

Package ‘retestR’

October 8, 2026

Title Model-Based Anomaly Detection for Repeat Test-Takers

Version 0.1.0

Description Flags repeat test-takers whose second-attempt performance departs from what a growth model predicts, using independent evidence sources: model-expected score gain (accounting for regression to the mean, time between attempts and remediation), differential performance on exposed versus new items (Sinharay, 2017, <[doi:10.3102/1076998616673872](https://doi.org/10.3102/1076998616673872)>), and differential response speed under a lognormal response-time model (van der Linden, 2006, <[doi:10.3102/10769986031002181](https://doi.org/10.3102/10769986031002181)>). Evidence is combined into a risk index calibrated by parametric bootstrap under the no-misconduct model, so flagging thresholds carry explicit false-positive rates.

License MIT + file LICENSE

URL <https://github.com/edidatasolutions/retestR>,
<https://edidatasolutions.github.io/retestR/>

BugReports <https://github.com/edidatasolutions/retestR/issues>

Encoding UTF-8

Depends R (>= 4.1)

Imports stats, utils

Suggests knitr, markdown

VignetteBuilder knitr

Config/roxygen2/version 8.1.0

NeedsCompilation no

Author Daniel Edi [aut, cre, cph] (ORCID:
<<https://orcid.org/0000-0001-5475-819X>>)

Maintainer Daniel Edi <danieledi2026@gmail.com>

Repository CRAN

Date/Publication 2026-10-08 09:10:07 UTC

Contents

rt_data	2
rt_evidence	3
rt_fit	4
rt_raw_gain	4
rt_risk	5
rt_simulate	6
Index	8

rt_data	<i>Assemble two-attempt repeater data</i>
---------	---

Description

Assemble two-attempt repeater data

Usage

rt_data(responses, persons, bank)

Arguments

responses	Long data frame, one row per person x attempt x item: ‘person’, ‘attempt’ (1 or 2), ‘item’, ‘x’ (0/1), and optionally ‘rt’ (response time in seconds).
persons	One row per repeater: ‘person’ plus the covariates used by the growth model (e.g. ‘days_between’, ‘remediation’).
bank	Calibrated item bank: ‘item’, ‘b’ (Rasch difficulty), ‘exposed’ (TRUE for items that may be compromised, e.g. long-running operational items; FALSE for new items), and optionally ‘beta’ (lognormal time intensity, log-seconds).

Details

Assumes no item is administered to the same person on both attempts (legitimate item memory would otherwise look like preknowledge).

Value

An ‘rt_data’ object.

Examples

```
bank <- data.frame(item = paste0("Q", 1:20), b = rnorm(20),
  exposed = rep(c(TRUE, FALSE), each = 10))
persons <- data.frame(person = c("A", "B"), days_between = c(60, 200),
  remediation = c(0, 1))
resp <- data.frame(person = rep(c("A", "B"), each = 20),
  attempt = rep(rep(1:2, each = 10), 2),
```

```

      item = c(paste0("Q", c(1:5, 11:15, 6:10, 16:20)),
               paste0("Q", c(6:10, 16:20, 1:5, 11:15))),
      x = rbinom(40, 1, 0.6))
str(rt_data(resp, persons, bank)$responses)

```

rt_evidence

Evidence statistics for each repeater

Description

Each statistic is a posterior-predictive z-score, positive in the suspicious direction:

‘z_gain’ Attempt-2 total score against the distribution predicted from attempt 1 plus the fitted expected growth for this person’s covariates. A large gain after long study and remediation is expected; the same gain after two weeks is not.

‘z_exposed’ Attempt-2 score on exposed items against the prediction from attempt 1, expected growth and attempt-2 new items. Gains concentrated on exposed items are the signature of preknowledge.

‘z_rt’ Mean log-speed on exposed minus new items (lognormal RT model with known time intensities); present when response times are.

Usage

```
rt_evidence(fit)
```

Arguments

`fit` An ‘rt_fit’.

Value

Data frame: ‘person’, ‘S1’, ‘S2’, ‘z_gain’, ‘z_exposed’, and ‘z_rt’.

Examples

```

sim <- rt_simulate(n_persons = 200, form_exposed = 20, form_new = 10, seed = 1)
fit <- rt_fit(sim)
ev <- rt_evidence(fit)
# preknowledge shows up on exposed items and in speed, not only in the gain
aggregate(ev[c("z_gain", "z_exposed", "z_rt")],
          list(preknowledge = sim$truth$preknowledge), mean)

```

rt_fit	<i>Fit the expected-gain model</i>
--------	------------------------------------

Description

Marginal maximum likelihood on a theta grid, in two stages: (1) the attempt-1 ability distribution of repeaters, $N(\mu_1, s_1)$; (2) growth ' $\theta_2 = \theta_1 + X \beta + N(0, \sigma_{\text{growth}})$ ' from attempt-1 responses and attempt-2 responses to ****new items only****. Because exposed items never enter the growth model, preknowledge cannot inflate the expected gain, and conditioning on the full attempt-1 likelihood handles regression to the mean.

