

Package ‘rstansim’

September 21, 2017

Type Package

Title Simulation Studies with Stan

Version 0.1.1

Date 2017-09-16

Description Provides a set of functions to facilitate and ease the running of simulation studies of Bayesian models using 'stan'. Provides functionality to simulate data, fit models, and manage simulation results.

License GPL (>= 3) | file LICENSE

Encoding UTF-8

LazyData true

Imports loo (>= 1.0.0), foreach (>= 1.4.3), doSNOW (>= 1.0.14), doRNG (>= 1.6.6), rstan (>= 2.14.1)

Depends R (>= 3.2.3), Rcpp (>= 0.12.12)

RoxygenNote 6.0.1

Suggests testthat, knitr, rmarkdown, ggplot2, ggjoy

VignetteBuilder knitr

URL <https://github.com/Ewan-Keith/rstansim#rstansim>

BugReports <https://github.com/Ewan-Keith/rstansim/issues>

NeedsCompilation no

Author Ewan Keith [aut, cre]

Maintainer Ewan Keith <ewan.keith100@gmail.com>

Repository CRAN

Date/Publication 2017-09-21 17:06:14 UTC

R topics documented:

collect_simulations	2
extract_data	3
extract_data.stansim_collection	4
extract_data.stansim_simulation	5
extract_refitted	7
extract_refitted.stansim_collection	7
extract_refitted.stansim_simulation	8
extract_time_elapsed	9
extract_time_elapsed.stansim_collection	10
extract_time_elapsed.stansim_simulation	11
fit_models	12
print.stansim_collection	14
print.stansim_data	15
print.stansim_simulation	16
refit	16
rename	18
rename.stansim_collection	18
rename.stansim_simulation	19
rstansim	20
simulate_data	20
stansim_collection	21
stansim_data	22
stansim_simulation	23
Index	24

collect_simulations	<i>Group stansim objects into a collection</i>
---------------------	--

Description

collect_simulations() groups together an arbitrary number of objects with class stansim_simulation or stansim_collection into a single stansim_collection object. Allows for multiple simulations to be stored, saved, analysed and managed in a single object.

Usage

```
collect_simulations(collection_name, object, ...)
```

Arguments

collection_name	A name attached to the stansim_collection object to help identify it. It is strongly recommended that an informative and unique name is assigned.
object	An object of class stansim_simulation or stansim_collection. Must be provided.

... Any further `stansim_simulation` or `stansim_collection` objects to be grouped into a single `stansim_collection` object.

Value

An S3 object of class `stansim_simulation` recording relevant simulation data.

Examples

```
## Not run:
# group together stansim_simulation objects
collection_basic <- collect_simulations("Linear Regression Study", simulation1,
                                       simulation2)

# group together stansim_simulations and stansim_collections
collection_extended <- collect_simulations("Extended Lin Reg Study", collection_basic,
                                           simulation3)

# group together multiple stansim_collections
merged_collections <- collect_simulations("merged collections", collection_extended,
                                          collection_additional)

## End(Not run)
```

extract_data	<i>Extract data from rstansim objects</i>
--------------	---

Description

Generic function for extracting data from `rstansim` objects. Default arguments will return full data as a dataframe, otherwise rows will be filtered based on provided arguments.

Usage

```
extract_data(object, ...)
```

Arguments

`object` An S3 object of class `stansim_simulation` or `stansim_collection`.
... Arguments for filtering returned data, see specific methods for further detail.

Value

A dataframe containing the specified data.

```
extract_data.stansim_collection
```

Extract data from a stansim_collection object

Description

Applied to an object of type `stansim_collection`, `extract_data()` will return the object's simulation data as a dataframe, subject to the filtering specified by the function arguments.

