

Package ‘risq’

September 23, 2026

Title Representativity Indicators for Survey Quality

Version 3.0.1

Maintainer Reijer Idema <r.idema@cbs.nl>

Description Calculate representativity indicators for survey quality based on survey data and response models. Use partial indicators to analyse the impact of individual variables and categories. Monitor changes in representativity during data collection. Improve representativity through adaptive survey design. Supports both R-indicators and coefficients of variation. See also Schouten, Cobben, Bethlehem (2009) <<https://api.semanticscholar.org/CorpusID:33654901>>, Shlomo, Skinner, Schouten (2012) <[doi:10.1016/j.jspi.2011.07.008](https://doi.org/10.1016/j.jspi.2011.07.008)> and Schouten, Shlomo (2017) <[doi:10.1111/insr.12159](https://doi.org/10.1111/insr.12159)>.

License EUPL-1.2

URL <https://github.com/reijeridema/risq>

BugReports <https://github.com/reijeridema/risq/issues>

Encoding UTF-8

Depends R (>= 2.10)

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

LazyData true

VignetteBuilder knitr

Config/roxygen2/markdown TRUE

Config/roxygen2/version 8.0.0

Config/testthat/edition 3

NeedsCompilation no

Author Reijer Idema [aut, cre],
Barry Schouten [aut]

Repository CRAN

Date/Publication 2026-09-23 14:50:02 UTC

Contents

<code>-.bootstrap</code>	2
<code>bootstrap</code>	3
<code>cv</code>	4
<code>cv_by_cat</code>	5
<code>cv_by_var</code>	6
<code>hlc</code>	7
<code>mean.bootstrap</code>	8
<code>print.risq</code>	9
<code>quantile.bootstrap</code>	9
<code>ri</code>	10
<code>risq</code>	11
<code>ri_by_cat</code>	12
<code>ri_by_var</code>	13
<code>rr</code>	14
<code>sample</code>	15
<code>sample.risq</code>	16
<code>var</code>	16
<code>var.bootstrap</code>	17
Index	18

<code>-.bootstrap</code>	<i>Bootstrap Subtraction</i>
--------------------------	------------------------------

Description

Calculate the difference between two `bootstrap` objects. This can be used to measure the progression between two target variables.

Usage

```
## S3 method for class 'bootstrap'
x - y
```

Arguments

- `x` A `bootstrap` object.
- `y` A `bootstrap` object.

Details

For the result to be meaningful, the `bootstrap` objects `x` and `y` must have been made with the same input arguments for the `bootstrap()` function, except for the target variable.

Value

A `bootstrap` object with the difference between `x` and `y`.

See Also

Other bootstrap methods: [bootstrap\(\)](#), [mean.bootstrap\(\)](#), [quantile.bootstrap\(\)](#), [var.bootstrap\(\)](#)

Examples

```
# Note: a low iteration count is used to limit computing time of example.
risq_hlc <- risq(predictor = ~ age + gender, data = hlc)
a <- bootstrap(risq_hlc, ri, target = "response_1", seed = 0, iterations = 5)
b <- bootstrap(risq_hlc, ri, target = "response_2", seed = 0, iterations = 5)
mean(b - a)
var(b - a)
```

bootstrap	<i>RISQ Bootstrap</i>
-----------	-----------------------

Description

Draw random samples from a [risq](#) object and apply a representativity indicator or coefficient of variation function to these samples. The result is a bootstrap object from which statistical quantities can be derived using bootstrap class methods.

Usage

```
bootstrap(x, fun, ..., seed = NULL, iterations = 1000L)
```

Arguments

x	The risq object to sample from.
fun	The function to apply to the bootstrap samples. Can be any representativity indicator or coefficient of variation function.
...	Other arguments to pass into fun. Must not include include_se.
seed	An optional integer that is used to seed the random number generator for sampling. If omitted, the current seed is left intact. For details, see set.seed() .
iterations	An optional positive integer specifying the number of samples to draw and apply fun to.

Value

A bootstrap object.

