

Package ‘splithalfr’

June 7, 2019

Title Extensible Bootstrapped Split-Half Reliabilities

Version 1.0.10

Date 2019-06-06

Author Thomas Pronk [aut, cre]

Description Calculates scores and estimates bootstrapped split-half reliabilities for reaction time tasks and questionnaires. The 'splithalfr' can be extended with custom scoring algorithms for user-provided datasets. For more information, see Parsons, Kruijt, & Fox (2018) <doi:10.31234/osf.io/6ka9z>.

Depends R (>= 3.6)

License GPL-3

Encoding UTF-8

LazyData true

Suggests psych (>= 1.8.12), knitr (>= 1.20), rmarkdown (>= 1.10), testthat (>= 2.1.0)

Imports dplyr (>= 0.8.1), rlang (>= 0.3.4)

RoxygenNote 6.1.1.9000

VignetteBuilder knitr

URL <https://thomaspronk.com>,
<https://www.uva.nl/en/profile/p/r/t.pronk/t.pronk.html>

NeedsCompilation no

Maintainer Thomas Pronk <pronkthomas@gmail.com>

Repository CRAN

Date/Publication 2019-06-07 08:00:03 UTC

R topics documented:

ds_aat	2
ds_iat	3
ds_rapi	4

ds_vpt	5
mean_fr_by_split	6
mean_sb_by_split	6
sh_apply	7
splithalfr	8

Index	10
--------------	-----------

ds_aat	<i>Example Approach Avoidance Task (AAT) Measurement Data in JASMIN2 Format</i>
--------	---

Description

The JASMIN2 AAT was an irrelevant feature task, in which participants were instructed to approach/avoid left/right rotated stimuli. This particular AAT presented stimuli from a "test" category, which were math-related pictures, and from a "control" category, which were pictures unrelated to math. It registered approach responses by participants pressing (and holding) the arrow down key, while avoid responses were given via the arrow up key. Upon a response, the stimulus zoomed in or out, until it disappeared from the screen. The first response to a stimulus was logged, as well as the final response, as defined by the stimulus completely zooming in or out. Upon a correct final response the next trial started and upon an incorrect final response the current trial was repeated. The first and final response to each trial was logged. The dataset contains one row per trial. This dataset was graciously provided by [Eva Schmitz](#).

Usage

ds_aat

Format

An object of class `data.frame` with 24480 rows and 13 columns.

Details

- `UserID`. Identifies participants
- `approach_left`. If `TRUE`, participants were instructed to approach left rotated stimuli. If `FALSE`, participants were instructed to approach right rotated stimuli.
- `trial_type`. Values: "approach" or "avoid"
- `cat`. Stimulus category: practice, test, or control
- `stim`. Stimulus
- `response`. Initial response; 1 = correct, 2 = incorrect, 3 = timeout (no response in 4000 ms), 4 = invalid key
- `rt`. Response time in milliseconds
- `sust`. Was approach or avoid response sustained until the stimulus was completely zoomed in or out?

- final_response. Final response; the response that ended the current trial. Possible values are the same as for response
- block. Counts blocks, starting at zero
- trial. Counts trials in blocks, starting at zero
- attempt. Counts attempts (responses) in trials, starting at zero
- block_type. Type of block: "assess1" and "assess2" for assessment trials with salient stimuli

 ds_iat

Example Implicit Association Task (IAT) Data in JASMIN2 Format

Description

The JASMIN2 IAT closely followed the original IAT procedure (Greenwald, McGhee, & Schwartz, 1998), except that target and attribute trials did not alternate. Upon a correct response, the next trial started. Upon an incorrect response, the current trial was repeated. The response to each trial was logged. This particular dataset is from a Drinker Identity IAT (Lindgren et al., 2013) that was used in a cross-cultural study. Drinker was target 1, non-drinker was target 2, me was attribute 1, and not-me was attribute 2. The dataset contains one row per trial. This dataset was graciously provided by Helle Larsen.

Usage

ds_iat

Format

An object of class `data.frame` with 3568 rows and 11 columns.

