

# Package ‘curtail’

January 2, 2016

**Type** Package

**Title** A package for creating and assessing performance of curtailed tests

**Version** 0.1

**Date** 2016-01-01

**Author** Marjolein Fokkema

**Maintainer** Marjolein Fokkema

**Description** A package for creating and assessing the performance of curtailed tests. A curtailed test is a variable-length test, which allows for early stopping of item administration when further items are unlikely or unable to change the final (classification) decision. The package allows for creating and assessing deterministically and stochastically (based on empirical proportions) curtailed tests. In future versions of the package, model-based stochastic curtailment may be added. Curtailed tests are created by the creating look-up tables, providing a user with item-specific cut-off values for early stopping of item administration. Curtailment performance is assessed in terms of efficiency (number of items administered to arrive at a final decision) and accuracy (concordance of the final decision based on curtailed testing with the final decision that would be obtained by administering the full-length test).

**License** GPL-2

**LazyData** TRUE

## R topics documented:

Curtail . . . . .	2
itemscores . . . . .	3
stochCurtail . . . . .	3
stochCurtailXval . . . . .	4
stochTable . . . . .	6
Table . . . . .	7
<b>Index</b>	<b>8</b>

---

Curtail

---

*Simulate deterministic curtailment on item scores of a test*


---

### Description

Simulates deterministic curtailment on an existing dataset of item scores, based on a user-specified cut-off value.

### Usage

```
Curtail(dataset.test, Xstar, highest = NULL, lowest = NULL)
```

### Arguments

<code>dataset.test</code>	A dataframe containing item scores only.
<code>Xstar</code>	Cut-off value to be used for classifying observations as 'at risk' (test-score values greater than or equal to the cut-off value) or 'not at risk' (test-score values less than cut-off value)
<code>highest</code>	Highest possible item score. If not provided, function will take the max of all observed item scores.
<code>lowest</code>	Lowest possible item score. If not provided, function will take the min of all observed item scores.

### Details

Note that currently, only items with the same number of response categories, and the same ordering of the response categories, are supported.

The code is still under development and might change in future versions.

### Value

The function returns a list with the following objects:

<code>item.scores</code>	Original dataset of item scores
<code>cumulative.scores</code>	Dataset of cumulative scores on every item
<code>current.item</code>	Vector with the number of the current item (last item administered in curtailed administration) for each observation in the dataset
<code>current.score</code>	Vector with the current score (cumulative score for last item administered) for each observation in the dataset
<code>curtailed.test.length.distribution</code>	List with descriptive statistics on efficiency: mean, standard deviation and median number of items administered
<code>accuracy</code>	List with descriptive statistics on accuracy: number of 'at risk' and 'not at risk' observations flagged as 'at risk' and 'not at risk' after application of curtailment

## References

Fokkema, M., Smits, N., Kelderman, H., Carlier, I.V. & Van Hemert, A.M. (2014). Curtailment: A method to reduce the length of self-report questionnaires while maintaining diagnostic accuracy. *Psychiatry Research* 215(2), 477-482.

Finkelman, M.D., Smits, N., Kim, W. & Riley, B. (2012). Curtailment and stochastic curtailment to shorten the CES-D. *Applied Psychological Measurement*, 36(8), 632-658.

## Examples

```
Curtail(itemscores, 19)
```

---

itemscores	<i>Dataset for illustrating curtailment functions</i>
------------	---

---

## Description

Dataset containing 1000 simulated responses to 20 items on a mental-health questionnaire. Items are scored 0, 1, 2 or 3. The dataset is used in the examples to illustrate the functions. Observations are classified as 'at risk' with a total score after administration of item 20, equal to or greater than 19 (i.e., a cut-off score of 19 is used).

## Examples

```
summary(itemscores)
```

---

stochCurtail	<i>Simulates stochastic curtailment on item scores of a test</i>
--------------	--

---

## Description

Simulates stochastic curtailment on an existing test dataset of item scores, based on an existing training dataset of item scores, and a user-specified cut-off value.