Usage

```
## S3 method for class 'stansim_collection'
extract_data(object, sim_names = "all",
             datasets = "all", parameters = "all", estimates = "all",
             values = NULL, param_expand = TRUE, ...)
```

Arguments

<code>object</code>	An object of S3 class <code>stansim_collection</code> .
<code>sim_names</code>	Either a character vector containing the names of the <code>stansim_simulation</code> objects grouped in the collection, or the string "all". The former will only return values for the corresponding simulations, the latter applies no filtering on stan-sim simulations.
<code>datasets</code>	Either a character vector containing the names of datasets (as provided to the original <code>stansim()</code> call) fitted, or the string "all". The former will only return values for the corresponding datasets, the latter applies no filtering on datasets
<code>parameters</code>	Either a character vector containing the names of stan model parameters present in the fitted stan models, or the string "all". The former will only return values for the corresponding parameters, the latter applies no filtering on parameters. See also the effect of the <code>param_expand</code> argument.
<code>estimates</code>	Either a character vector containing the names of parameter estimates calculated (e.g. <code>c("2.5" "all")</code>). The former will only return values for the corresponding estimates, the latter applies no filtering on estimates
<code>values</code>	Either a function taking a single numeric argument that returns a Boolean value, or NULL. The former will only return values for which the provided function is TRUE, the latter applies no filtering on values.
<code>param_expand</code>	If TRUE then any provided <code>parameters</code> arguments, without specified dimension, will be expanded to capture all dimensions of that parameter. For example, "eta" becomes <code>c("eta[1]", "eta[2]", "eta[3]", ...)</code> . Expansion isn't carried out if a <code>parameters</code> dimension is specified (e.g. <code>parameters = "eta[1]"</code>) or if <code>param_expand = FALSE</code> .
<code>...</code>	other arguments not used by this method

Value

A dataframe containing the specified data.

Examples

```
## Not run:
# extract full dataset
extract_data(collection)

# extract all parameter means, 2.5% & 97.5% percentiles
extract_data(collection, estimates = c("2.5%", "mean", "97.5%"))

# extract all Rhat estimates greater than 1.1
extract_data(collection, estimates = "Rhat",
             values = function(x) x > 1.1)

# extract all "eta" parameters
extract_data(collection, parameters = "eta")

# extract all "eta[1]" parameters
extract_data(collection, parameters = "eta[1]",
             param_expand = FALSE)

# extract all rows for dataset "data_file-12.rds"
extract_data(collection, datasets = "data_file-12.rds")
# extract all rows for sim_names "simulation1"
extract_data(collection, sim_names = "simulation1")

## End(Not run)
```

extract_data.stansim_simulation

Extract data from a stansim_simulation object

Description

Applied to an object of type `stansim_simulation`, `extract_data()` will return the object's simulation data as a dataframe, subject to the filtering specified by the function arguments.

Usage

```
## S3 method for class 'stansim_simulation'
extract_data(object, datasets = "all",
             parameters = "all", estimates = "all", values = NULL,
             param_expand = TRUE, ...)
```

Arguments

`object` An object of S3 class `stansim_simulation`.

datasets	Either a character vector containing the names of datasets (as provided to the original <code>fit_models()</code> call) fitted, or the string "all". The former will only return values for the corresponding datasets, the latter applies no filtering on datasets
parameters	Either a character vector containing the names of stan model parameters present in the fitted stan models, or the string "all". The former will only return values for the corresponding parameters, the latter applies no filtering on parameters. See also the effect of the <code>param_expand</code> argument.
estimates	Either a character vector containing the names of parameter estimates calculated (e.g. <code>c("2.5" "all")</code>). The former will only return values for the corresponding estimates, the latter applies no filtering on estimates
values	Either a function taking a single numeric argument that returns a Boolean value, or NULL. The former will only return values for which the provided function is TRUE, the latter applies no filtering on values.
param_expand	If TRUE then any provided <code>parameters</code> arguments, without specified dimension, will be expanded to capture all dimensions of that parameter. For example, "eta" becomes <code>c("eta[1]", "eta[2]", "eta[3]", ...)</code> . Expansion isn't carried out if a <code>parameters</code> dimension is specified (e.g. <code>parameters = "eta[1]"</code>) or if <code>param_expand = FALSE</code> .
...	other arguments not used by this method

Value

A dataframe containing the specified data.