See Also

risq object constructor: [risq\(\)](#)

Representativity indicator: [ri\(\)](#), [ri_by_var\(\)](#), [ri_by_cat\(\)](#)

Coefficient of variation: [cv\(\)](#), [cv_by_var\(\)](#), [cv_by_cat\(\)](#)

Other bootstrap methods: [-.bootstrap\(\)](#), [mean.bootstrap\(\)](#), [quantile.bootstrap\(\)](#), [var.bootstrap\(\)](#)

Examples

```
# Note: a low iteration count is used to limit computing time of example.
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
bs <- bootstrap(risq_hlc, ri, target = "response_1", iterations = 5)
mean(bs)
var(bs)
```

cv	<i>Coefficient of Variation</i>
----	---------------------------------

Description

Estimate the bias-adjusted coefficient of variation for a target variable, within the context of the given [risq](#) object.

Usage

```
cv(x, target, include_se = TRUE)
```

Arguments

x	A risq object.
target	Name of the target variable. Must be the name of a logical variable in the risq object data.
include_se	An optional logical specifying whether to include the standard error or not. If omitted, the standard error is included.

Value

Named list with the bias-adjusted coefficient of variation value and, if `include_se` is `TRUE`, the associated standard error `se`.

See Also

[risq](#) object constructor: [risq\(\)](#)
 Other coefficient of variation functions: [cv_by_cat\(\)](#), [cv_by_var\(\)](#)

Examples

```
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
cv(risq_hlc, "response_1")
```

cv_by_cat*Coefficient of Variation by Category*

Description

Estimate unconditional and conditional partial coefficient of variation values by category for a selection of categorical variables, given a [risq](#) object and target variable.

Usage

```
cv_by_cat(  
  x,  
  target,  
  variables,  
  type = c("unconditional", "conditional"),  
  include_se = TRUE  
)
```

Arguments

x	A risq object.
target	Name of the target variable. Must be the name of a logical variable in the risq object data.
variables	A character vector that specifies the variables for which to estimate partial coefficient of variation values. Must be the names of categorical variables in the risq object data. May include names of variables that are not part of the model.
type	An optional string that specifies the type of partial coefficient of variation to estimate. Must be either "unconditional" or "conditional". Defaults to "unconditional".
include_se	An optional logical specifying whether to include the standard error or not. If omitted, the standard error is included.

Value

A data frame with columns

- variable: name of the variable for which the row holds values,
- category: category within the variable for which the row holds values,
- value: estimate for the partial coefficient of variation,
- se: standard error for the partial coefficient of variation (only if `include_se` is `TRUE`).

Note that estimates are NA for empty categories and that conditional estimates are NA for a variable if that variable is the only variable in the predictor of the [risq](#) object model.

See Also

risq object constructor: [risq\(\)](#)
Other coefficient of variation functions: [cv\(\)](#), [cv_by_var\(\)](#)

Examples

```
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
cv_by_cat(risq_hlc, "response_1", c("gender", "age", "job"))
```

cv_by_var	<i>Coefficient of Variation by Variable</i>
-----------	---

Description

Estimate bias-adjusted unconditional or conditional partial coefficient of variation values by variable for a selection of categorical variables, given a [risq](#) object and target variable.

Usage

```
cv_by_var(  
  x,  
  target,  
  variables,  
  type = c("unconditional", "conditional"),  
  include_se = TRUE  
)
```

Arguments

x	A risq object.
target	Name of the target variable. Must be the name of a logical variable in the risq object data.
variables	A character vector that specifies the variables for which to estimate partial coefficient of variation values. Must be the names of categorical variables in the risq object data. May include names of variables that are not part of the model.
type	An optional string that specifies the type of partial coefficient of variation to estimate. Must be either "unconditional" or "conditional". Defaults to "unconditional".
include_se	An optional logical specifying whether to include the standard error or not. If omitted, the standard error is included.

Value

A data frame with columns

- variable: name of the variable for which the row holds values,
- value: estimate for the partial coefficient of variation,
- se: standard error for the partial coefficient of variation (only if `include_se` is TRUE).

Note that conditional estimates are NA for a variable if that variable is the only variable in the predictor of the `risq` object model.

See Also

`risq` object constructor: [risq\(\)](#)

Other coefficient of variation functions: [cv\(\)](#), [cv_by_cat\(\)](#)

Examples

```
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
cv_by_var(risq_hlc, "response_1", c("gender", "age", "job"))
```

hlc

Household Living Conditions data

Description

A subset of data from a survey on Household Living Conditions conducted by Statistics Netherlands.

