Package 'TailID'

September 14, 2025

Title Detect Sensitive Points in the Tail

Version 1.0.0

Description

The goal of 'TailID' is to detect sensitive points in the tail of a dataset using techniques from Extreme Value Theory (EVT). It utilizes the Generalized Pareto Distribution (GPD) for assessing tail behavior and detecting inconsistent points with the Identical Distribution hypothesis of the tail. For more details see Manau (2025)<doi:10.4230/LIPIcs.ECRTS.2025.20>.

License GPL (>= 3)

Encoding UTF-8

RoxygenNote 7.3.3

Imports ggplot2, grDevices, ismev, scales

Language en-US

NeedsCompilation no

Author Blau Manau [aut, cre] (ORCID: <https://orcid.org/0009-0007-7227-4448>)

Maintainer Blau Manau <blau.manau@bsc.es>

Repository CRAN

Date/Publication 2025-09-14 16:00:02 UTC

Contents

---

candidate_selection   *Selects the candidates of the tail*

Description

This function selects the candidates of the tail that can be inconsistent to the ID hypothesis

Usage

```
candidate_selection(sample, pc_max, pc_min)
```

Arguments

sample	A numeric vector.
pc_max	A number between pm_max and 1 indicating the threshold of maximum sensitive points to consider.
pc_min	A number between pm_min and 1 indicating the threshold of minimum sensitive points to consider.

Value

A vector of indices corresponding to the detected sensitive points.

Examples

```
candidate_selection(rnorm(1000), 0.99, 0.99)
candidate_selection(c(rnorm(10^3,10,1),rnorm(10,20,3)), 0.9, 0.9)
```

CI_shapeGPD   *Computes a Confidence Interval for a GPD shape*

Description

This function computes a confidence interval for a GPD shape.

Usage

```
CI_shapeGPD(sample, threshold, parameter, conf_level)
```

Arguments

sample	A numeric vector.
threshold	A number between 0 and 1 indicating the threshold of extreme values to consider.
parameter	A number indicating the shape value.
conf_level	A number between 0 and 1 indicating the confidence level for the detection.

Value

A Confidence Interval vector.

Examples

```
CI_shapeGPD(rnorm(1000), 0.8, 1, 0.95)
CI_shapeGPD(c(rnorm(10^3,10,1),rnorm(10,20,3)), 0.8, 12, 0.9999)
```

---

plot_TailID	*Saves the plots corresponding to the TailID detection*

Description

This function saves the plots corresponding to the TailID detection, which includes: targeted candidates plot, shape variation plot, and inconsistent detected points.

Usage

```
plot_TailID(output_dir, sample, pm_max, pm_min, pc_max, pc_min, conf_level)
```

Arguments

output_dir	Path to save the plots.
sample	A numeric vector.
pm_max	A number between 0 and 1 indicating the threshold of maximum extreme values to consider.
pm_min	A number between 0 and 1 indicating the threshold of minimum extreme values to consider.
pc_max	A number between pm_max and 1 indicating the threshold of maximum sensitive points to consider.
pc_min	A number between pm_min and 1 indicating the threshold of minimum sensitive points to consider.
conf_level	A number between 0 and 1 indicating the confidence level for the detection.

Value

A vector of indices corresponding to the detected sensitive points.

Examples

```
output_dir <- file.path(tempdir(), "output")
if (dir.exists(output_dir) || dir.create(output_dir, recursive = TRUE)) {
  plot_TailID(output_dir, rnorm(1000), 0.85, 0.85, 0.999, 0.999, 0.95)
}
if (dir.exists(output_dir) || dir.create(output_dir, recursive = TRUE)) {
 plot_TailID(output_dir, c(rnorm(10^3, 10, 1), rnorm(10, 20, 3)), 0.85, 0.85, 0.99, 0.99, 0.99999)
}
```

---

shape_evaluation	*Evaluation of the shape parameter to return inconsistent points and shape parameters computed and its confidence intervals*

Description

This function detects the points of the tail that are inconsistent with the ID hypothesis by evaluation the shape variation of the GPD, and also returns the shape parameters computed and its confidence intervals

Usage

```
shape_evaluation(sample, candidates, pm_max, pm_min, conf_level)
```

Arguments

sample	A numeric vector.
candidates	A list of indices of the sample.
pm_max	A number between 0 and 1 indicating the threshold of maximum extreme values to consider.
pm_min	A number between 0 and 1 indicating the threshold of maximum extreme values to consider.
conf_level	A number between 0 and 1 indicating the confidence level for the detection.

Value

A vector of indices corresponding to the detected sensitive points.

Examples

```
shape_evaluation(rnorm(1000),candidate_selection(rnorm(1000), 0.99, 0.99),
 0.8, 0.8, 0.95)
shape_evaluation(c(rnorm(10^3,10,1),rnorm(10,20,3)),
candidate_selection(c(rnorm(10^3,10,1),rnorm(10,20,3)), 0.9, 0.9),
0.8, 0.8, 0.9999)
```

TailID	*Detects the sensitive points in the Tail*

Description

This function returns the points of the tail that are inconsistent with the ID hypothesis.

Usage

```
TailID(sample, pm_max, pm_min, pc_max, pc_min, conf_level)
```

Arguments

`sample`	A numeric vector.
`pm_max`	A number between 0 and 1 indicating the threshold of maximum extreme values to consider.
`pm_min`	A number between 0 and 1 indicating the threshold of minimum extreme values to consider.
`pc_max`	A number between pm_max and 1 indicating the threshold of maximum sensitive points to consider.
`pc_min`	A number between pm_min and 1 indicating the threshold of minimum sensitive points to consider.
`conf_level`	A number between 0 and 1 indicating the confidence level for the detection.

Value

A vector of indices corresponding to the detected sensitive points.

Examples

```
TailID(rnorm(1000), 0.8, 0.8, 0.99, 0.99, 0.95)
TailID(c(rnorm(10^3,10,1),rnorm(10,20,3)), 0.8, 0.8, 0.9, 0.9, 0.9999)
```

Index