Package ‘trimr’

October 14, 2022

Version 1.1.1
Date 2022-05-03
Title An Implementation of Common Response Time Trimming Methods
Description Provides various commonly-used response time trimming methods, including the recursive / moving-criterion methods reported by Van Selst and Jolicoeur (1994). By passing trimming functions raw data files, the package will return trimmed data ready for inferential testing.

Depends R (>= 4.0)
Imports stats, dplyr
License GPL-3
LazyData true

URL https://github.com/JimGrange/trimr

BugReports https://github.com/JimGrange/trimr/issues
Suggests knitr, rmarkdown
VignetteBuilder knitr
RoxygenNote 7.1.1
NeedsCompilation no
Author James Grange [cre, aut], Ed Berry [ctb]
Maintainer James Grange <grange.jim@gmail.com>
Repository CRAN
Date/Publication 2022-05-05 02:20:02 UTC

R topics documented:

absoluteRT .......................................................... 2
exampleData .......................................................... 3
hybridRecursive ..................................................... 4
linearInterpolation .................................................. 5
Description

`absoluteRT` takes a data frame of RT data and returns trimmed rt data that fall between set minimum and maximum limits.

Usage

```r
absoluteRT(
  data,
  minRT,
  maxRT,
  pptVar = "participant",
  condVar = "condition",
  rtVar = "rt",
  accVar = "accuracy",
  omitErrors = TRUE,
  returnType = "mean",
  digits = 3
)
```

Arguments

data: A data frame with columns containing: participant identification number ("pptVar"); condition identification, if applicable ("condVar"); response time data ("rtVar"); and accuracy ("accVar"). The RT can be in seconds (e.g., 0.654) or milliseconds (e.g., 654). Typically, "condition" will consist of strings. Accuracy must be coded as 1 for correct and 0 for error responses.

minRT: The lower criteria for acceptable response time. Must be in the same form as rt column in data frame (e.g., in seconds or milliseconds).

maxRT: The upper criteria for acceptable response time. Must be in the same form as rt column in data frame (e.g., in seconds or milliseconds).

pptVar: The quoted name of the column in the data that identifies participants.

condVar: The quoted name of the column in the data that includes the conditions.

rtVar: The quoted name of the column in the data containing reaction times.

accVar: The quoted name of the column in the data containing accuracy, coded as 0 or 1 for incorrect and correct trial, respectively.
omitErrors

If set to TRUE, error trials will be removed before conducting trimming procedure. Final data returned will not be influenced by errors in this case.

returnType

Request nature of returned data. "raw" returns trial-level data excluding trimmed data; "mean" returns mean response times per participant for each experimental condition identified; "median" returns median response times per participant for each experimental condition identified.

digits

How many decimal places to round to after trimming?

Details

By passing a data frame containing raw response time data, together with trimming criteria, the function will return trimmed data, either in the form of trial-level data or in the form of means/medians for each subject & condition.

Examples

# load the example data that ships with trimr
data(exampleData)

# perform the trimming, returning mean RT
trimmedData <- absoluteRT(data = exampleData, minRT = 150, maxRT = 2500, returnType = "mean")

exampleData

Example response time data set

Description

An example data set containing multiple participants’ data for a response time study involving two experimental conditions. The data set also includes This is a synthetic data set and has no theoretical basis.

Usage

e.exampleData

Format

A data frame with 20518 rows and 4 variables:

- participant: participant identification number
- condition: the experimental condition (2 in this example)
- rt: response time, coded in milliseconds
- accuracy: accuracy of the response; 1 = correct, 0 = error
Description

`hybridRecursive` takes a data frame of RT data and returns trimmed rt data. The returned value is the average returned from the nonRecursive and the modifiedRecursive procedures as described in van Selst & Jolicoeur (1994).

Usage

```r
hybridRecursive(
  data,
  minRT,
  pptVar = "participant",
  condVar = "condition",
  rtVar = "rt",
  accVar = "accuracy",
  omitErrors = TRUE,
  digits = 3
)
```

Arguments

- **data**: A data frame with columns containing: participant identification number ('pptVar'); condition identification, if applicable ('condVar'); response time data ('rtVar'); and accuracy ('accVar'). The RT can be in seconds (e.g., 0.654) or milliseconds (e.g., 654). Typically, "condition" will consist of strings. Accuracy must be coded as 1 for correct and 0 for error responses.
- **minRT**: The lower criteria for acceptable response time. Must be in the same form as rt column in data frame (e.g., in seconds OR milliseconds). All RTs below this value are removed before proceeding with SD trimming.
- **pptVar**: The quoted name of the column in the data that identifies participants.
- **condVar**: The quoted name of the column in the data that includes the conditions.
- **rtVar**: The quoted name of the column in the data containing reaction times.
- **accVar**: The quoted name of the column in the data containing accuracy, coded as 0 or 1 for incorrect and correct trial, respectively.
- **omitErrors**: If set to TRUE, error trials will be removed before conducting trimming procedure. Final data returned will not be influenced by errors in this case.
- **digits**: How many decimal places to round to after trimming?