## Usage

```
stochCurtail(dataset.train, dataset.test = NULL, Xstar, gamma0=.95, gamma1=.95)
```

## Arguments

dataset.train	A dataframe, containing item scores only, which will be used to derive the probabilities of obtaining a final test score greater than, or equal to, the cut-off value, based on the current cumulative score
dataset.test	A dataframe containing item scores only. Curtailment will be simulated on these observations. When no test dataset is specified, curtailment will be simulated on the training dataset.
Xstar	Cut-off value to be used for classifying observations as 'at risk' (test-score values greater than or equal to the cut-off value) or 'not at risk' (test-score values less than cut-off value)

gamma0	The threshold for the probability, calculated using the 'not-at-risk' training observations, that the classification decision based on the stochastically curtailed version will match that of the full-length instrument.
gamma1	The threshold for the probability, calculated using the 'at-risk' training observations, that the classification decision based on the stochastically curtailed version will match that of the full-length instrument.

## Details

The code is still under development and might change in future versions.

## Value

The function returns a list with the following objects:

item.scores	Original test dataset of item scores
cumulative.scores	Cumulative scores on every item for the test dataset
current.item	Vector with the number of the current item (last item administered in curtailed administration) for each observation in the test dataset
current.score	Vector with the current score (cumulative score for last item administered) for each observation in the test dataset
curtailed.test.length.distribution	List with descriptive statistics on efficiency of the stochastically curtailed test in the test dataset: mean, standard deviation and median number of items administered
accuracy	List with descriptive statistics on accuracy: number of 'at risk' and 'not at risk' observations flagged as 'at risk' and 'not at risk' after application of stochastic curtailment in the test dataset

## Examples

```
## obtain a test and training dataset
set.seed(32061983)
samp <- sample(1:1000, 500); train <- samp[1:500]
trainingdata <- itemscores[train,]
testdata <- itemscores[-train,]
stochCurtail(trainingdata, testdata, 19)
## try lower gamma values for earlier stopping, but lower accuracy:
stochCurtail(trainingdata, testdata, 19, gamma0=.75, gamma1=.75)
```

---

stochCurtailXval

*Simulate stochastic curtailment using k-fold CV*

---

## Description

Simulates stochastic curtailment on an existing dataset of item scores, using leave-one-out (LOO) cross validation, and a user-specified cut-off value. For a given observation, all other observations will be used to derive the probabilities of obtaining a final test score greater than or equal to the cut-off value.

**Usage**

```
stochCurtailXval(dataset, Xstar, gamma0=.95, gamma1=.95)
```

**Arguments**

dataset	A dataframe, containing item scores only.
Xstar	Cut-off value to be used for classifying observations as 'at risk' (test-score values greater than or equal to the cut-off value) or 'not at risk' (test-score values less than cut-off value)
gamma0	The threshold for the probability, calculated using the 'not-at-risk' training observations, that the classification decision based on the stochastically curtailed version will match that of the full-length instrument.
gamma1	The threshold for the probability, calculated using the 'at-risk' training observations, that the classification decision based on the stochastically curtailed version will match that of the full-length instrument.

**Details**

The code is still under development and might change in future versions.

**Value**

The function returns an augmented dataset with the following variables:

Variables for all but the last item (therefore, variable names consist of the following characters, appended with the item number):

item	the original item scores
test	the cumulative sum score for the items up to this number
Pkplus	see Finkelman et al. (2012)
Pkmin	see Finkelman et al. (2012)

Further, the following variables are included:

plusflag	for internal calculations
minflag	for internal calculations
SCrisk	whether the observation is flagged as 'at risk' in curtailed test administration
SCnorisk	whether the observation is flagged as 'not at risk' in curtailed test administration
currit	the number of the current item (last item administered in curtailed administration) for each observation in the dataset
currts	the current test score (cumulative score for last item administered) for each observation in the dataset

**References**

Fokkema, M., Smits, N., Kelderman, H., Carlier, I.V. & Van Hemert, A.M. (2014). Curtailment: A method to reduce the length of self-report questionnaires while maintaining diagnostic accuracy. *Psychiatry Research* 215(2), 477-482.

