Package 'bhrcr'

January 28, 2019

Type Package

Title Bayesian Hierarchical Regression on Clearance Rates in the Presence of Lag and Tail Phases

Version 1.0.3

Date 2019-01-27

Author Colin B. Fogarty [cre] <cfogarty@mit.edu>,
Saeed Sharifi-Malvajerdi [aut] <saeedsh@wharton.upenn.edu>,
Feiyu Zhu [aut] <feiyuzhu@sas.upenn.edu>

Maintainer Colin B. Fogarty <cfogarty@mit.edu>

Description An implementation of the Bayesian Clearance Estimator (Fogarty et al. (2015) <doi:10.1111/biom.12307>).
It takes serial measurements of a response on an individual (e.g., parasite load after treatment) that is decaying over time and performs Bayesian hierarchical regression of the clearance rates on the given covariates. This package provides tools to calculate WWARN PCE (WorldWide Antimalarial Resistance Network's Parasite Clearance Estimator) estimates of the clearance rates as well. A tutorial appeared in Sharifi-Malvajerdi et al. (2019) <doi: 10.1186/s12936-018-2631-8>.

Depends R (>= 3.0.0)

Imports stats, MCMCpack, MASS, msm, mvtnorm, survival, AER, Cairo, graphics, grDevices, utils

Suggests knitr, rmarkdown

VignetteBuilder knitr

License GPL-3

RoxygenNote 6.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2019-01-28 19:10:12 UTC

R topics documented:

`bhrcr-package` *Bayesian Hierarchical Regression on Clearance Rates*

Description

"bhrcr" provides tools for calculating, analyzing, and visualizing parasite clearance rates in the presence of "lag" and "tail" phases through the use of a Bayesian hierarchical linear model. The main function for the Bayesian hierarchical linear model is clearanceEstimatorBayes. Also the function calculatePCE performs the method presented in Flegg et al (2011). The hierarchical approach enables us to appropriately incorporate the uncertainty in both estimating clearance rates in patients and assessing the potential impact of covariates on these rates into the posterior intervals generated for the parameters associated with each covariate. Furthermore, it permits users to incorporate information about individuals for whom there exists only one observation time before censoring, which alleviates a systematic bias affecting inference when these individuals are excluded. The detailed model and simulation study are presented in the paper "Bayesian Hierarchical Regression on Clearance Rates in the Presence of Lag and Tail Phases with an Application to Malaria Parasites" by Fogarty et al. (2015).

References

Flegg, J. A., Guerin, P. J., White, N. J., & Stepniewska, K. (2011). Standardizing the measurement of parasite clearance in falciparum malaria: the parasite clearance estimator. Malaria journal, 10(1), 339.

Fogarty, C. B., Fay, M. P., Flegg, J. A., Stepniewska, K., Fairhurst, R. M., & Small, D. S. (2015). Bayesian hierarchical regression on clearance rates in the presence of "lag" and "tail" phases with an application to malaria parasites. Biometrics, 71(3), 751-759.

`calculatePCE` *WWARN Parasite Clearance Estimator (PCE)*

Description

This function is a wrapper function of the WWARN PCE method to calculate the parasite clearance rates of a given data set of patient profiles. The function returns the output data frame and it also saves more comprehensive outputs under a folder named "PceEstimates" under the current working directory.

Usage

```
calculatePCE(data, detect.limit = 15, outlier.detect = TRUE, ...)
```

Arguments

`data`	a data frame containing the profiles of patients. This data frame must contain `id`, `time`, and `count` columns, in that order. The first column represents the IDs of patients. The second and third columns contain parasite measurements (per microliter) in different times.
`detect.limit`	detection limit of the parasite density in blood
`outlier.detect`	indicator of whether or not to use Flegg's outlier detection method
`...`	additional parameters.

Details

This function gives users a way to calculate the parasite clearance rates by using the method in Flegg et al. (2011). Users can compare the results with that given by our method of Bayesian heierarchical model. The output is saved under a folder named "PceEstimates" under the current working directory. `data` should be a data frame in the form of the example data `pursat` provided in this package. `detect.limit` is the detection limit of the parasite density in blood. The default value is set to be 15. `outlier.detect` is an indicator users can turn off if the dataset has already been cleaned. Otherwise, it is always recommended to set `outlier.detect = TRUE` to let the program automatically detect outliers in the dataset.