Examples

```
## Not run:
# extract full dataset
extract_data(simulation)

# extract all parameter means, 2.5% & 97.5% percentiles
extract_data(simulation, estimates = c("2.5%", "mean", "97.5%"))

# extract all Rhat estimates greater than 1.1
extract_data(simulation, estimates = "Rhat",
             values = function(x) x > 1.1)

# extract all "eta" parameters
extract_data(simulation, parameters = "eta")

# extract all "eta[1]" parameters
extract_data(simulation, parameters = "eta[1]",
             param_expand = FALSE)

# extract all rows for dataset "data_file-12.rds"
extract_data(simulation, datasets = "data_file-12.rds")

## End(Not run)
```

extract_refitted	<i>Extract identifiers of refitted items from rstansim objects</i>
------------------	--

Description

Generic function for extracting refitted identifiers from rstansim objects. Default arguments will return full data as a dataframe, otherwise rows will be filtered based on provided arguments (stansim_collections only).

Usage

```
extract_refitted(object, ...)
```

Arguments

object	An S3 object of class stansim_simulation or stansim_collection.
...	Arguments for filtering returned refitted records, see specific methods for further detail.

Value

The specified refitted identifiers, see specific methods for further detail.

extract_refitted.stansim_collection	<i>Extract details of refitted datasets from a stansim_collection object</i>
-------------------------------------	--

Description

Applied to an object of type stansim_collection, extract_refitted() will return a dataframe of simulation-dataset pairings that have been refitted since the initial simulation.

Usage

```
## S3 method for class 'stansim_collection'
extract_refitted(object, sim_names = "all",
  datasets = "all", ...)
```

Arguments

object	An object of S3 class stansim_collection.
sim_names	Either a character vector containing the names of the stansim_simulation objects grouped in the collection, or the string "all". The former will only return values for the corresponding simulations, the latter applies no filtering on stansim simulations.
datasets	Either a character vector containing the names of datasets (as provided to the original stansim call) fitted, or the string "all". The former will only return values for the corresponding datasets, the latter applies no filtering on stansim datasets.
...	other arguments not used by this method.

Value

A dataframe with the simulation titles and dataset names of refitted datasets.

Examples

```
## Not run:
# extract all refitted indicators
extract_refitted(collection)

# extract only datasets from "simulation1"
extract_refitted(collection, sim_names = "simulation1")

# extract only indicators for dataset "data-file_12.rds"
extract_refitted(collection, datasets = "data-file_12.rds")

## End(Not run)
```

```
extract_refitted.stansim_simulation
```

Extract names of refitted datasets from a stansim_simulation object

Description

Applied to an object of type stansim_simulation, extract_refitted() will return a vector of the names of datasets that have been refitted since the initial simulation.

Usage

```
## S3 method for class 'stansim_simulation'
extract_refitted(object, ...)
```


Arguments

object An object of S3 class stansim_simulation.
... other arguments not used by this method.

Value

A character vector of the refitted dataset names.

Examples

```
## Not run:  
# extract names of refitted datasets  
extract_refitted(simulation)  
  
## End(Not run)
```

extract_time_elapsed *Extract time_elapsed from rstansim objects*

Description

Generic function for extracting the time taken to fit rstansim models Default arguments will return full data as a dataframe, otherwise rows will be filtered based on provided arguments.

Usage

```
extract_time_elapsed(object, ...)
```

Arguments

object An S3 object of class stansim_simulation or stansim_collection.
... Arguments for filtering returned data, see specific methods for further detail.

Value

A dataframe containing the specified data.

```
extract_time_elapsed.stansim_collection
```

Extract time_elapsed from a stansim_collection object

Description

Applied to an object of type `stansim_collection`, `extract_time_elapsed()` will return the time taken to fit the models contained within the object a dataframe, subject to the filtering specified by the function arguments.