Details

Overview of columns:

- UserID. Identifies participants
- t1_left. If TRUE, the first combination block had target 1 on the left (and target 2 on the right)
- a1_left. If TRUE, the first combination block had attribute 1 on the left (and attribute 2 on the right)
- cat. Category that stimulus belonged to
- stim. Stimulus
- response. Response; 1 = correct, 2 = incorrect, 3 = timeout (no response in 4000 ms), 4 = invalid key
- rt. Response time in milliseconds. Note that some response times may exceed the timeout window due to clock errors on the computer that the IAT was administered
- block. Counts blocks, starting at zero
- trial. Counts trials in blocks, starting at zero

- `attempt`. Counts attempts (responses) in trials, starting at zero
- `block_type`. Type of block

The variable `block_type` can have these values:

- `tar_discr`: target discrimination
- `att_discr`: attribute discrimination
- `tar1att1_1`: target 1 with attitude 1, practice block
- `tar1att1_2`: target 1 with attitude 1, test block
- `tar_rev`: reverse target discrimination
- `tar1att2_1`: target 1 with attitude 2, practice block
- `tar1att2_2`: target 1 with attitude 2, test block

ds_rapi

Example 23-item Rutgers Alcohol Problem Inventory (RAPI) data

Description

The RAPI is a questionnaire which asks how often a participant experienced each of 23 alcohol-related problems within the last year (White & Labouvie, 1989). The dataset contains one row per participant.

Usage

```
ds_rapi
```

Format

An object of class `data.frame` with 426 rows and 24 columns.

Details

The dataset contains the following columns:

- `twnr`. Identifies participants
- `V1` to `V23`. Answers on each of the 23 RAPI items

Each item is answered on a four-point scale with the following answer options:

- 0 = None
- 1 = 1-2 times
- 2 = 3-5
- 3 = More than 5 times

`ds_vpt`*Example Visual Probe Task (VPT) Measurement Data in JASMINI Format*

Description

The JASMINI VPT distinguished between "test" stimuli, which are in some way assumed to be salient to the participant and "control" stimuli, which are not. Test and control stimuli were presented in pairs, with one left and one right, followed by a probe that was an arrow pointing up or down. Participants needed to indicate whether the arrow pointed up or down. Upon a correct response the next trial started and upon an incorrect response the current trial was repeated. Only the first response to a new trial was logged. This particular VPT was part of the pre-measurement of a cognitive bias modification study. The "test" stimuli were alcoholic beverages and the "control" stimuli were non-alcoholic beverages, selected from the Amsterdam Beverage Picture Set (Pronk, Deursen, Beraha, Larsen, & Wiers, 2015). The dataset contains one row per trial. This dataset was graciously provided by Marilisa Boffo.

Usage

`ds_vpt`

Format

An object of class `data.frame` with 19520 rows and 12 columns.

Details

- `UserID`. Identifies participants
- `patt`. Probe-at-test. If "yes", the probe was positioned at the test stimulus. If "no", the probe was positioned at the control stimulus.
- `phor`. Probe horizontal position. Values: "left" or "right"
- `thor`. Test horizontal position. Values: "left" or "right"
- `keep`. If "yes" the probe was superimposed on the stimuli. If "no" the probe replaced the stimuli.
- `pdir`. Probe direction. Values: "up" or "down"
- `stim`. Stimulus
- `response`. Response; 1 = correct, 2 = incorrect, 3 = timeout (no response in 5000 ms), 4 = invalid key
- `rt`. Response time in milliseconds
- `block`. Counts blocks, starting at zero
- `trial`. Counts trials in blocks, starting at zero
- `block_type`. Type of block: "assess" for assessment trials with salient stimuli

mean_fr_by_split	<i>Flanagan-Rulon reliability coefficient averaged over each split. Can be applied to output of sh_apply</i>
------------------	--

Description

Flanagan-Rulon reliability coefficient averaged over each split. Can be applied to output of sh_apply

Usage

```
mean_fr_by_split(ds)
```

Arguments

ds (data frame) a data frame with columns "split", "score_1", and "score_2"

Value

(numeric) mean Flanagan-Rulon coefficient

Examples

```
# N.B. This example uses R script from the vignette: "rapi_sum"
data("ds_rapi", package = "splithalfr")
rapi_fn_sets <- function (ds) {
  return (list(
    items = unlist(ds[paste("V", 1 : 23, sep = "")])
  ))
}
rapi_fn_score <- function (sets) {
  return (sum(sets$items))
}
ds_splits = sh_apply(ds_rapi, "twrn", rapi_fn_sets, rapi_fn_score, 10)
mean_fr_by_split(ds_splits)
```

mean_sb_by_split	<i>Spearman-Brown reliability coefficient averaged over each split. Can be applied to output of sh_apply</i>
------------------	--

Description

Spearman-Brown reliability coefficient averaged over each split. Can be applied to output of sh_apply

Usage

```
mean_sb_by_split(ds)
```

Arguments

ds (data frame) a data frame with columns "split", "score_1", and "score_2"