Value

All results are saved under a folder named "PceEstimates" under the current working directory.

`output`	Output data frame. If `outlier.detect = TRUE`, the cleaned data frame will be returned.

Author(s)

Colin B. Fogarty <cfogarty@mit.edu>, Saeed Sharifi-Malvajerdi <saeedsh@wharton.upenn.edu>, Feiyu Zhu <feiyuzhu@sas.upenn.edu>

References

Flegg, J. A., Guerin, P. J., White, N. J., & Stepniewska, K. (2011). Standardizing the measurement of parasite clearance in falciparum malaria: the parasite clearance estimator. Malaria journal, 10(1), 339.

Examples

```
data("pursat")
output <- calculatePCE(data = pursat, detect.limit = 15, outlier.detect = TRUE)
```

---

clearanceEstimatorBayes

Bayesian Hierarchical Regression on Clearance Rates

Description

`clearanceEstimatorBayes` estimates the parasite clearance rates by using a Bayesian hierarchical model. Moreover, it provides regression analysis of clearance rates on given covariates.

Usage

```
clearanceEstimatorBayes(data, covariates = NULL, seed = 1234,
  detect.limit = 40, outlier.detect = TRUE, conf.level = 0.95,
  niteration = 1e+05, burnin = 500, thin = 50,
  filename = "output.csv")
```

Arguments

`data`	a data frame containing the profiles of patients. This data frame must contain `id`, `time`, and `count` columns, in that order. The first column represents the IDs of patients. The second and third columns contain parasite measurements (per microliter) in different times.
`covariates`	an optional data frame containing individual level covariates. This argument may be `NULL`, in which case estimation of clearance rates is of primary interest.
`seed`	a user-specified number used to initialize a pseudorandom number generator. The default value is set to be 1234 for reproducibility. If `seed = NULL`, then its value will be automatically obtained from the system clock.
`detect.limit`	detection limit of the parasite density in blood (parasites per microliter)
`outlier.detect`	indicator of whether or not to use Flegg's outlier detection method. `outlier.detect = TRUE` is recommended.
`conf.level`	required confidence level for reporting credible intervals
`niteration`	total number of simulations after the burn-in period

burnin	length of the burn-in period in the MCMC used in clearanceEstimatorBayes
thin	step size of the thinning process in the MCMC used in clearanceEstimatorBayes
filename	the name of the csv file used to store some output elements. This file contains id, clearance.mean, lag.median, and tail.median.

Details

This function estimates parasite clearance rates, along with the effect of covariates on them, by using the Bayesian hierarchical model which was introduced in Fogarty et al. (2015). A change point model is used on the log of the parasite densities to account for three potential phases: (1) a constant phase (the lag phase); (2) a phase with a linear decrease (decay phase); (3) another constant phase (the tail phase). Hence the estimation of the parasite clearance rate is only based on observations within the decay phase. The Bayesian approach allows us to treat the delineation between lag, decay, and tail phases within an individual's clearance profile as themselves being random variables, thus taking into account the additional uncertainty of boundaries between phases. Details are in Fogarty et al. (2015).

Value

The function summary (i.e., summary.bhrcr) can be used to obtain a summary of the results. clearanceEstimatorBayes returns an object of class "bhrcr" which is a list containing:

CALL	function call
clearance.post	posterior distributions of clearance rates
clearance.mean	mean values of the posterior distributions of clearance rates
clearance.median	median values of the posterior distributions of clearance rates
intercept.post	posterior distributions of the intercepts (alpha_i's) in the model
gamma.post	posterior distribution of gamma
gamma.post.thin	thinned posterior sample of gamma
gamma.mean	mean values of the posterior distribution of gamma
gamma.median	median values of the posterior distribution of gamma
gamma.CI	Credible intervals for gamma
halflifeslope.post	posterior distribution for the effect of covariates on log half-lives
halflifeslope.mean	mean values of the posterior distribution for the effect of covariates on log half-lives
halflifeslope.median	median values of the posterior distribution for the effect of covariates on log half-lives
halflifeslope.CI	Credible intervals for the effect of covariates on log half-lives
predicted.pce	PCE estimates