Usage

hlc

Format

hlc is a data frame with 35455 rows and 8 columns:

- age: age category (factor),
- gender: gender (factor),
- marital_status: marital status (factor),
- household: household type (factor),
- job: employment status (factor),
- house_value: house value category (factor),
- urbanisation: degree of urbanisation (factor),
- response_1: survey response status (logical).
- response_2: survey response status (logical).
- response_3: survey response status (logical).

Source

Statistics Netherlands.

mean.bootstrap

Arithmetic Mean of a Bootstrap

Description

Calculate the arithmetic mean over all bootstrap iterations for each indicator in the [bootstrap](#) object.

Usage

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

Arguments

x A bootstrap object.

... Additional arguments to control the mean calculation. For details see [base::mean\(\)](#).

Value

For a bootstrap using a non-partial indicator function, a single mean value. For a bootstrap using a partial indicator function, a data frame with the same shape as the output of that partial indicator function.

See Also

Other bootstrap methods: [-.bootstrap\(\)](#), [bootstrap\(\)](#), [quantile.bootstrap\(\)](#), [var.bootstrap\(\)](#)

Examples

```
# Note: a low iteration count is used to limit computing time of example.
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
bs <- bootstrap(
  risq_hlc,
  ri_by_var,
  target = "response_1",
  variables = "age",
  iterations = 5
)
mean(bs)
```

print.risq	<i>Print a RISQ object</i>
------------	----------------------------

Description

Print a description of a [risq](#) object.

Usage

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

Arguments

x	A risq object.
...	Unused.

Value

No return value.

See Also

Other risq methods: [risq\(\)](#), [sample.risq\(\)](#)

quantile.bootstrap	<i>Quantiles of a Bootstrap</i>
--------------------	---------------------------------

Description

Calculate sample quantiles over all bootstrap iterations for each indicator in the [bootstrap](#) object.

Usage

```
## S3 method for class 'bootstrap'  
quantile(x, ...)
```

Arguments

x	A bootstrap object.
...	Additional arguments to control the quantile calculation. For details see stats::quantile() .

Value

For a bootstrap using a non-partial indicator function, a vector with quantile values. For a bootstrap using a partial indicator function, a data frame with the same shape as the output of that partial indicator function, but with a quantile column for each entry in probs, instead of a single value column.

See Also

Other bootstrap methods: [-.bootstrap\(\)](#), [bootstrap\(\)](#), [mean.bootstrap\(\)](#), [var.bootstrap\(\)](#)

Examples

```
# Note: a low iteration count is used to limit computing time of example.
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
bs <- bootstrap(risq_hlc, cv, target = "response_1", iterations = 5)
quantile(bs, probs = c(0.05, 0.95))
```

ri	<i>Representativity Indicator</i>
----	-----------------------------------

Description

Estimate the bias-adjusted representativity indicator for a target variable, within the context of the given [risq](#) object.

Usage

```
ri(x, target, include_se = TRUE)
```

Arguments

x	A risq object.
target	Name of the target variable. Must be the name of a logical variable in the risq object data.
include_se	An optional logical specifying whether to include the standard error or not. If omitted, the standard error is included.

Value

Named list with the bias-adjusted representativity indicator value and, if include_se is TRUE, the associated standard error se.

See Also

risq object constructor: [risq\(\)](#)

Other representativity indicator functions: [ri_by_cat\(\)](#), [ri_by_var\(\)](#)

Examples

```
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
ri(risq_hlc, "response_1")
```

risq

*RISQ: Representativity Indicators for Survey Quality***Description**

Build a risq object that can serve as input for RISQ functions.

Usage

```
risq(
  predictor,
  family = c("binomial", "gaussian"),
  data,
  weights = NULL,
  strata = NULL
)
```

Arguments

predictor	An object of class <code>formula</code> describing the linear model of explanatory variables used to predict the response. The left-hand side must be empty.
family	An optional string that specifies the regression type. Use "binomial" for logistic regression or "gaussian" for linear regression. If omitted, logistic regression is used.
data	A data frame with sample data of a survey. Must contain all variables used in predictor. Response variables must be logical. Other variables must be factors if they are to be used for the calculation of partial indicators.
weights	An optional vector with inclusion weights for the sampling units. Values must be strictly positive. If omitted, the inclusion weights are set to 1.
strata	An optional factor with the strata membership of the sampling units. If omitted, a default is created based on the inclusion weights. The default uses a stratum for each unique weight, up to a maximum of 20. If the number of unique weights exceeds 20, a single stratum is used.