Usage

```
## S3 method for class 'stansim_collection'
extract_time_elapsed(object, sim_names = "all",
  datasets = "all", chains = "all", stages = "all", elapsed = NULL, ...)
```

Arguments

<code>object</code>	An object of S3 class <code>stansim_collection</code> .
<code>sim_names</code>	Either a character vector containing the names of the <code>stansim_simulation</code> objects grouped in the collection, or the string "all". The former will only return values for the corresponding simulations, the latter applies no filtering on stan-sim simulations.
<code>datasets</code>	Either a character vector containing the names of datasets (as provided to the original <code>fit_models()</code> calls) fitted, or the string "all". The former will only return values for the corresponding datasets, the latter applies no filtering on datasets
<code>chains</code>	Either a character vector containing the numbers of the stan model chains to return, or the string "all". The former will only return values for the corresponding chains, the latter applies no filtering on chains.
<code>stages</code>	Either a character vector containing the names of model fitting stages, <code>c("warmup", "sample", "total")</code> or the string "all". The former will only return values for the corresponding stages, the latter applies no filtering on estimates.
<code>elapsed</code>	Either a function taking a single numeric argument that returns a Boolean value, or NULL. The former will only return elapsed times for which the provided function is TRUE, the latter applies no filtering on elapsed times
<code>...</code>	other arguments not used by this method

Value

A dataframe containing the specified data.

Examples

```
## Not run:
# extract full dataset
extract_time_elapsed(simulation)

# extract all rows for dataset "data_file-12.rds"
extract_time_elapsed(simulation, datasets = "data_file-12.rds")

# extract results for chains 1 and 3
extract_time_elapsed(simulation, chains = c(1, 3))

# extract results for only the warmup stage
extract_time_elapsed(simulation, stages = "warmup")

# extract all elapsed times greater than 60 seconds
extract_time_elapsed(simulation,
                    elapsed = function(x) x > 60)

## End(Not run)
```

```
extract_time_elapsed.stansim_simulation
```

Extract time_elapsed from a stansim_simulation object

Description

Applied to an object of type `stansim_simulation`, `extract_time_elapsed()` will return the time taken to fit the models contained within the object a dataframe, subject to the filtering specified by the function arguments.

Usage

```
## S3 method for class 'stansim_simulation'
extract_time_elapsed(object, datasets = "all",
                    chains = "all", stages = "all", elapsed = NULL, ...)
```

Arguments

<code>object</code>	An object of S3 class <code>stansim_simulation</code> .
<code>datasets</code>	Either a character vector containing the names of datasets (as provided to the original <code>fit_models()</code> call) fitted, or the string "all". The former will only return values for the corresponding datasets, the latter applies no filtering on datasets
<code>chains</code>	Either a character vector containing the numbers of the stan model chains to return, or the string "all" The former will only return values for the corresponding chains, the latter applies no filtering on chains.

stages	Either a character vector containing the names of model fitting stages, <code>c("warmup", "sample", "total")</code> or the string "all". The former will only return values for the corresponding stages, the latter applies no filtering on estimates.
elapsed	Either a function taking a single numeric argument that returns a Boolean value, or NULL. The former will only return elapsed times for which the provided function is TRUE, the latter applies no filtering on elapsed times
...	other arguments not used by this method

Value

A dataframe containing the specified data.

Examples

```
## Not run:
# extract full dataset
extract_time_elapsed(simulation)

# extract all rows for dataset "data_file-12.rds"
extract_time_elapsed(simulation, datasets = "data_file-12.rds")

# extract results for chains 1 and 3
extract_time_elapsed(simulation, chains = c(1, 3))

# extract results for only the warmup stage
extract_time_elapsed(simulation, stages = "warmup")

# extract all elapsed times greater than 60 seconds
extract_time_elapsed(simulation,
                    elapsed = function(x) x > 60)

## End(Not run)
```

fit_models

Fit a stan model to multiple datasets

Description

`fit_models()` fits a stan model across multiple datasets, collates, and returns summary information and data for all fitted models as a `stansim_simulation` object. All fitted models have basic reproducibility information recorded; such as parameter inits and seeds, along with parameter estimates, and simulation information such as time and date ran.

