Package: rakeR (via r-universe)

August 23, 2024

Title Easy Spatial Microsimulation (Raking) in R

Version 0.2.1.9000

Date 2018-02-26

Description Functions for performing spatial microsimulation ('raking') in R.

Depends R (>= 3.4.0)

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

Imports ipfp, wrswoR

Suggests testthat, readr

URL https://philmikejones.github.io/rakeR/

BugReports https://github.com/philmikejones/rakeR/issues

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

RemoteUrl https://github.com/philmikejones/raker

RemoteRef HEAD

RemoteSha 676dbb0c36e7cf6beeac6f3fb8d350d1758123cc

Contents

extract *extract*

Description

Deprecated. Use rk_extract() instead.

Usage

```
extract(weights, inds, id)
```

Arguments

weights	A weight table, typically produced by rakeR::weight()
inds	The individual level data
id	The unique id variable in the individual level data (inds), usually the first column

Value

A data frame with zones and aggregated simulated values for each variable

Examples

```
## Not run:
Deprecated. Use rk_extract()

## End(Not run)
```

extract_weights *extract_weights*

Description

Deprecated: use rakeR::rk_extract()

Usage

```
extract_weights(weights, inds, id)
```

Arguments

weights	A weight table, typically produced using rakeR::rk_weight()
inds	The individual level data
id	The unique id variable in the individual level data (inds), usually the first column

Value

A data frame with zones and aggregated simulated values for each variable

Examples

```
## Not run:
extract_weights() is deprecated, use rk_extract() instead

## End(Not run)
```

`integerise` *integerise*

Description

Deprecated. Use rk_integerise()

Usage

```
integerise(weights, inds, method = "trs", seed = 42)
```

Arguments

`weights`	A matrix or data frame of fractional weights, typically provided by `rakeR::weight()`
`inds`	The individual–level data (i.e. one row per individual)
`method`	The integerisation method specified as a character string. Defaults to `"trs"`; currently other methods are not implemented.
`seed`	The seed to use, defaults to 42.

Value

A data frame of integerised cases

Examples

```
## Not run:
Deprecated. Use rk_integerise()

## End(Not run)
```

rake *rake*

Description

Deprecated. Use rk_rake()

Usage

```
rake(cons, inds, vars, output = "fraction", iterations = 10, ...)
```

Arguments

cons	A data frame of constraint variables
inds	A data frame of individual–level (survey) data
vars	A character string of variables to iterate over
output	A string specifying the desired output, either "fraction" (rk_extract()) or "integer" (rk_integerise())
iterations	The number of iterations to perform. Defaults to 10.
...	Additional arguments to pass to depending on desired output: • if "fraction" specify 'id' (see rk_extract() documentation) • if "integer" specify 'method' and 'seed' (see rk_integerise() documentation)

Value

A data frame with extracted weights (if output == "fraction", the default) or integerised cases (if output == "integer")

Examples

```
## Not run:
Deprecated. Use rk_rake()

## End(Not run)
```

rk_extract *rk_extract*

Description

Extract aggregate weights from individual weight table

Usage

```
rk_extract(weights, inds, id)
```

Arguments

weights	A weight table, typically produced by rakeR::weight()
inds	The individual level data
id	The unique id variable in the individual level data (inds), usually the first column

Details

Extract aggregate weights from individual weight table, typically produced by rakeR::rk_weight()

Extract cannot operate with numeric variables because it creates a new variable for each unique factor of each variable If you want numeric information, like income, you need to cut() the numeric values, or use integerise() instead.

Value

A data frame with zones and aggregated simulated values for each variable

Examples

```
## Not run
## Use weights object from rk_weight()
## ext_weights <- rk_extract(weights = weights, inds = inds, id = "id")
```

rk_integerise *rk_integerise*

Description

Generate integer cases from numeric weights matrix.

Usage

```
rk_integerise(weights, inds, method = "trs", seed = 42)
```

Arguments

weights	A matrix or data frame of fractional weights, typically provided by rakeR::rk_weight()
inds	The individual–level data (i.e. one row per individual)
method	The integerisation method specified as a character string. Defaults to "trs"; currently other methods are not implemented.
seed	The seed to use, defaults to 42.

Details

Extracted weights (using rakeR::rk_extract()) are more 'precise' than integerised weights (although the user should be careful this is not spurious precision based on context) as they return fractions. Nevertheless, integerised weights are useful in cases when:

- Numeric information (such as income) is required, as this needs to be cut() to work with rakeR::rk_extract()
- Simulated 'individuals' are required for case studies of key areas.
- Input individual-level data for agent-based or dynamic models are required

The default integerisation method uses the 'truncate, replicate, sample' method developed by Robin Lovelace and Dimitris Ballas http://www.sciencedirect.com/science/article/pii/S0198971513000240

Other methods (for example proportional probabilities) may be implemented at a later date.