Value

A risq object that can serve as input for other functions in this package.

See Also

Response rate: `rr()`

Representativity indicator: `ri()`, `ri_by_var()`, `ri_by_cat()`

Coefficient of variation: `cv()`, `cv_by_var()`, `cv_by_cat()`

Other risq methods: `print.risq()`, `sample.risq()`

Examples

```
risq(predictor = ~ gender + age, data = hlc)
```

ri_by_cat

Representativity Indicator by Category

Description

Estimate unconditional or conditional partial representativity indicator values by category for a selection of categorical variables, given a `risq` object and target variable.

Usage

```
ri_by_cat(
  x,
  target,
  variables,
  type = c("unconditional", "conditional"),
  include_se = TRUE
)
```

Arguments

<code>x</code>	A <code>risq</code> object.
<code>target</code>	Name of the target variable. Must be the name of a logical variable in the <code>risq</code> object data.
<code>variables</code>	A character vector that specifies the variables for which to estimate partial representativity indicator values. Must be the names of categorical variables in the <code>risq</code> object data. May include names of variables that are not part of the model.
<code>type</code>	An optional string that specifies the type of partial representativity indicator to estimate. Must be either "unconditional" or "conditional". Defaults to "unconditional".
<code>include_se</code>	An optional logical specifying whether to include the standard error or not. If omitted, the standard error is included.

Value

A data frame with columns

- `variable`: name of the variable for which the row holds values,
- `category`: category within the variable for which the row holds values,
- `value`: estimate for the partial representativity indicator,
- `se`: standard error for the partial representativity indicator (only if `include_se` is `TRUE`).

Note that estimates are NA for empty categories and that conditional estimates are NA for a variable if that variable is the only variable in the predictor of the `risq` object model.

See Also

`risq` object constructor: [risq\(\)](#)

Other representativity indicator functions: [ri\(\)](#), [ri_by_var\(\)](#)

Examples

```
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
ri_by_cat(risq_hlc, "response_1", c("gender", "age", "job"))
```

ri_by_var

Representativity Indicator by Variable

Description

Estimate bias-adjusted unconditional or conditional partial representativity indicator values by variable for a selection of categorical variables, given a [risq](#) object and target variable.

Usage

```
ri_by_var(
  x,
  target,
  variables,
  type = c("unconditional", "conditional"),
  include_se = TRUE
)
```

Arguments

<code>x</code>	A <code>risq</code> object.
<code>target</code>	Name of the target variable. Must be the name of a logical variable in the <code>risq</code> object data.
<code>variables</code>	A character vector that specifies the variables for which to estimate partial representativity indicator values. Must be the names of categorical variables in the <code>risq</code> object data. May include names of variables that are not part of the model.
<code>type</code>	An optional string that specifies the type of partial representativity indicator to estimate. Must be either "unconditional" or "conditional". Defaults to "unconditional".
<code>include_se</code>	An optional logical specifying whether to include the standard error or not. If omitted, the standard error is included.

Value

A data frame with columns

- `variable`: name of the variable for which the row holds values,
- `value`: estimate for the partial representativity indicator,
- `se`: standard error for the partial representativity indicator (only if `include_se` is TRUE).

Note that conditional estimates are NA for a variable if that variable is the only variable in the predictor of the `risq` object model.

See Also

`risq` object constructor: [risq\(\)](#)

Other representativity indicator functions: [ri\(\)](#), [ri_by_cat\(\)](#)

Examples

```
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
ri_by_var(risq_hlc, "response_1", c("gender", "age", "job"))
```

`rr`

Response Rate

Description

Calculate the response rate for a target variable, within the context of the given [risq](#) object.

Usage

```
rr(x, target)
```

Arguments

x	A risq object.
target	Name of the target variable. Must be the name of a logical variable in the data component of the risq object.

Value

Response rate numeric value.

See Also

risq object constructor: [risq\(\)](#)

Examples

```
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
rr(risq_hlc, "response_1")
```

sample

Random Sampling

Description

Generic function for taking samples. Defaults to [base::sample\(\)](#).

