Package ‘tailplots’

May 8, 2026

Title Estimators and Plots for Gamma and Pareto Tail Detection

Version 0.1.1

Author Bernhard Klar [aut, cre],
Lucas Iglesias [aut]

Maintainer Bernhard Klar <Bernhard.Klar@kit.edu>

Description Estimators for two functionals used to detect Gamma, Pareto or Lognormal distributions, as well as distributions exhibiting similar tail behavior, as introduced by Iwashita and Klar (2023) <doi:10.1111/stan.12316> and Klar (2024) <doi:10.1080/00031305.2024.2413081>.
One of these functionals, g, originally proposed by Asmussen and Lehtomaa (2017) <doi:10.3390/risks5010010>, distinguishes between log-convex and log-concave tail behavior.
Furthermore the characterization of the lognormal distribution is based on the work of Mosimann (1970) <doi:10.2307/2284599>.
The package also includes methods for visualizing these estimators and their associated confidence intervals across various threshold values.

Encoding UTF-8

RoxygenNote 7.3.2

Imports Rcpp

LinkingTo Rcpp

Suggests actuar

Config/testthat/edition 3

NeedsCompilation yes

Repository CRAN

Date/Publication 2025-09-08 14:00:03 UTC

Contents

`gamma_tail` *Estimate of tail functional g and confidence intervals for g and alpha*

Description

This function computes the estimate of g and the associated confidence interval for g as well as $alpha$, the corresponding shape parameter under the assumption of a gamma model, according to Iwashita and Klar (2024). Three methods are implemented to compute the confidence intervals: a method based on the unbiased variance estimators of the underlying U-statistics, and two resampling methods (jackknife and bootstrap).

Usage

```
gamma_tail(
  x,
  d,
  confint = FALSE,
  method = c("unbiased", "bootstrap", "jackknife"),
  R = 1000,
  conf.level = 0.95
)
```

Arguments

`x`	a vector containing the sample data.
`d`	the threshold for the computation of g.
`confint`	a boolean value indicating whether a confidence interval should be computed.
`method`	the method used for computing the confidence intervals (options include unbiased variance estimator, jackknife, and bootstrap).
`R`	the number of the bootstrap replicates.
`conf.level`	the confidence level for the interval.

Details

The function g introduced by Asmussen and Lehtomaa (2017) is used to distinguish between log-concave and log-convex tail behavior. It is defined as:

$$g(d) = E\left[\frac{|X_1 - X_2|}{X_1 + X_2}\middle| X_1 + X_2 > d\right]$$

where X_1, X_2 are independent and identically distributed (i.i.d.) positive random variables. For gamma distributions, g takes a constant value, making it a useful tool for detecting gamma-tailed distributions.

This function estimates $g(d)$ using U-statistics. The estimator $\hat{g}(d)$ is given by:

$$\hat{g}(d) = \frac{U_n^{(1)}(d)}{U_n^{(2)}(d)}, \quad d > 0,$$

where

$$U_n^{(1)}(d) = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \frac{|X_i - X_j|}{X_i + X_j} 1(X_i + X_j > d),$$

$$U_n^{(2)}(d) = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} 1(X_i + X_j > d).$$

Confidence intervals for $g(d)$, based on the following variance estimation methods, are also provided:

- Unbiased Variance Estimator
- Bootstrap Resampling
- Jackknife Resampling

The $(1-\gamma)$ confidence interval for $\hat{g}_n(d)$ is given by:

$$\left[\max\left\{ \hat{g}_n(d) \; - \; z_{1-\gamma/2} \, \frac{\hat{\sigma}_d}{\sqrt{n\, U_n^{(2)}(d)}}, \; 0 \right\}, \;\; \min\left\{ \hat{g}_n(d) \; + \; z_{1-\gamma/2} \, \frac{\hat{\sigma}_d}{\sqrt{n\, U_n^{(2)}(d)}}, \; 1 \right\} \right].$$

Here, $z_{1-\gamma/2} = \Phi^{-1}(1 - \frac{\gamma}{2})$ is the appropriate quantile of the standard normal distribution and $\hat{\sigma}_d$ is an estimator of the standard deviation based on one of the methods above.

Value

A matrix containing:

`threshold`	The value of the threshold d.
`g.estimate`	Estimate of g.
`g.ci1`	The lower bound of the confidence interval for g (if `confint = TRUE`).
`g.ci2`	The upper bound of the confidence interval for g (if `confint = TRUE`).
`alpha`	Estimate of the shape parameter under a gamma model.
`alpha.ci1`	The lower bound of the confidence interval for alpha (if `confint = TRUE`).
`alpha.ci2`	The upper bound of the confidence interval for alpha (if `confint = TRUE`).

