Package: TopicTestlet (via r-universe)

June 1, 2026

Title A Topic Testlet Model for Calibrating Testlet Constructed Responses

Version 0.1.0

Description Implements the Topic Testlet Model (TTM) as described by Xiong et al. (2025) <doi:10.1111/jedm.70001>. The package integrates Latent Dirichlet Allocation (LDA) with the Partial Credit Model to account for local item dependence in testlets using latent topics from student textual responses.

Depends R (>= 4.0.0)

Imports topicmodels, tm, stats

License GPL (>= 3)

Encoding UTF-8

RoxygenNote 7.3.2

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author Jiawei Xiong [aut, cre], Cheng Tang [ctb], Qidi Liu [ctb]

Maintainer Jiawei Xiong <jiawei.xiong@uga.edu>

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

Date/Publication 2025-12-01 14:00:02 UTC

RemoteUrl https://github.com/cran/TopicTestlet

RemoteRef HEAD

RemoteSha ad4937f53f37551b245cb135edf9ff9fce5d2adf

Contents

aggregate_responses *Concatenate Student Responses*

Description

Aggregates all written responses within a testlet for each student.

Usage

```
aggregate_responses(response_matrix)
```

Arguments

response_matrix	An N x J matrix of character strings (essays).

Value

A named character vector of length N.

ttm_est *Fit the Topic Testlet Model (TTM)*

Description

Calibrates the TTM using score data and pre-computed topic proportions. Uses a Variational Expectation-Maximization (VEM) approach to estimate student ability (theta), topic penalties (lambda), and item parameters (b).

Usage

```
ttm_est(scores, delta, max_iter = 100, tol = 1e-04)
```

Arguments

scores	An N x J numeric matrix of item scores (0, 1, ...).
delta	An N x K numeric matrix of topic proportions (from ttm_lda).
max_iter	Maximum number of EM iterations.
tol	Convergence tolerance.

Value

A list containing:

`theta`	Vector of estimated student abilities.
`lambda`	Matrix of estimated topic penalties.
`gamma`	Vector of person-specific testlet effects.
`item_params`	List of step difficulties for each item.
`AIC`	Akaike Information Criterion.
`BIC`	Bayesian Information Criterion.

`ttm_lda` *Fit LDA and Extract Topic Proportions*

Description

Fits a Latent Dirichlet Allocation model to the text and returns the person-specific topic proportion matrix (delta).

Usage

```
ttm_lda(text_vector, k, seed = 1234)
```

Arguments

`text_vector`	A character vector of aggregated student responses.
`k`	The number of latent topics.
`seed`	Integer seed for reproducibility.

Value

A matrix of dimension N x K containing topic proportions (delta).

`ttm_perplexity` *Calculate Perplexity for Different Topic Numbers*

Description

Calculates the perplexity of LDA models over a range of K topics to help determine the optimal number of topics.

Usage

```
ttm_perplexity(text_vector, k_range = 2:5, seed = 1234)
```

Arguments

`text_vector`	A character vector of aggregated student responses (length N).
`k_range`	A numeric vector indicating the number of topics to try (e.g., 2:10).
`seed`	Integer seed for reproducibility.

Value

A data frame containing K and the corresponding perplexity score.

Index