Package: NobBS (via r-universe)

September 5, 2024

Title Nowcasting by Bayesian Smoothing

Version 1.0.0

Description A Bayesian approach to estimate the number of
occurred-but-not-yet-reported cases from incomplete,
time-stamped reporting data for disease outbreaks. 'NobBS'
learns the reporting delay distribution and the time evolution
of the epidemic curve to produce smoothed nowcasts in both
stable and time-varying case reporting settings, as described
in McGough et al. (2020) <doi:10.1371/journal.pcbi.1007735>.

Depends R (>= 3.3.0)

SystemRequirements JAGS (http://mcmc-jags.sourceforge.net/) for
analysis of Bayesian hierarchical models

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports dplyr, rlang, rjags, coda, magrittr

RoxygenNote 7.2.3

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

RemoteUrl https://github.com/sarahhbellum/nobbs

RemoteRef HEAD

RemoteSha 7216e335aceae9f23f15d141abaf5436b527919b

Contents

---

denguedat   *denguedat: Dengue fever reporting data from Puerto Rico*

Description

Surveillance data from CDC Division of Vector-Borne Diseases. 1990-2010 case reporting data included. The first column, `onset_week`, indicates the week of symptom onset. The second column, `report_week`, indicates the week of case report. The third column, `gender`, indicates the gender of the infected individual (randomly assigned with 0.5:0.5 probability of "Male"/"Female"). This column may be used to produce stratified nowcasts using the function `NobBS.strat`.

Usage

```
data(denguedat)
```

Format

A data frame.

Examples

```
data(denguedat)
nowcast <- NobBS(denguedat, as.Date("1990-04-09"),units="1 week",onset_date="onset_week",
report_date="report_week")
nowcast$estimates
```

NobBS   *Produce smooth Bayesian nowcasts of incomplete, time-stamped reporting data.*

Description

Nowcasting is useful to estimate the true number of cases when they are unknown or incomplete in the present because of reporting delays. 'NobBS' is a Bayesian nowcasting approach that learns from the reporting delay distribution as well as the temporal evolution of the epidemic curve to estimate the number of occurred but not yet reported cases for a given date.

Usage

```
NobBS(
  data,
  now,
  units,
  onset_date,
  report_date,
  moving_window = NULL,
```

```
  max_D = NULL,
  cutoff_D = NULL,
  proportion_reported = 1,
  quiet = TRUE,
 specs = list(dist = c("Poisson", "NB"), alpha1.mean.prior = 0, alpha1.prec.prior =
   0.001, alphat.shape.prior = 0.001, alphat.rate.prior = 0.001, beta.priors = NULL,
   param_names = NULL, conf = 0.95, dispersion.prior = NULL, nAdapt = 1000, nChains = 1,
     nBurnin = 1000, nThin = 1, nSamp = 10000)
)
```

Arguments

data	A time series of reporting data in line list format (one row per case), with a column onset_date indicating date of case onset, and a column report_date indicating date of case report.
now	An object of datatype Date indicating the date at which to perform the nowcast.
units	Time scale of reporting. Options: "1 day", "1 week".
onset_date	In quotations, the name of the column of datatype Date designating the date of case onset. e.g. "onset_week"
report_date	In quotations, the name of the column of datatype Date designating the date of case report. e.g. "report_week"
moving_window	Size of moving window for estimation of cases (numeric). The moving window size should be specified in the same date units as the reporting data (i.e. specify 7 to indicate 7 days, 7 weeks, etc). Default: NULL, i.e. takes all historical dates into consideration.
max_D	Maximum possible delay observed or considered for estimation of the delay distribution (numeric). Default: (length of unique dates in time series)-1 ; or, if a moving window is specified, (size of moving window)-1
cutoff_D	Consider only delays d<=max_D? Default: TRUE. If cutoff_D=TRUE, delays beyond max_D are ignored. If cutoff_D=FALSE, max_D is interpreted as delays>=max_D but within the moving window given by moving_window.
proportion_reported	
	A decimal greater than 0 and less than or equal to 1 representing the proportion of all cases expected to be reported. Default: 1, e.g. 100 percent of all cases will eventually be reported. For asymptomatic diseases where not all cases will ever be reported, or for outbreaks in which severe under-reporting is expected, change this to less than 1.
quiet	Suppress all output and progress bars from the JAGS process. Default: TRUE.
specs	A list with arguments specifying the Bayesian model used: dist (Default: "Poisson"), beta.priors (Default: 0.1 for each delay d), nSamp (Default: 10000), nBurnin (Default: 1000), nAdapt (Default: 1000), nChains (Default: 1), nThin (Default: 1), alphat.shape.prior (Default: 0.001), alphat.rate.prior (Default: 0.001), alpha1.mean.prior (Default: 0), alpha1.prec.prior (Default: 0.001), dispersion.prior (Default: NULL, i.e. no dispersion. Otherwise, enter c(shape,rate) for a Gamma distribution.), conf (Default: 0.95), param_names (Default: NULL, i.e. output for all parameters is provided: c("lambda","alpha","beta.logged","tau2.alpha"

See McGough et al. 2019 (https://www.biorxiv.org/content/10.1101/663823v1)
for detailed explanation of these parameters.).

Value

The function returns a list with the following elements: `estimates`, a 5-column data frame contain-
ing estimates for each date in the window of predictions (up to "now") with corresponding date of
case onset, lower and upper bounds of the prediction interval, and the number of cases for that onset
date reported up to 'now'; `estimates.inflated`, a Tx4 data frame containing estimates inflated
by the proportion_reported for each date in the time series (up to "now") with corresponding date of
case onset, lower and upper bounds of the prediction interval, and the number of cases for that onset
date reported up to 'now'; `nowcast.post.samples`, vector of 10,000 samples from the posterior
predictive distribution of the nowcast, and `params.post`, a 10,000xN dataframe containing 10,000
posterior samples for the "N" parameters specified in specs[["param_names"]]. See McGough et al.
2019 (https://www.biorxiv.org/content/10.1101/663823v1) for detailed explanation of parameters.