Usage

```
rt_fit(
  data,
  growth = ~log(days_between) + remediation,
  grid = seq(-5, 5, by = 0.2)
)
```

Arguments

data	An 'rt_data' (or 'rt_sim') object.
growth	One-sided formula for mean growth, evaluated in 'data\$persons'.
grid	Theta grid.

Value

An 'rt_fit' object.

Examples

```
sim <- rt_simulate(n_persons = 200, form_exposed = 20, form_new = 10, seed = 1)
fit <- rt_fit(sim)
fit  # growth coefficients: intercept, log(days_between), remediation
```

rt_raw_gain	<i>Raw-gain flagging (the common practice, as a baseline)</i>
-------------	---

Description

Raw-gain flagging (the common practice, as a baseline)

Usage

```
rt_raw_gain(data, threshold = 0.2)
```

Arguments

data	An 'rt_data' or 'rt_sim'.
threshold	Flag gains in proportion correct at or above this value.

Value

Data frame: 'person', 'p1', 'p2', 'gain', 'flag'.

Examples

```
sim <- rt_simulate(n_persons = 200, form_exposed = 20, form_new = 10, seed = 1)
raw <- rt_raw_gain(sim, threshold = 0.2)
# raw gains also flag honest candidates who remediated
table(flagged = raw$flag, remediation = sim$truth$remediation)
```

rt_risk	<i>Calibrated risk index for repeat test-takers</i>
---------	---

Description

Combines evidence statistics into 'T = sum(z)' and calibrates it by parametric bootstrap: 'n_null' complete replicate administrations are simulated under the fitted no-misconduct model (same persons, forms, covariates; attempt-1 ability drawn from each person's posterior; honest growth; honest response times), and the full evidence pipeline is rerun on each. Because the null distribution comes from the same pipeline, the correlation between evidence sources is accounted for, and 'p_value' is an honest false-positive rate for an honest repeater.

Usage

```
rt_risk(fit, evidence = NULL, n_null = 10, alpha = 0.01, seed = NULL)
```

Arguments

fit	An 'rt_fit'.
evidence	Which statistics to combine: any of "gain", "exposed", "rt". The default combines 'exposed' and 'rt' (when response times are available). 'z_gain' is always reported but not combined by default: in known-truth simulations it mostly repeats the exposed-item signal with extra noise, and adding it lowered detection at a fixed false-positive rate.
n_null	Null replicates (each is a full administration).
alpha	Flagging level.
seed	Optional seed.

Value

An 'rt_risk' data frame: evidence columns, 'T', per-component empirical p-values, 'p_value', 'q_value' (Benjamini-Hochberg), 'flag'.

```
sim <- rt_simulate(n_persons = 200, form_exposed = 20, form_new = 10, seed = 1)
fit <- rt_fit(sim)
risk <- rt_risk(fit, n_null = 2, alpha = 0.01, seed = 1)
risk
table(flagged = risk$flag, preknowledge = sim$truth$preknowledge)
```

rt_simulate	<i>Simulate repeat test-takers with known misconduct</i>
-------------	--

Honest repeaters grow by $g_0 + g_1 * \log(\text{days} / 30) + g_2 * \text{remediation}$ plus normal noise. A fraction $p_{\text{preknowledge}}$ obtained a random share (known_frac) of the exposed pool between attempts: on those items they answer correctly with probability known_p and respond speedup log-units faster.

```
rt_simulate(  
  n_persons = 2000,  
  n_exposed_pool = 300,  
  n_new_pool = 200,  
  form_exposed = 40,  
  form_new = 20,  
  theta_mean = -0.5,  
  theta_sd = 0.7,  
  growth = c(0.1, 0.1, 0.4),  
  growth_sd = 0.25,  
  p_remediation = 0.4,  
  p_preknowledge = 0.05,  
  known_frac = 0.6,  
  known_p = 0.95,  
  speedup = 1,  
  rt_sd = 0.5,  
  seed = NULL  
)
```

n_persons	Number of repeaters.
n_exposed_pool, n_new_pool	
	Bank sizes.
form_exposed, form_new	
	Items per form from each pool; forms are disjoint across a person's two attempts.
theta_mean, theta_sd	
	Attempt-1 ability of repeaters.

growth	Coefficients 'c(g0, g1, g2)'.
growth_sd	SD of individual growth.
p_remediation	Share of repeaters who completed remediation.
p_preknowledge	Share with preknowledge at attempt 2.
known_frac, known_p, speedup	Preknowledge strength.
rt_sd	Residual SD of log response time.
seed	Optional seed.

Value

An 'rt_sim': '\$data' (an 'rt_data') and '\$truth' (per-person 'theta1', 'theta2', 'growth_mean', 'preknowledge', 'remediation').

Examples

```
sim <- rt_simulate(n_persons = 200, form_exposed = 20, form_new = 10, seed = 1)
head(sim$truth)
table(sim$truth$preknowledge)
```

Index

rt_data, [2](#)
rt_evidence, [3](#)
rt_fit, [4](#)
rt_raw_gain, [4](#)
rt_risk, [5](#)
rt_simulate, [6](#)