Raw stan posterior samples are not returned, rather the user specifies the estimates they wish to record (e.g. posterior percentiles, Rhat, etc.) and the parameters for which they wish to record these estimates. All data is collated into a single, **tidy** dataframe for further analysis.

By default the function caches completed runs as it progresses, so that progress is not lost in the case of function failure. By simply running the function with the same calls in the same working

directory it will pick up where it left off. When the function terminates as expected this cache is removed.

Usage

```
fit_models(sim_name = paste0("Stansim_", Sys.time()), sim_data = NULL,
  stan_args = list(), calc_loo = FALSE, use_cores = 1L,
  parameters = "all", probs = c(0.025, 0.25, 0.5, 0.75, 0.975),
  estimates = c("mean", "se_mean", "sd", "n_eff", "Rhat"),
  stan_warnings = "catch", cache = TRUE, seed = floor(stats::runif(1, 1,
  1e+05)))
```

Arguments

sim_name	A name attached to the <code>stansim_simulation</code> object to help identify it. It is strongly recommended that an informative name is assigned, especially if <code>stansim_simulation</code> objects are to be combined in to a <code>stansim_collection</code> object for management of results.
sim_data	Either an object of class <code>stansim_data</code> or a vector of strings pointing to the location of <code>.rds</code> files containing the simulation data. See the vignette on producing simulation data for details on the formatting of these datasets.
stan_args	A list of function arguments to be used by the internal <code>rstan::sampling()</code> function when fitting the models. If not specified then the <code>rstan::sampling()</code> function defaults are used.
calc_loo	If TRUE then model fit statistics will be calculated using the <code>loo</code> package. If TRUE there must be a valid <code>log_lik</code> quantity specified in the generated quantities section of the provided stan model.
use_cores	Number of cores to use when running in parallel. Each stan model is fitted serially regardless of the number of chains ran, as parallelisation across models is more flexible than within.
parameters	A character vector indicating which parameters should have estimates returned and stored from the fitted models. By default all parameters are returned, for non-scalar parameters you cannot select subsets of the parameter (e.g. must request <code>theta</code> rather than <code>theta[1]</code>).
probs	A numeric vector of values between 0 and 1. Corresponding quantiles will be estimated and returned for all fitted models.
estimates	A character vector of non-quantile estimates to be returned for each model parameter. Argument must be some subset of the default character vector.
stan_warnings	How warnings returned by individual stan instances should be handled. "catch" records all warnings in the returned object alongside other instance level data, "print" simply prints warnings to the console as the models are fit (default stan behaviour), and "suppress" suppresses all warnings without recording them.
cache	If TRUE then the results for each instance are written to a local, temporary file so that data is not lost should the function not terminate properly. This temporary data is removed upon the model terminating as expected. If FALSE no data is

written and results are only returned upon the correct termination of the whole function. The default value of TRUE is recommended unless there are relevant write-permission restrictions.

seed Set a seed for the function.

Value

An S3 object of class `stansim_simulation` recording relevant simulation data.

Examples

```
## Not run:
# specify arguments for stan
StanArgs <- list(file = '8schools.stan',
                 iter = 1000, chains = 4)

# get number of cores
core_num <- parallel::detectCores()

# get the list of data file locations
datasets <- dir("data/repo", full.names = TRUE)

# fit the model to all datasets using specified stan arguments
# store the specified estimates for all parameters
simulation <- fit_models(
  sim_name = "stansim simulation",
  sim_data = datasets,
  stan_args = StanArgs,
  calc_loo = T,
  use_cores = core_num,
  probs = c(.025, .5, .975),
  estimates = c("mean", "n_eff", "Rhat")
)

## End(Not run)
```

```
print.stansim_collection
```

Print a summary for a stansim_collection object

Description

Print basic information regarding a `stansim_collection` object, including collection title, the simulations within the collection, and any datasets that have been refitted.