eta.post	posterior distribution of eta
changelag.post	posterior distributions of changetime between lag and decay phases
changetail.post	posterior distributions of changetime between decay and tail phases
lag.median	median values of the posterior distributions of changetime between lag and decay phases
tail.median	median values of the posterior distributions of changetime between decay and tail phases
var.epsilon.post	posterior variance of epsilon after simulation
var.error.post	thinned posterior sample of variance of epsilon
var.alpha.post	posterior distribution of variance of alpha
var.beta.post	posterior distribution of variance of beta
index	a list containing each patient's indices in the data
counts	Original parasite counts of all patients
counts.current	Parasite counts of all patients after sampling censored measurements
t.overall	measurement times of all patients
p.lag	posterior value of the priori probability of there being a lag phase after simulation
p.lag.thin	thinned posterior sample of the priori probability of there being a lag phase
p.tail	posterior value of the priori probability of there being a tail phase after simulation
p.tail.thin	thinned posterior sample of the priori probability of there being a tail phase
var1.post	posterior distribution of c^2
var2.post	posterior distribution of d^2
mu1.post	posterior distribution of a
mu2.post	posterior distribution of b
detect.limit	the detection limit of parasitemia
lag.post	posterior distributions of index of changetime between lag and decay phases
lag2.post	posterior distributions of index of changetime between decay and tail phases
theta.post	posterior distributions of log-parasite-count's mean in lag phase
theta2.post	posterior distributions of log-parasite-count's mean in tail phase
burnin	length of the burn-in period

Author(s)

Colin B. Fogarty <cfogarty@mit.edu>, Saeed Sharifi-Malvajerdi <saeedsh@wharton.upenn.edu>, Feiyu Zhu <feiyuzhu@sas.upenn.edu>

References

Flegg, J. A., Guerin, P. J., White, N. J., & Stepniewska, K. (2011). Standardizing the measurement of parasite clearance in falciparum malaria: the parasite clearance estimator. Malaria journal, 10(1), 339.

Fogarty, C. B., Fay, M. P., Flegg, J. A., Stepniewska, K., Fairhurst, R. M., & Small, D. S. (2015). Bayesian hierarchical regression on clearance rates in the presence of "lag" and "tail" phases with an application to malaria parasites. Biometrics, 71(3), 751-759.

Examples

```
data("pursat")
data("pursat_covariates")
out <- clearanceEstimatorBayes(data = pursat, covariates = pursat_covariates, outlier.detect = TRUE,
                               niteration = 200, burnin = 50, thin = 10)
```

`diagnostics` *Diagnostics Function for MCMC*

Description

`diagnostics` provides diagnostic analysis for the MCMC process used in the main function `clearanceEstimatiorBayes`.

Usage

```
diagnostics(object, ...)
```

Arguments

`object`	an object of class `bhrcr`, given by `clearanceEstimatorBayes`.
`...`	additional parameters.

Details

This function provides diagnostic analysis such as trace plots, ACF and PACF plots for some important parameters in the simulation process of Gibbs sampling. With these diagnostic plots, we can be assured that we get the results after we have reached stationarity and have thinned sufficiently.

Value

the directory location under which all the output is saved.

Author(s)

Colin B. Fogarty <cfogarty@mit.edu>, Saeed Sharifi-Malvajerdi <saeedsh@wharton.upenn.edu>, Feiyu Zhu <feiyuzhu@sas.upenn.edu>

Examples

```
data("posterior")
diagnostics(posterior)


data("pursat")
data("pursat_covariates")
out <- clearanceEstimatorBayes(data = pursat, covariates = pursat_covariates,
                               niteration = 200, burnin = 50, thin = 10)
diagnostics(out)
```

`plot.bhrcr` *Bayesian Clearance Estimator Plotting*

Description

`plot.bhrcr` plots the posterior results from `clearanceEstimatorBayes`.

Usage

```
## S3 method for class 'bhrcr'
plot(x, plot.post = T, id.plot = NULL, thin = NULL,
  ...)
```

Arguments

`x`	output given by `clearanceEstimatorBayes`
`plot.post`	indicator of whether or not the posterior samples should be plotted
`id.plot`	patients' IDs
`thin`	an optional vector showing which posterior samples to be plotted
`...`	additional arguments passed to the `plot.bhrcr` function

Details

This function plots clearance profile of each individual along with their fitted Bayesian model, Flegg's PCE estimates, and posterior samples.