Finkelman, M.D., Smits, N., Kim, W. & Riley, B. (2012). Curtailment and stochastic curtailment to shorten the CES-D. *Applied Psychological Measurement*, 36(8), 632-658.

## Examples

```
## Not run:
tmp <- stochCurtailXval(itemscores, 19, gamma0=.95, gamma1=.95)
table(tmp$test20>=19, tmp$SCrisk) # accuray
table(tmp$test20>=19, tmp$SCnorisk) # accuracy
table(tmp$currit); mean(tmp$currit); sd(tmp$currit) # efficiency

## End(Not run)
```

---

stochTable	<i>Create a look-up table for stochastic curtailment</i>
------------	--

---

## Description

Creates a look-up table for early stopping of item administration, for stochastic curtailment.

## Usage

```
stochTable(dataset.train, Xstar, gamma0 = 0.95, gamma1 = 0.95)
```

## Arguments

dataset.train	Dataframe, containing item scores only
Xstar	cut-off value for the final test score to be used
gamma0	The threshold for the probability, calculated using the 'not-at-risk' training observations, that the classification decision based on the stochastically curtailed version will match that of the full-length instrument.
gamma1	The threshold for the probability, calculated using the 'at-risk' training observations, that the classification decision based on the stochastically curtailed version will match that of the full-length instrument.

## Details

The code is still under development and might change in future versions.

## Value

Returns a look-up table with cut-off values for every item, to be used for early stopping of item administration.

## References

Fokkema, M., Smits, N., Kelderman, H., Carlier, I.V. & Van Hemert, A.M. (2014). Curtailment: A method to reduce the length of self-report questionnaires while maintaining diagnostic accuracy. *Psychiatry Research* 215(2), 477-482.

Finkelman, M.D., Smits, N., Kim, W. & Riley, B. (2012). Curtailment and stochastic curtailment to shorten the CES-D. *Applied Psychological Measurement*, 36(8), 632-658.

**Examples**

```
stochTable(itemscores, 19)
# try lower gamma values for earlier stopping:
stochTable(itemscores, 19, gamma0 = .75, gamma1 = .75)
```

---

Table	<i>Create a look-up table for deterministic curtailment</i>
-------	---

---

**Description**

Creates a look-up table for early stopping of item administration, for deterministic curtailment.

**Usage**

```
Table(dataset.train = NULL, Xstar, nitems = NULL, highest = NULL, lowest = NULL)
```

**Arguments**

dataset.train	Optional, a dataframe, containing item scores only
Xstar	cut-off value for the final test score to be used
nitems	number of items in questionnaire (assumed equal for all items)
highest	Optional, highest theoretically possible item score (assumed equal for all items)
lowest	Optional, lowest theoretically possible item score (assumed equal for all items)

**Details**

Note that either dataset.train argument needs to be specified, or the nitems, highest and lowest arguments.

The code is still under development and might change in future versions.

**Value**

Returns a look-up table with cut-off values for every item, to be used for early stopping of item administration.

**References**

Fokkema, M., Smits, N., Kelderman, H., Carlier, I.V. & Van Hemert, A.M. (2014). Curtailment: A method to reduce the length of self-report questionnaires while maintaining diagnostic accuracy. *Psychiatry Research* 215(2), 477-482.

Finkelman, M.D., Smits, N., Kim, W. & Riley, B. (2012). Curtailment and stochastic curtailment to shorten the CES-D. *Applied Psychological Measurement*, 36(8), 632-658.

**Examples**

```
Table(itemscores, 19)
Table(Xstar = 19, nitems = 20, highest = 3, lowest = 0)
```

# Index

## \*Topic **curtail, curtailment**

- Curtail, [2](#)
- stochCurtail, [3](#)
- stochCurtailXval, [4](#)
- stochTable, [6](#)
- Table, [7](#)

Curtail, [2](#)

itemscores, [3](#)

stochCurtail, [3](#)  
stochCurtailXval, [4](#)  
stochTable, [6](#)

Table, [7](#)