Usage

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

Arguments

x An object of S3 class stansim_collection.
... other arguments not used by this method

See Also

S3 class [stansim_collection](#).

Examples

```
## Not run:  
# print stansim_collection summary info  
print(collection)  
  
## End(Not run)
```

```
print.stansim_data     Print a summary for a stansim_data object
```

Description

Print basic information regarding a stansim_data object, including data title, number of datasets, variables recorded, and the name of the model from which the data was simulated from.

Usage

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

Arguments

x An object of S3 class stansim_data.
... other arguments not used by this method

See Also

S3 class [stansim_data](#).

Examples

```
## Not run:  
# print stansim_data summary info  
print(simulated_data)  
  
## End(Not run)
```

```
print.stansim_simulation
```

Print a summary for a stansim_simulation object

Description

Print basic information regarding a stansim_simulation object, including simulation title, time/date ran, number of models fitted, parameters and estimates recorded, and the titles of any datasets that were refitted.

Usage

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

Arguments

x An object of S3 class stansim_simulation.
 ... other arguments not used by this method

See Also

S3 class [stansim_simulation](#).

Examples

```
## Not run:
# print stansim_simulation summary info
print(simulation)

## End(Not run)
```

```
refit
```

Refit specified datasets in a stansim_simulation object

Description

refit Takes a stansim_simulation object and a vector of characters corresponding to the names of datasets fitted within the stansim_simulation object, and refits the stan model for each of these instances. This allows users to refit any specific models using new stan arguments if need be (e.g. if the model fails to converge in the original run).

Usage

```
refit(object, datasets, stan_args = list(), calc_loo = FALSE,
       use_cores = 1L, cache = TRUE, seed = floor(stats::runif(1, 1, 1e+05)))
```


Arguments

object	An object of S3 class <code>stansim_simulation</code> .
datasets	The full names of the data files to be refitted. These must be consistent both with the dataset names stored within the <code>stansim_simulation</code> object, and with the copies of the data files relative to the current working directory. This is best ensured by running <code>refit</code> from the same working directory as the original <code>fit_models()</code> call.
stan_args	A list of function arguments to be used by the internal <code>rstan::sampling()</code> function when fitting the models. If not specified then the <code>sampling()</code> defaults are used.
calc_loo	If TRUE then model fit statistics will be calculated using the <code>loo</code> package. If TRUE there must be a valid <code>log_lik</code> quantity specified in the generated quantities section of the provided stan model.
use_cores	Number of cores to use when running in parallel. Each stan model is fitted serially regardless of the number of chains ran as parallelisation across models is more flexible than within.
cache	If TRUE then the results for each instance are written to a local, temporary file so that data is not lost should the function not terminate properly. This temporary data is removed upon the model terminating as expected. If FALSE no data is written and results are only returned upon the correct termination of the whole function. The default value of TRUE is recommended unless there are relevant write-permission restrictions.
seed	Set a seed for the <code>fit_models()</code> function.

Value

An S3 object of class `stansim_simulation` recording relevant simulation data.

Examples

```
## Not run:
# refit datasets "data_file-12.rds" & "data_file-08.rds"
refit(simulation,
      datasets = c("data_file-12.rds", "data_file-08.rds")
      use_cores = 4)

# refit dataset "data_file-12.rds" using a larger number of samples
refit(simulation,
      datasets = "data_file-12.rds",
      stan_args = list(iter = 4000),
      use_cores = 4)

## End(Not run)
```

rename	<i>Rename a stansim object</i>
--------	--------------------------------

Description

Generic function for renaming rstansim objects.

Usage

```
rename(object, new_name)
```

Arguments

object	An S3 object of class stansim_simulation or stansim_collection.
new_name	New object name.