References

Examples

```r
# load the example data that ships with trimr
data(exampleData)

# perform the trimming, returning mean RT
trimmedData <- hybridRecursive(data = exampleData, minRT = 150)
```

---

`linearInterpolation`  
SDs used for the recursive / moving criterion trimming methods

Description

A data frame containing the SDs used for each sample size as trimming criterion for the nonRecursive function and the modifiedRecursive function

Usage

`linearInterpolation`

Format

A data frame with 97 rows and 3 columns:

<table>
<thead>
<tr>
<th>sampleSize</th>
<th>Sample size of the data set being passed</th>
</tr>
</thead>
<tbody>
<tr>
<td>nonRecursive</td>
<td>The standard deviation to use as the criterion for the nonRecursive function</td>
</tr>
<tr>
<td>modifiedRecursive</td>
<td>The standard deviation to use as the criterion for the modifiedRecursive function</td>
</tr>
</tbody>
</table>

Description

modifiedRecursive takes a data frame of RT data and returns trimmed rt data that fall below a set standard deviation above the each participant’s mean for each condition, with the criterion changing as more trials are removed, as described in van Selst & Jolicoeur (1994).
Usage

modifiedRecursive(
  data,
  minRT,
  pptVar = "participant",
  condVar = "condition",
  rtVar = "rt",
  accVar = "accuracy",
  omitErrors = TRUE,
  returnType = "mean",
  digits = 3
)

Arguments

data          A data frame with columns containing: participant identification number ('ppt-Var'); condition identification, if applicable ('condVar'); response time data ('rt-Var'); and accuracy ('accVar'). The RT can be in seconds (e.g., 0.654) or milliseconds (e.g., 654). Typically, "condition" will consist of strings. Accuracy must be coded as 1 for correct and 0 for error responses.
minRT         The lower criteria for acceptable response time. Must be in the same form as rt column in data frame (e.g., in seconds OR milliseconds). All RTs below this value are removed before proceeding with SD trimming.
pptVar        The quoted name of the column in the data that identifies participants.
condVar       The quoted name of the column in the data that includes the conditions.
rtVar         The quoted name of the column in the data containing reaction times.
accVar        The quoted name of the column in the data containing accuracy, coded as 0 or 1 for incorrect and correct trial, respectively.
omitErrors    If set to TRUE, error trials will be removed before conducting trimming procedure. Final data returned will not be influenced by errors in this case.
returnType    Request nature of returned data. "raw" returns trial-level data excluding trimmed data; "mean" returns mean response times per participant for each experimental condition identified; "median" returns median response times per participant for each experimental condition identified.
digits        How many decimal places to round to after trimming?

References


Examples

# load the example data that ships with trimr
data(exampleData)
# perform the trimming, returning mean RT
trimmedData <- modifiedRecursive(data = exampleData, minRT = 150,
returnType = "mean")

## nonRecursive

### Description

`nonRecursive` takes a data frame of RT data and returns trimmed rt data that fall below a set standard deviation above the each participant's mean for each condition. The SD used for trimming is proportional to the number of trials in the data being passed, as described in van Selst & Jolicoeur (1994).

### Usage

```r
nonRecursive(
  data,
  minRT,
  pptVar = "participant",
  condVar = "condition",
  rtVar = "rt",
  accVar = "accuracy",
  omitErrors = TRUE,
  returnType = "mean",
  digits = 3
)
```

### Arguments

- **data**: A data frame with columns containing: participant identification number ('pptVar'); condition identification, if applicable ('condVar'); response time data ('rtVar'); and accuracy ('accVar'). The RT can be in seconds (e.g., 0.654) or milliseconds (e.g., 654). Typically, "condition" will consist of strings. Accuracy must be coded as 1 for correct and 0 for error responses.
- **minRT**: The lower criteria for acceptable response time. Must be in the same form as `rtVar` column in data frame (e.g., in seconds OR milliseconds). All RTs below this value are removed before proceeding with SD trimming.
- **pptVar**: The quoted name of the column in the data that identifies participants.
- **condVar**: The quoted name of the column in the data that includes the conditions.
- **rtVar**: The quoted name of the column in the data containing reaction times.
- **accVar**: The quoted name of the column in the data containing accuracy, coded as 0 or 1 for incorrect and correct trial, respectively.
- **omitErrors**: If set to TRUE, error trials will be removed before conducting trimming procedure. Final data returned will not be influenced by errors in this case.
- **returnType**: The type of return, either "mean" or "median".
- **digits**: The number of decimals to round the results to.