Value

the directory location under which all the plots are saved.

Author(s)

Colin B. Fogarty <cfogarty@mit.edu>, Saeed Sharifi-Malvajerdi <saeedsh@wharton.upenn.edu>, Feiyu Zhu <feiyuzhu@sas.upenn.edu>

Examples

```
data("posterior")
plot(posterior)


data("pursat")
data("pursat_covariates")
out <- clearanceEstimatorBayes(data = pursat, covariates = pursat_covariates,
                               niteration = 200, burnin = 50, thin = 10)
plot(out)
```

`posterior` *Description of the Dataset* `posterior.rda`

Description

The output given by `clearanceEstimatorBayes` in the slow demo example, which is calculated and saved ahead of time to save users' time.

Usage

```
data("posterior")
```

Format

A data frame of class "bhrcr".

`print.bhrcr` *Print Function for the Bayesian Clearance Estimator*

Description

`print.bhrcr` prints the estimated effect of covariates on both log clearance rates and log half-lives.

Usage

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

Arguments

x	an object of class bhrcr, given by clearanceEstimatorBayes.
...	additional parameters.

Details

This function prints the posterior mean value of gamma, which represents the effect of covariates on log clearance rates. It also prints the estimated impact of covariates on log half-lives.

Value

the input object is returned silently.

Author(s)

Colin B. Fogarty <cfogarty@mit.edu>, Saeed Sharifi-Malvajerdi <saeedsh@wharton.upenn.edu>, Feiyu Zhu <feiyuzhu@sas.upenn.edu>

Examples

```
data("posterior")
print(posterior)


data("pursat")
data("pursat_covariates")
out <- clearanceEstimatorBayes(data = pursat, covariates = pursat_covariates,
                               niteration = 200, burnin = 50, thin = 10)
print(out)
```

---

pursat *Description of the Dataset* pursat.rda

Description

The pursat dataset consists of P. falciparum clearance profiles of 110 patients, measured in 2009 and 2010 in the Pursat province of Western Cambodia. Parasite densities were measured every 6 hours, and the detection limit was 15 parasites per microliter. All 110 individuals were observed until no parasites were found in their blood.

Usage

```
data("pursat")
```

Format

A data frame with following variables

id patients' IDs

time measurement times

count parasite counts in blood

`pursat_covariates` *Description of the Dataset* `pursat_covariates.rda`

Description

Individual level covariates of patients in the `pursat` dataset.

Usage

```
data("pursat_covariates")
```

Format

A data frame with following variables

Sex A factor variable with two levels `F` and `M`

agegroup 21+ (21 years of age or older), or 21- (younger than 21 years)

vvkv whether or not an individual was from Veal Veng or Kranvanh

HbE the number of alleles of Hemoglobin E variant

athal the number of alleles of alpha-thalassaemia variant

g6pd the number of alleles of G6PD deficient variant

lnPf0 Log initial parasite density

year2010 `TRUE` if 2010, `FALSE` if 2009

group 1 if group 1, 0 if group 2

summary.bhrcr *Summary Statistics for the Bayesian Clearance Estimator*

Description

summary.bhrcr provides summary statistics for the effect of covariates on both log clearance rates and log half-lives.

Usage

```
## S3 method for class 'bhrcr'
summary(object, ...)
```

Arguments

object	an object of class bhrcr, given by clearanceEstimatorBayes.
...	additional parameters.

Details

This function provides mean, median, and credible intervals for gamma, which represents the effect of covariates on log clearance rates. It also provides those statistics for the effect of covariates on log half-lives.

Value

the input object is returned silently.

Author(s)

Colin B. Fogarty <cfogarty@mit.edu>, Saeed Sharifi-Malvajerdi <saeedsh@wharton.upenn.edu>, Feiyu Zhu <feiyuzhu@sas.upenn.edu>

Examples

```
data("posterior")
summary(posterior)


data("pursat")
data("pursat_covariates")
out <- clearanceEstimatorBayes(data = pursat, covariates = pursat_covariates,
                               niteration = 200, burnin = 50, thin = 10)
summary(out)
```

Index