Value

(numeric) mean Spearman-Brown coefficient

Examples

```
# N.B. This example uses R script from the vignette: "rapi_sum"
data("ds_rapi", package = "splithalfr")
rapi_fn_sets <- function (ds) {
  return (list(
    items = unlist(ds[paste("V", 1 : 23, sep = "")])
  ))
}
rapi_fn_score <- function (sets) {
  return (sum(sets$items))
}
ds_splits = sh_apply(ds_rapi, "twnr", rapi_fn_sets, rapi_fn_score, 10)
mean_sb_by_split(ds_splits)
```

sh_apply	<i>Apply a scoring algorithm to each participant using full or split data set</i>
----------	---

Description

Apply a scoring algorithm to each participant using full or split data set

Usage

```
sh_apply(ds, participant_id, fn_sets, fn_score, split_count = 0,
  show_progress = TRUE)
```

Arguments

ds (data frame) data frame containing data to score.

participant_id (character) name of column that identifies participants in ds.

fn_sets (function) receives data from a single participant, should return a list of sets that may be split. Elements of sets that are data frames are split by row. Elements of sets that are lists or vectors are split by element.

fn_score (function) receives full or split sets, should return a single number.

split_count (numeric) Default: 0. If 0, applies fn_score on full set. If > 0, applies fn_score to split sets, split_count times.

show_progress (logical) Default: TRUE. If TRUE, prints current split number each split.

Value

(data frame) If `split_count == 0`, returns a data frame with a column for `participant_id` and a column named "score" for `fn_score` applied to the full data of each participant. If `split_count > 0`, it splits each element returned by `fn_sets` into two halves that differ at most by one in size, applies `fn_score` on split data, and returns a data frame with a column for `participant_id`, a column "split" that counts splits, and "score_1" and "score_2" with the score of each split.

Examples

```
# N.B. This example uses R script from the vignette: "rapi_sum"
data("ds_rapi", package = "splithalfr")
rapi_fn_sets <- function (ds) {
  return (list(
    items = unlist(ds[paste("V", 1 : 23, sep = "")])
  ))
}
rapi_fn_score <- function (sets) {
  return (sum(sets$items))
}
# Calculate scores per participant on full data
sh_apply(ds_rapi, "twnr", rapi_fn_sets, rapi_fn_score)
# Calculate split scores per participant ten times
sh_apply(ds_rapi, "twnr", rapi_fn_sets, rapi_fn_score, 10)
```

splithalfr

splithalfr: Extensible Bootstrapped Split-Half Reliabilities

Description

Calculates scores and estimates bootstrapped split-half reliabilities for reaction time (RT) tasks and questionnaires. The `splithalfr` can be extended with custom scoring algorithms for user-provided datasets.

Details

The `splithalfr` vignettes demonstrate how to write a custom scoring algorithm based on included example scoring algorithms and datasets:

- `vignette("rapi_sum")` Sum-score for data of the 23-item version of the Rutgers Alcohol Problem Index (White & Labouvie, 1989)
- `vignette("vpt_diff_of_means")` Difference of mean RTs for correct responses, after removing RTs below 200 ms and above 520 ms, on Visual Probe Task data (Mogg & Bradley, 1999)
- `vignette("aat_double_diff_of_medians")` Double difference of medians for correct responses on Approach Avoidance Task data (Heuer, Rinck, & Becker, 2007)
- `vignette("iat_dscore_ri")` Improved d-score algorithm for data of an Implicit Association Task that requires a correct response in order to continue to the next trial (Greenwald, Nosek, & Banaji, 2003)

Validation of algorithms

- The R script included in each vignette has been validated by comparing the *splithalfr* score for a single participant with the same score calculated via Excel. The materials for each test can be found in [the *splithalfr* GitHub repository](#).
- The *splithalfr* splitting algorithm has been validated via a set of simulations that are not included in this package. The R script for these simulations can be found [here](#).

Related packages

These R packages offer bootstrapped split-half reliabilities for specific scoring algorithms and are available via CRAN at the time of this writing: [multicon](#), [psych](#), and [splithalfr](#).

Index

*Topic **datasets**

- ds_aat, [2](#)
- ds_iat, [3](#)
- ds_rapi, [4](#)
- ds_vpt, [5](#)

- ds_aat, [2](#)
- ds_iat, [3](#)
- ds_rapi, [4](#)
- ds_vpt, [5](#)

- mean_fr_by_split, [6](#)
- mean_sb_by_split, [6](#)

- sh_apply, [7](#)
- splithalfr, [8](#)
- splithalfr-package (splithalfr), [8](#)