Usage

```
sample(x, ...)
```

Arguments

x	Object to sample.
...	Further arguments to control the sampling.

Value

A sample of x.

See Also

[base::sample\(\)](#), [sample.risq\(\)](#)

sample.risq	<i>Sample a RISQ object</i>
-------------	-----------------------------

Description

Take a sample of the specified size from a [risq](#) object. The sampling can be done either with or without replacement, and can be weighted or unweighted. The default behaviour is unweighted without replacement.

Usage

```
## S3 method for class 'risq'
sample(x, size = nrow(x$data), ...)
```

Arguments

x	A risq object.
size	An optional positive integer specifying the number of items to sample. Defaults to the number of records in the risq object data.
...	Other arguments to be passed into base::sample() to control the sampling of records from the risq object.

Value

A risq object with the same model as x, but with a sample of the data, weights and strata of x.

See Also

Other risq methods: [print.risq\(\)](#), [risq\(\)](#)

var	<i>Variance</i>
-----	-----------------

Description

Generic function for calculating variance. Defaults to [stats::var\(\)](#).

Usage

```
var(x, ...)
```

Arguments

x	Object to calculate variance for.
...	Further arguments to control variance calculation.

Value

The variance of x.

See Also

[stats::var\(\)](#), [var.bootstrap\(\)](#)

var.bootstrap

Variance of a Bootstrap

Description

Calculate the variance over all bootstrap iterations for each indicator in the [bootstrap](#) object.

Usage

```
## S3 method for class 'bootstrap'
var(x, ...)
```

Arguments

x A bootstrap object.
 ... Additional arguments to control the variance calculation. For details see [stats::var\(\)](#).

Value

For a bootstrap using a non-partial indicator function, a single variance value. For a bootstrap using a partial indicator function, a data frame with the same shape as the output of that partial indicator function.

See Also

Other bootstrap methods: [-.bootstrap\(\)](#), [bootstrap\(\)](#), [mean.bootstrap\(\)](#), [quantile.bootstrap\(\)](#)

Examples

```
# Note: a low iteration count is used to limit computing time of example.
risq_hlc <- risq(predictor = ~ gender + age, data = hlc)
bs <- bootstrap(
  risq_hlc,
  ri_by_cat,
  target = "response_1",
  variables = "age",
  iterations = 5
)
var(bs)
```

Index

- * **bootstrap methods**
 - .bootstrap, 2
 - bootstrap, 3
 - mean.bootstrap, 8
 - quantile.bootstrap, 9
 - var.bootstrap, 17
- * **coefficient of variation functions**
 - cv, 4
 - cv_by_cat, 5
 - cv_by_var, 6
- * **datasets**
 - hlc, 7
- * **representativity indicator functions**
 - ri, 10
 - ri_by_cat, 12
 - ri_by_var, 13
- * **risq methods**
 - print.risq, 9
 - risq, 11
 - sample.risq, 16
- .bootstrap, 2
- base::mean(), 8
- base::sample(), 15, 16
- bootstrap, 2, 3, 8, 9, 17
- bootstrap(), 2, 3, 8, 10, 17
- cv, 4
- cv(), 3, 6, 7, 12
- cv_by_cat, 5
- cv_by_cat(), 3, 4, 7, 12
- cv_by_var, 6
- cv_by_var(), 3, 4, 6, 12
- formula, 11
- hlc, 7
- mean.bootstrap, 8
- mean.bootstrap(), 3, 10, 17
- print.risq, 9
- print.risq(), 12, 16
- quantile.bootstrap, 9
- quantile.bootstrap(), 3, 8, 17
- ri, 10
- ri(), 3, 12–14
- ri_by_cat, 12
- ri_by_cat(), 3, 10, 12, 14
- ri_by_var, 13
- ri_by_var(), 3, 10, 12, 13
- risq, 3–6, 9, 10, 11, 12–14, 16
- risq(), 3, 4, 6, 7, 9, 10, 13–16
- rr, 14
- rr(), 12
- sample, 15
- sample.risq, 16
- sample.risq(), 9, 12, 15
- set.seed(), 3
- stats::quantile(), 9
- stats::var(), 16, 17
- var, 16
- var.bootstrap, 17
- var.bootstrap(), 3, 8, 10, 17