Value

A data frame of integerised cases

Examples

```
cons <- data.frame(
  "zone"      = letters[1:3],
  "age_0_49"  = c(8, 2, 7),
  "age_gt_50" = c(4, 8, 4),
  "sex_f"     = c(6, 6, 8),
  "sex_m"     = c(6, 4, 3),
  stringsAsFactors = FALSE
)

inds <- data.frame(
  "id"     = LETTERS[1:5],
  "age"    = c(
    "age_gt_50", "age_gt_50", "age_0_49", "age_gt_50", "age_0_49"
  ),
  "sex"    = c("sex_m", "sex_m", "sex_m", "sex_f", "sex_f"),
  "income" = c(2868, 2474, 2231, 3152, 2473),
  stringsAsFactors = FALSE
)
vars <- c("age", "sex")

weights     <- rk_weight(cons = cons, inds = inds, vars = vars)
weights_int <- rk_integerise(weights, inds = inds)
```

rk_rake *rk_rake*

Description

A convenience function wrapping `rk_weight()` and `rk_extract()` or `rk_weight()` and `rk_integerise()`

Usage

```
rk_rake(cons, inds, vars, output = "fraction", iterations = 10, ...)
```

Arguments

`cons`	A data frame of constraint variables
`inds`	A data frame of individual–level (survey) data
`vars`	A character string of variables to iterate over
`output`	A string specifying the desired output, either "fraction" (rk_extract()) or "integer" (rk_integerise())
`iterations`	The number of iterations to perform. Defaults to 10.
`...`	Additional arguments to pass to depending on desired output: • if "fraction" specify 'id' (see rk_extract() documentation) • if "integer" specify 'method' and 'seed' (see rk_integerise() documentation)

Value

A data frame with extracted weights (if output == "fraction", the default) or integerised cases (if output == "integer")

Examples

```
## Not run:
frac_weights <- rake(cons, inds, vars, output = "fraction",
                     id = "id")

int_weight <- rake(cons, inds, vars, output = "integer",
                   method = "trs", seed = "42")

## End(Not run)
```

rk_weight *rk_weight*

Description

Produces fractional weights using the iterative proportional fitting algorithm.

Usage

```
rk_weight(cons, inds, vars = NULL, iterations = 10)
```

Arguments

cons	A data frame containing all the constraints. This should be in the format of one row per zone, one column per constraint category. The first column should be a zone code; all other columns must be numeric counts.
inds	A data frame containing individual-level (survey) data. This should be in the format of one row per individual, one column per constraint. The first column should be an individual ID.
vars	A character vector of variables that constrain the simulation (i.e. independent variables)
iterations	The number of iterations the algorithm should complete. Defaults to 10

Details

rk_weight() requires three arguments:

- A data frame of constraints (e.g. census tables)
- A data frame of individual data (e.g. a survey)
- A character vector of constraint variable names

The first column of each data frame should be an ID. The first column of `cons` should contain the zone codes. The first column of `inds` should contain the individual unique identifier.

Both data frames should only contain:

- an ID column (zone ID `cons` or individual ID `inds`).
- constraints `inds` or constraint category `cons`.
- `inds` can optionally contain additional dependent variables that do not influence the weighting process.

No other columns should be present (the user can merge these back in later).

It is essential that the levels in each `inds` constraint (i.e. column) match exactly with the column names in `cons`. In the example below see how the column names in cons (`'age_0_49'`, `'sex_f'`, `...`) match exactly the levels in the appropriate `inds` variables.

The columns in `cons` must be arranged in alphabetical order because these are created alphabetically when they are 'spread' in the individual-level data.

Value

A data frame of fractional weights for each individual in each zone with zone codes recorded in column names and individual id recorded in row names.

Examples

```
# SimpleWorld
cons <- data.frame(
  "zone"      = letters[1:3],
  "age_0_49"  = c(8, 2, 7),
  "age_gt_50" = c(4, 8, 4),
  "sex_f"     = c(6, 6, 8),
  "sex_m"     = c(6, 4, 3),
  stringsAsFactors = FALSE
)
inds <- data.frame(
  "id"     = LETTERS[1:5],
  "age"    = c(
    "age_gt_50", "age_gt_50", "age_0_49", "age_gt_50", "age_0_49"
  ),
  "sex"    = c("sex_m", "sex_m", "sex_m", "sex_f", "sex_f"),
  "income" = c(2868, 2474, 2231, 3152, 2473),
  stringsAsFactors = FALSE
)
# Set variables to constrain over
vars <- c("age", "sex")
weights <- rk_weight(cons = cons, inds = inds, vars = vars)
print(weights)
```

`weight` *weight*

Description

Deprecated. Use rk_weight()

Usage

```
weight(cons, inds, vars = NULL, iterations = 10)
```

Arguments

`cons`	A data frame containing all the constraints. This should be in the format of one row per zone, one column per constraint category. The first column should be a zone code; all other columns must be numeric counts.
`inds`	A data frame containing individual-level (survey) data. This should be in the format of one row per individual, one column per constraint. The first column should be an individual ID.

vars	A character vector of variables that constrain the simulation (i.e. independent variables)
iterations	The number of iterations the algorithm should complete. Defaults to 10

Value

A data frame of fractional weights for each individual in each zone with zone codes recorded in column names and individual id recorded in row names.

Examples

```
# Deprecated. Use rk_weight()
```

Index