Value

A stansim S3 object.

rename.stansim_collection	<i>Rename a stansim_collection object</i>
---------------------------	---

Description

Change the collection_name value of a `"stansim_collection"` object.

Usage

```
## S3 method for class 'stansim_collection'
rename(object, new_name)
```

Arguments

object	An object of S3 class stansim_collection.
new_name	New object name. Must be of type character.

Value

An S3 object of class stansim_collection.

Examples

```
## Not run:  
# rename a stansim_collection to "new collection name"  
rename(collection, "new collection name")  
  
## End(Not run)
```

rename.stansim_simulation

Rename a stansim_simulation object

Description

Change the sim_name of a "stansim_simulation" object.

Usage

```
## S3 method for class 'stansim_simulation'  
rename(object, new_name)
```

Arguments

object	An object of S3 class stansim_simulation.
new_name	New object name.

Value

An S3 object of class stansim_simulation.

Examples

```
## Not run:  
# rename a stansim_simulation to "new simulation name"  
rename(simulation, "new simulation name")  
  
## End(Not run)
```

`rstansim`*rstansim: running simulation studies with Stan*

Description

`rstansim` provides a set of helper functions and utilities to make it easier to setup, run, analyse, and manage simulation studies in R that use Stan for estimating Bayesian models.

Details

The package addresses three aspects of running a simulation study:

- Data simulation [currently not covered].
- Model fitting and capture of relevant data.
- Management of simulation results.

To learn more about `rstansim`, start with the vignettes: `'browseVignettes(package = "rstansim")'`

`simulate_data`*Simulate datasets from a stan model*

Description

`simulate_data()` takes a specified stan model and allows the user to simulate data from it based on specified parameter values. The user then specifies which data they wish to save and how many simulations they wish to run. The data will be saved as individual `.rds` files in the directory specified by `path`.

By default an object of class `stansim_data` will be returned, providing an index of the saved data that can then be provided directly to a `stansim()` call.

To allow for simulated data to be directly fed into stan model that simulated them as input data, the `sim_drop` argument is provided. If `sim_drop` is true then any stan data object with a name beginning with `"sim_"` will have this string removed from it's name. For example, the simulated data `"sim_x"` would be returned simply as `"x"`. This helps avoid the issue of overlapping data names for both input and output

Usage

```
simulate_data(file, data_name = paste0("Simdata_", Sys.time()),
  input_data = NULL, vars = "all", param_values = NULL, nsim = 1,
  path = NULL, seed = floor(stats::runif(1, 1, 1e+05)),
  return_object = TRUE, use_cores = 1, sim_drop = TRUE,
  recursive = TRUE)
```

Arguments

file	A character string containing either the file location of the model code (ending in ".stan"), a character string containing the model specification or the name of a character string object in the workspace.
data_name	A name attached to the <code>stansim_data</code> object to help identify it. It is strongly recommended that an informative name is assigned. This will also be the name stem for the saved <code>.rds</code> files.
input_data	Values for the data field in the provided stan model. Values must be provided for all entries even if they are not used in the 'generate quantities' model section producing the simulated data.
vars	The names of the stan variables to return. Defaults to "all", otherwise a vector of variable names should be provided.
param_values	A list containing the named values for the stan model parameters used to simulate data. If a parameter's value is not specified here it will be initialised randomly. Recommended to specify all parameter values.
nsim	The number of simulated datasets to produce.
path	The name of the directory to save the simulated data to, if this doesn't exist it will be created. Defaults to NULL in which the datasets are saved to the working directory
seed	Set a seed for the function.
return_object	if FALSE then no <code>stansim_data</code> object is returned.
use_cores	Number of cores to use when running in parallel.
sim_drop	If TRUE then any simulated data objects beginning in "sim_" will have this removed. So "sim_x" becomes "x".
recursive	logical. Should elements of the path other than the last be created? If true, like the Unix command <code>mkdir -p</code> .

Value

An object of S3 class `stansim_data` or NULL.