### Example Usage

```r
# Load example data
exampleData <- read.csv("exampleData.csv")

# Call nonRecursive function
trimmedData <- nonRecursive(data = exampleData, minRT = 150,
returnType = "mean")

# Summary statistics
summary(trimmedData)
```

### Note

- This function is designed for non-recursive trimming, which means it does not consider the possibility of nested conditions within the data.
- The SD for trimming is calculated based on the number of trials in each condition, ensuring that the trimming process is fair across different condition sizes.
- The function ensures that only trials above the specified threshold are considered for trimming, providing a more accurate representation of the data.

---

**References**

returnType Request nature of returned data. "raw" returns trial-level data excluding trimmed data; "mean" returns mean response times per participant for each experimental condition identified; "median" returns median response times per participant for each experimental condition identified.

digits How many decimal places to round to after trimming?

References

Examples

# load the example data that ships with trimr
data(exampleData)

# perform the trimming, returning mean RT
trimmedData <- nonRecursive(data = exampleData, minRT = 150,
returnType = "mean")

sdTrim
RT trimming with standard deviation criterion

Description
sdTrim takes a data frame of RT data and returns trimmed rt data that fall below a set set criterion (based on standard deviations above a particular mean). The criterion can be based on the mean of the whole set of data, based on the mean per experimental condition, based on the mean per participant, or based on the mean of each participant in each experimental condition.

Usage

sdTrim(
data, minRT, sd, pptVar = "participant", condVar = "condition", rtVar = "rt", accVar = "accuracy", perCondition = TRUE, perParticipant = TRUE, omitErrors = TRUE, returnType = "mean", digits = 3)

)
Arguments

- **data**: A data frame with columns containing: participant identification number ('pptVar'); condition identification, if applicable ('condVar'); response time data ('rtVar'); and accuracy ('accVar'). The RT can be in seconds (e.g., 0.654) or milliseconds (e.g., 654). Typically, "condition" will consist of strings. Accuracy must be coded as 1 for correct and 0 for error responses.

- **minRT**: The lower criteria for acceptable response time. Must be in the same form as rt column in data frame (e.g., in seconds OR milliseconds). All RTs below this value are removed before proceeding with SD trimming.

- **sd**: The upper criteria for standard deviation cut-off.

- **pptVar**: The quoted name of the column in the data that identifies participants.

- **condVar**: The quoted name of the column in the data that includes the conditions.

- **rtVar**: The quoted name of the column in the data containing reaction times.

- **accVar**: The quoted name of the column in the data containing accuracy, coded as 0 or 1 for incorrect and correct trial, respectively.

- **perCondition**: Set to TRUE if the user wishes the trimming to occur per condition of the experimental design.

- **perParticipant**: Set to TRUE if the user wishes the trimming to occur per participant.

- **omitErrors**: If set to TRUE, error trials will be removed before conducting trimming procedure. Final data returned will not be influenced by errors in this case.

- **returnType**: Request nature of returned data. "raw" returns trial-level data excluding trimmed data; "mean" returns mean response times per participant for each experimental condition identified; "median" returns median response times per participant for each experimental condition identified.

- **digits**: How many decimal places to round to after trimming?

Details

By passing a data frame containing raw response time data, together with trimming criteria, the function will return trimmed data, either in the form of trial-level data or in the form of means/medians for each subject & condition.

Examples

```r
# load the example data that ships with trimr
data(exampleData)

# perform the trimming with SD trimming per condition, returning mean RT
trimmedData <- sdTrim(data = exampleData, minRT = 150, sd = 2.5, perCondition = TRUE, perParticipant = FALSE, returnType = "mean")
```
Index

* datasets
  exampleData, 3
  linearInterpolation, 5

absoluteRT, 2
exampleData, 3
hybridRecursive, 4
linearInterpolation, 5
modifiedRecursive, 5
nonRecursive, 7
sdTrim, 8