Package ‘MonotonicityTest’

December 18, 2024

Type Package

Title Nonparametric Bootstrap Test for Regression Monotonicity

Version 1.0

Author Dylan Huynh

Maintainer Dylan Huynh <dylanhuynh@utexas.edu>

Description Implements nonparametric bootstrap tests for detecting monotonicity in regression functions from Hall, P. and Heckman, N. (2000) <doi:10.1214/aos/1016120363> Includes tools for visualizing results using Nadaraya-Watson kernel regression and supports efficient computation with C++.

License GPL

Encoding UTF-8

RoxygenNote 7.3.2

LinkingTo Rcpp, RcppEigen

Imports Rcpp (>= 1.0.13-1), parallel, stats, graphics, grDevices

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation yes

Contents

`create_kernel_plot` *Generate Kernel Plot*

Description

Creates a scatter plot of the input vectors ‘X‘ and ‘Y‘, and overlays a Nadaraya-Watson kernel regression curve using the specified bandwidth.

Usage

```
create_kernel_plot(X, Y, bandwidth = bw.nrd(X) * (length(X)^-0.1))
```

Arguments

X	Vector of x values.
Y	Vector of y values.
bandwidth	Kernel bandwidth used for the Nadaraya-Watson estimator. Default is calculated as `bw.nrd(X) * (length(X) ^ -0.1)`.

Value

A recorded plot object containing the scatter plot with the kernel regression curve.

References

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and Its Applications*, **9**(1), 141–142.

Watson, G. S. (1964). Smooth estimates of regression functions. *Sankhyā: The Indian Journal of Statistics, Series A*, 359-372.

---

monotonicity_test *Perform Monotonicity Test*

Description

Performs a monotonicity test between the vectors X and Y as described in Hall and Heckman (2000). This function uses a bootstrap approach to test for monotonicity in a nonparametric regression setting.

Usage

```
monotonicity_test(
  X,
  Y,
  bandwidth = bw.nrd(X) * (length(X)^-0.1),
  boot_num = 200,
  m = floor(0.05 * length(X)),
  ncores = 1,
  negative = FALSE
)
```

Arguments

X	Numeric vector of predictor variable values. Must not contain missing or infinite values.
Y	Numeric vector of response variable values. Must not contain missing or infinite values.
bandwidth	Numeric value for the kernel bandwidth. Default is calculated as `bw.nrd(X) * (length(X) ^ -0.1)`.
boot_num	Integer specifying the number of bootstrap samples. Default is 200.

`m` Integer parameter used in the calculation of the test statistic. Corresponds to the minimum window size to calculate the test statistic over. Default is `floor(0.05) * length(X)`.

`ncores` Integer specifying the number of cores to use for parallel processing. Default is `1`.

`negative` Logical value indicating whether to test for a monotonic decreasing (negative) relationship. Default is `FALSE`.

Details

The test evaluates the following hypotheses:

H_0: The regression function is monotonic

- *Non-decreasing* if `negative` = `FALSE`
- *Non-increasing* if `negative` = `TRUE`

H_A: The regression function is not monotonic

Value

A list with the following components:

`p` The p-value of the test.

`dist` The distribution of test statistics from bootstrap samples.

`stat` The test statistic calculated from the original data.

Note

For large datasets (e.g., $n \geq 6500$) this function may require significant computation time due to having to compute the statistic for every possible interval. Consider reducing `boot_num`, using a subset of the data, or using parallel processing with `ncores` to improve performance.

References

Hall, P., & Heckman, N. E. (2000). Testing for monotonicity of a regression mean by calibrating for linear functions. *The Annals of Statistics*, **28**(1), 20–39.

Examples

```
# Generate sample data
X <- runif(100)
Y <- X^2 + rnorm(100, sd = 0.1)

# Perform monotonicity test
result <- monotonicity_test(X, Y, boot_num=1)
print(result$p)  # Display the p-value
```

Index