<code>stansim_collection</code>	<i>Construct an S3 object of type <code>stansim_collection</code></i>
---------------------------------	---

Description

A constructor function for creating S3 objects of class `stansim_collection`. `stansim_collection` objects are the preferred means of managing and storing the results of multiple stan simulations ran using `stansim()`.

Usage

```
stansim_collection(collection_name, data, refitted, simulations)
```

Arguments

collection_name	A name attached to the stansim_collection object to help identify it. It is strongly recommended that an informative name is assigned.
data	The tidy dataframe containing data from all grouped simulations.
refitted	A dataframe containing a row for every refitted dataset, listing the simulation to which it belonged and the dataset name.
simulations	A list of all simulation data other than estimated values and the list of refitted data.

Value

An S3 object of class stansim_collection.

stansim_data	<i>Construct an S3 object of type stansim_data</i>
--------------	--

Description

A constructor function for creating S3 objects of class stansim_data. stansim_data objects are the in-memory option for storing the results of a call to stansim_simulate() and feeding the results to stansim().

Usage

```
stansim_data(data_name, datasets, compiled_model, input_data, param_values,
             vars)
```

Arguments

data_name	A name attached to the stansim_data object to help identify it. This also forms the stem of the individual .rds file names after a call to write_data(). It is strongly recommended that an informative name is assigned.
datasets	A vector of names of simulated datasets.
compiled_model	An object of S4 class stanmodel, this should be the model provided to stansim_simulate() to simulate data from.
input_data	Values for the data field in the provided stan model.
param_values	A list containing the named values for the stan model parameters used to simulate data.
vars	The names of the stan variables saved.

Value

An S3 object of class stansim_data.

stansim_simulation	<i>Construct an S3 object of type stansim_simulation</i>
--------------------	--

Description

A constructor function for creating S3 objects of class `stansim_simulation`. `stansim_simulation` objects are the basic unit of output from calls to the `fit_models()` function and collects the specified data for all stan models fitted.

Usage

```
stansim_simulation(sim_name, stansim_uni_list, start_time, end_time, raw_call,
  seed)
```

Arguments

<code>sim_name</code>	The name to be given to the simulation represented by the <code>stansim_simulation</code> object.
<code>stansim_uni_list</code>	A list of objects with S3 class <code>stan_sim_uni</code> . This is an unexported class used to store the outcomes of individual simulation runs internal to the <code>fit_models()</code> function.
<code>start_time</code>	System time when <code>fit_models()</code> was called.
<code>end_time</code>	System time when the results from <code>fit_models()</code> were returned.
<code>raw_call</code>	The values of all arguments provided to <code>fit_models()</code> when first ran. This is used for any refitting of datasets using the <code>refit()</code> method.
<code>seed</code>	The global seed for the <code>fit_models()</code> call.

Value

An S3 object of class `stansim_simulation` recording relevant simulation data.

Index

`collect_simulations`, [2](#)

`extract_data`, [3](#)
`extract_data.stansim_collection`, [4](#)
`extract_data.stansim_simulation`, [5](#)
`extract_refitted`, [7](#)
`extract_refitted.stansim_collection`, [7](#)
`extract_refitted.stansim_simulation`, [8](#)
`extract_time_elapsed`, [9](#)
`extract_time_elapsed.stansim_collection`,
[10](#)
`extract_time_elapsed.stansim_simulation`,
[11](#)

`fit_models`, [12](#)

`print.stansim_collection`, [14](#)
`print.stansim_data`, [15](#)
`print.stansim_simulation`, [16](#)

`refit`, [16](#)
`rename`, [18](#)
`rename.stansim_collection`, [18](#)
`rename.stansim_simulation`, [19](#)
`rstansim`, [20](#)
`rstansim-package (rstansim)`, [20](#)

`simulate_data`, [20](#)
`stansim_collection`, [15](#), [21](#)
`stansim_data`, [15](#), [22](#)
`stansim_simulation`, [16](#), [23](#)