Notes

'NobBS' requires that JAGS (Just Another Gibbs Sampler) is downloaded to the system. JAGS can
be downloaded at <http://mcmc-jags.sourceforge.net/>.

Examples

```
# Load the data
data(denguedat)
# Perform default 'NobBS' assuming Poisson distribution, vague priors, and default specifications.
nowcast <- NobBS(denguedat, as.Date("1990-04-09"),units="1 week",onset_date="onset_week",
report_date="report_week")
nowcast$estimates
```

NobBS.strat   *Stratified nowcasts of incomplete, time-stamped reporting data.*

Description

Produces nowcasts stratified by a single variable of interest, e.g. by geographic unit (province/state/region)
or by age group.

Usage

```
NobBS.strat(
  data,
  now,
  units,
  onset_date,
  report_date,
  strata,
  moving_window = NULL,
```

```
    max_D = NULL,
    cutoff_D = NULL,
    quiet = TRUE,
    proportion_reported = 1,
   specs = list(dist = c("Poisson", "NB"), alpha1.mean.prior = 0, alpha1.prec.prior =
     0.001, alphat.shape.prior = 0.001, alphat.rate.prior = 0.001, beta.priors = NULL,
    param_names = NULL, conf = 0.95, dispersion.prior = NULL, nAdapt = 1000, nChains = 1,
       nBurnin = 1000, nThin = 1, nSamp = 10000)
)
```

Arguments

data	A time series of reporting data in line list format (one row per case), with a column onset_date indicating date of case onset, and a column report_date indicating date of case report.
now	An object of datatype Date indicating the date at which to perform the nowcast.
units	Time scale of reporting. Options: "1 day", "1 week".
onset_date	In quotations, the name of the column of datatype Date designating the date of case onset. e.g. "onset_week"
report_date	In quotations, the name of the column of datatype Date designating the date of case report. e.g. "report_week"
strata	In quotations, the name of the column indicating the stratifying variable.
moving_window	Size of moving window for estimation of cases (numeric). The moving window size should be specified in the same date units as the reporting data (i.e. specify 7 to indicate 7 days, 7 weeks, etc). Default: NULL, i.e. takes all historical dates into consideration.
max_D	Maximum possible delay observed or considered for estimation of the delay distribution (numeric). Default: (length of unique dates in time series)-1 ; or, if a moving window is specified, (size of moving window)-1
cutoff_D	Consider only delays d<=max_D? Default: TRUE. If cutoff_D=TRUE, delays beyond max_D are ignored. If cutoff_D=FALSE, max_D is interpreted as delays>=max_D but within the moving window given by moving_window.
quiet	Suppress all output and progress bars from the JAGS process. Default: TRUE.
proportion_reported	
	A decimal greater than 0 and less than or equal to 1 representing the proportion of all cases expected to be reported. Default: 1, e.g. 100 percent of all cases will eventually be reported. For asymptomatic diseases where not all cases will ever be reported, or for outbreaks in which severe under-reporting is expected, change this to less than 1.
specs	A list with arguments specifying the Bayesian model used: dist (Default: "Poisson"), beta.priors (Default: 0.1 for each delay d), nSamp (Default: 10000), nBurnin (Default: 1000), nAdapt (Default: 1000), nChains (Default: 1), nThin (Default: 1), alphat.shape.prior (Default: 0.001), alphat.rate.prior (Default: 0.001), alpha1.mean.prior (Default: 0), alpha1.prec.prior (Default: 0.001), dispersion.prior (Default: NULL, i.e. no dispersion. Otherwise, enter c(shape,rate) for a Gamma distribution.), conf (Default: 0.95), param_names

(Default: NULL, i.e. output for all parameters is provided: c("lambda","alpha","beta.logged","tau2.alpha'
See McGough et al. 2019 (https://www.biorxiv.org/content/10.1101/663823v1)
for detailed explanation of these parameters.).

Value

The function returns a list with the following elements: `estimates`, a 5-column data frame contain-
ing estimates for each date in the window of predictions (up to "now") with corresponding date of
case onset, lower and upper bounds of the prediction interval, and the number of cases for that onset
date reported up to 'now'; `estimates.inflated`, a Tx4 data frame containing estimates inflated
by the proportion_reported for each date in the time series (up to "now") with corresponding date of
case onset, lower and upper bounds of the prediction interval, and the number of cases for that onset
date reported up to 'now'; `nowcast.post.samples`, vector of 10,000 samples from the posterior
predictive distribution of the nowcast, and `params.post`, a 10,000xN dataframe containing 10,000
posterior samples for the "N" parameters specified in specs[["param_names"]]. See McGough et al.
2019 (https://www.biorxiv.org/content/10.1101/663823v1) for detailed explanation of parameters.

Notes

'NobBS' requires that JAGS (Just Another Gibbs Sampler) is downloaded to the system. JAGS can
be downloaded at <http://mcmc-jags.sourceforge.net/>.

Examples

```
# Load the data
data(denguedat)
# Perform stratified 'NobBS' assuming Poisson distribution, vague priors, and default
# specifications.
nowcast <- NobBS.strat(denguedat, as.Date("1990-02-05"),units="1 week",onset_date="onset_week",
report_date="report_week",strata="gender")
nowcast$estimates
```

Index