use_ggplot = TRUE)
```

Arguments

x	A bdeconv_result object.
...	Unused.
use_ggplot	A logical value indicating whether the plots will be generated using ggplot2 or using base plot functions.

Details

By default it uses `ggplot2` package if it is available in the system library, otherwise it falls back to the base R plot function.

Value

A collection of plots illustrating various densities and related functions.

`print.bdeconv_result` *Printing summary of MCMC samples for* `bdeconv_result` *object*

Description

Printing summary of MCMC samples for `bdeconv_result` object

Usage

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

Arguments

`x`	A `bdeconv_result` object.
`...`	Unused.

Details

For episodic components, this gives five number summary for the observed surrogates (y), the latent surrogates corresponding to the consumption days (w), the true intake patterns (x) and the errors (u). For regular components, this gives five number summary for y, x and u only.

Value

Prints the five number summary for the relevant variables.

subset_result *Subset Methods for* bdeconv_result *object*

Description

Subset Methods for bdeconv_result object

Usage

```
## S3 method for class 'zi_res'
obj[i]

## S3 method for class 'nc_res'
obj[i]

## S3 method for class 'mv_res'
obj[i, j]
```

Arguments

obj	A bdeconv_result object.
i	An integer row index (or indices) to select some specific MCMC iterations.
j	And integer column index (or indices) to select some specific set of variables (For multivariate case only).

Value

A bdeconv_result object containing the selected iterations of posterior samples for the specified variables (in the multivariate case), suitable for further analysis and visualization.

Index

∗ **datasets**
BayesDecon_data_simulated, 2
[.mv_res *(*subset_result*)*, 9
[.nc_res *(*subset_result*)*, 9
[.zi_res *(*subset_result*)*, 9

autoplot.bdeconv_result, 2

BayesDecon_data_simulated, 2
bdeconv, 3

dens_x, 5
dist_x, 6

mean.bdeconv_result, 7

plot.bdeconv_result, 7
print.bdeconv_result, 8

subset_result, 9