Package: srsbench (via r-universe)

July 10, 2026

Type Package

Title Evaluation Metrics for Spaced Repetition Schedulers

Version 0.1

Description Calibration and discrimination metrics for spaced-repetition memory models. Provides the sample-weighted binned root mean squared error (RMSE(bins)) used to rank schedulers in the open spaced repetition benchmark, together with log loss, the area under the ROC curve, and calibration curves.

License MIT + file LICENSE

Encoding UTF-8

URL https://github.com/chrislongros/srsbench

BugReports https://github.com/chrislongros/srsbench/issues

Suggests tinytest

Config/roxygen2/version 8.0.0

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

Date/Publication 2026-07-09 17:50:17 UTC

RemoteUrl https://github.com/chrislongros/srsbench

RemoteRef HEAD

RemoteSha 087d2524fd8fcb9d3d1d498705b5975105f5ed1b

Contents

calibration_bins 2
log_loss 2
rmse_bins 3
srs_auc 4

Index **5**

calibration_bins *Calibration curve in equal-width bins*

Description

Calibration curve in equal-width bins

Usage

```
calibration_bins(p, y, n = 10)
```

Arguments

p	Numeric vector of predicted recall probabilities in [0, 1].
y	Numeric vector of observed outcomes: 1 recalled, 0 forgotten.
n	Number of equal-width probability bins (default 10).

Value

A data frame with one row per bin: bin index, bin bounds, mean predicted probability, observed recall rate and review count. Empty bins have NA means and a count of 0.

Examples

```
calibration_bins(p = c(0.1, 0.2, 0.8, 0.9), y = c(0, 0, 1, 1), n = 5)
```

log_loss *Log loss (binary cross-entropy)*

Description

Log loss (binary cross-entropy)

Usage

```
log_loss(p, y, weights = NULL, eps = 1e-15)
```

Arguments

p	Numeric vector of predicted recall probabilities in [0, 1].
y	Numeric vector of observed outcomes: 1 recalled, 0 forgotten.
weights	Optional numeric vector of non-negative per-review weights (default 1).
eps	Small value used to clip p away from 0 and 1.

Value

A single non-negative number; lower is better.

Examples

```
log_loss(p = c(0.9, 0.1), y = c(1, 0))
```

`rmse_bins` *RMSE in bins for a spaced repetition scheduler*

Description

Computes the sample-weighted binned root mean squared error used to rank schedulers in the open spaced repetition benchmark. Reviews are grouped into bins along three log-quantised axes – the interval (`elapsed_days`), the review number (`i`) and the number of prior lapses (`lapse`) – then the mean predicted and mean observed recall are compared within each bin and combined, weighting each bin by its number of reviews.

Usage

```
rmse_bins(p, y, elapsed_days, i, lapse, weights = NULL)
```

Arguments

`p`	Numeric vector of predicted recall probabilities in `[0, 1]`.
`y`	Numeric vector of observed outcomes: `1` recalled, `0` forgotten.
`elapsed_days`	Numeric vector of review intervals in days.
`i`	Numeric vector giving each review's position in its card's history (1 for the first review, 2 for the second, and so on).
`lapse`	Numeric vector giving the number of lapses before each review.
`weights`	Optional numeric vector of non-negative per-review weights (default 1).

Value

A single non-negative number; lower is better calibrated.

References

RMSE(bins) is the metric defined by the open spaced repetition benchmark; this is an independent R implementation of its `rmse_matrix` definition. https://github.com/open-spaced-repetition/srs-benchmark

Examples

```
rmse_bins(p = c(0.9, 0.2), y = c(1, 0),
          elapsed_days = c(2, 100), i = c(2, 5), lapse = c(0, 1))
```

srs_auc *Area under the ROC curve*

Description

A rank-based (Mann-Whitney) estimate of discrimination, with no external dependencies. Ties in p are handled by mid-ranks.

Usage

```
srs_auc(p, y)
```

Arguments

p	Numeric vector of predicted recall probabilities in [0, 1].
y	Numeric vector of observed outcomes: 1 recalled, 0 forgotten.

Value

The AUC in [0, 1], or NA if y has only one class.

Examples

```
srs_auc(p = c(0.9, 0.8, 0.2, 0.1), y = c(1, 1, 0, 0))
```

Index

calibration_bins, 2

log_loss, 2

rmse_bins, 3

srs_auc, 4