Package 'sshist'

March 4, 2026

Type Package

Title Optimal Histogram Binning Using Shimazaki-Shinomoto Method

Version 0.1.3

Description Implements the Shimazaki-Shinomoto method for optimizing the bin width of a histogram. This method minimizes the mean integrated squared error (MISE) and features a 'C++' backend for high performance and shift-averaging to remove edge-position bias. Ideally suits for time-dependent rate estimation and identifying intrinsic data structures. Supports both 1D and 2D data distributions. For more details see Shimazaki and Shinomoto (2007) ``A Method for Selecting the Bin Size of a Time Histogram'' <doi:10.1162/neco.2007.19.6.1503>.

License GPL (>= 3)

URL https://github.com/celebithil/sshist,
https://www.neuralengine.org/res/histogram.html

BugReports https://github.com/celebithil/sshist/issues

Imports graphics, grDevices, Rcpp, stats

Suggests ggplot2, knitr, rmarkdown

LinkingTo Rcpp

VignetteBuilder knitr, rmarkdown

Config/testthat/edition 3

Encoding UTF-8

RoxygenNote 7.3.3

NeedsCompilation yes

Author Daniil Popov [aut, cre]

Maintainer Daniil Popov <popov.daniil@gmail.com>

Repository CRAN

Date/Publication 2026-03-04 10:40:02 UTC

Contents

Index   **7**

---

plot.sshist	*Plot method for sshist objects*

Description

Plot method for sshist objects

Usage

```
## S3 method for class 'sshist'
plot(x, ...)
```

Arguments

x	An object of class sshist.
...	Additional arguments passed to plot.

Value

No return value, called for side effects. Creates a two-panel plot showing: (1) the cost function curve with the optimal number of bins highlighted, and (2) the optimal histogram with the selected bin edges.

---

plot.sshist_2d	*Plot method for sshist_2d objects*

Description

Plot method for sshist_2d objects

Usage

```
## S3 method for class 'sshist_2d'
plot(x, ...)
```

Arguments

x	An object of class sshist.
...	Additional arguments passed to plot.

Value

No return value, called for side effects. Creates a two-panel plot showing: (1) a heatmap of the cost function landscape across different bin combinations with the optimal point highlighted in red, and (2) the optimal 2D histogram visualization using the selected bin numbers for both dimensions.

---

print.sshist *Print method for sshist objects*

Description

Print method for sshist objects

Usage

```
## S3 method for class 'sshist'
print(x, ...)
```

Arguments

x	An object of class sshist.
...	Additional arguments passed to print.

Value

Returns the input object x invisibly. The method is called for its side effect of printing a summary of the Shimazaki-Shinomoto histogram optimization results, including the optimal number of bins, bin width, and minimum cost value.

---

print.sshist_2d *Print method for sshist_2d objects*

Description

Print method for sshist_2d objects

Usage

```
## S3 method for class 'sshist_2d'
print(x, ...)
```

Arguments

x	An object of class sshist.
...	Additional arguments passed to print.

Value

Returns the input object x invisibly. The method is called for its side effect of printing a summary of the 2D Shimazaki-Shinomoto histogram optimization results, including the optimal number of bins for both X and Y dimensions, bin widths, and minimum cost value.

---

sshist   *Optimal Histogram Binning (Shimazaki-Shinomoto Method)*

Description

Computes the optimal bin width and number of bins using the Shimazaki-Shinomoto method. This implementation uses Rcpp for high performance and includes shift-averaging to remove edge-position bias.

Usage

```
sshist(x, n_max = NULL, sn = 30)
```

Arguments

x	A numeric vector of data. Missing values (NA) will be removed.
n_max	An integer specifying the maximum number of bins to test. If NULL (default), it is automatically determined based on data resolution and sample size to prevent overfitting.
sn	Integer. The number of shifts to use for averaging (default is 30).

Value

An object of class "sshist" containing:

opt_n The optimal number of bins.

opt_d The optimal bin width.

edges The sequence of break points for the optimal histogram.

cost A numeric vector of the cost function values.

n_tested The vector of N values tested.

data The original data (cleaned).

References

Shimazaki, H. and Shinomoto, S., 2007. A method for selecting the bin size of a time histogram. *Neural Computation*, 19(6), pp.1503-1527.

Examples

```
# 1. Basic usage with base graphics
data(faithful)
res <- sshist(faithful$waiting)
plot(res)

# 2. Usage with ggplot2
if (requireNamespace("ggplot2", quietly = TRUE)) {
  library(ggplot2)
  df <- data.frame(waiting = faithful$waiting)

  ggplot(df, aes(x = waiting)) +
    geom_histogram(breaks = res$edges, fill = "lightblue", color = "black") +
    ggtitle(paste("Optimal Shimazaki-Shinomoto Bins (N =", res$opt_n, ")")) +
    theme_minimal()
}
```

---

sshist_2d   *Optimal 2D Histogram Binning (Shimazaki-Shinomoto Method)*

Description

Computes the optimal number of bins for a 2-dimensional histogram. Includes checks for data resolution and biased variance calculation.

Usage

```
sshist_2d(x, y = NULL, n_min = 2, n_max = 200)
```

Arguments

x	Numeric vector (coord X) or a 2-column matrix.
y	Numeric vector (coord Y). Optional if x is a matrix.
n_min	Integer. Minimum number of bins (default 2).
n_max	Integer. Maximum number of bins (default 200). Algorithm will reduce this automatically if data resolution restricts it.

Value

An object of class "sshist_2d" containing optimal parameters.

Examples

```
# 1. Basic usage with base graphics
set.seed(42)
x <- rnorm(500)
y <- rnorm(500)
res <- sshist_2d(x, y)
```

```
plot(res)

# 2. Usage with ggplot2
if (requireNamespace("ggplot2", quietly = TRUE)) {
  library(ggplot2)
  df <- data.frame(x = x, y = y)

  ggplot(df, aes(x = x, y = y)) +
    geom_bin2d(bins = c(res$opt_nx, res$opt_ny)) +
    scale_fill_viridis_c() +
    ggtitle(paste0("Optimal 2D Bins: ", res$opt_nx, "x", res$opt_ny)) +
    theme_minimal()
}
```

Index