References

Iwashita, T. & Klar, B. (2024). A gamma tail statistic and its asymptotics. Statistica Neerlandica 78:2, 264-280. doi:10.1111/stan.12316

Asmussen, S. & Lehtomaa, J. (2017). Distinguishing Log-Concavity from Heavy Tails. *Risks 2017, 5, 10*. doi:10.3390/risks5010010

Examples

```
x <- rgamma(100, shape = 2, scale = 1)
gamma_tail(x, d = 2, confint = FALSE, method = "unbiased", R = 1000)
```

---

gamma_tailplot *Plot the estimated g and the corresponding confidence intervals*

Description

This function produces a tail plot for the estimate $\hat{g}$ over a range of thresholds for a given sample, including confidence intervals computed by one of three methods (unbiased, bootstrap or jackknife). The function also allows a choice between original and log scale.

Usage

```
gamma_tailplot(
  x,
  method = c("unbiased", "bootstrap", "jackknife"),
  R = 1000,
  conf.level = 0.95,
  ci.points = 101,
  xscale = "o"
)
```

Arguments

`x`	a vector containing the sample data.
`method`	the method used for computing the confidence intervals (options include unbiased variance estimator, jackknife, and bootstrap).
`R`	the number of the bootstrap replicates.
`conf.level`	the confidence level for the interval.
`ci.points`	the number of thresholds used in the calculation of the confidence intervals.
`xscale`	the scale of the x-axis (options include "o" = original, "l" = log scale, "b" = both).

Details

For more details about the estimator $\hat{g}$ and the computation of the confidence intervals see gamma_tail.

Value

A plot showing the estimated $g(d)$ versus threshold d, optionally on a logarithmic x-axis and including confidence intervals.

References

Iwashita, T. & Klar, B. (2024). A gamma tail statistic and its asymptotics. Statistica Neerlandica 78:2, 264-280. doi:10.1111/stan.12316

Examples

```
x <- rgamma(2e2, 0.5, 0.2)
gamma_tailplot(x, method="unbiased", xscale="o")
```

---

lnorm_tail — *Estimate of tail functional s and confidence intervals for s and sigma*

Description

This function computes the estimate of s and the associated confidence interval for s as well as the standard deviation sigma on the log-scale of the lognormal distribution. Three methods are implemented to compute the confidence intervals: a method based on the unbiased variance estimators of the underlying U-statistics and two resampling methods (jackknife and bootstrap).

Usage

```
lnorm_tail(
  x,
  u,
  confint = FALSE,
  method = c("unbiased", "bootstrap", "jackknife"),
  R = 1000,
  conf.level = 0.95
)
```

Arguments

`x`	a vector containing the sample data.
`u`	the threshold for the computation of s.
`confint`	a boolean value indicating whether the confidence interval should be computed.
`method`	the method used for computing the confidence intervals (options include unbiased variance estimator, jackknife, and bootstrap).
`R`	the number of the bootstrap replicates.
`conf.level`	the confidence level for the interval.

Details

The function s, defined by

$$s(u) = \mathbb{E}\left[\frac{|X_1 - X_2|}{X_1 + X_2} \;\middle|\; X_1\, X_2 > u \right],$$

where X_1, X_2 are independent and identically distributed (i.i.d.) positive random variables, takes a constant value if and only if X_1 follows a lognormal distribution. Thus, s can be used to detect distributions with lognormal tails. The characterization of the lognormal distribution is based on the work of Mosimann (1970). This function estimates $s(u)$ using U-statistics, similarly as in Iwashita and Klar (2024).

Value

A matrix containing:

`threshold`	The value of the threshold u.
`s.estimate`	Estimate of the tail functional s.
`s.ci1`	The lower bound of the confidence interval for s (if `confint = TRUE`).
`s.ci2`	The upper bound of the confidence interval for s (if `confint = TRUE`).
`sigma`	Estimate of the scale parameter under a lognormal model.
`sigma.ci1`	The lower bound of the confidence interval for sigma (if `confint = TRUE`).
`sigma.ci2`	The upper bound of the confidence interval for sigma (if `confint = TRUE`).

References

Mosimann, J. E. (1970). Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. Journal of the American Statistical Association, 65(330):930–945. doi:10.2307/2284599

Iwashita, T. & Klar, B. (2024). A gamma tail statistic and its asymptotics. Statistica Neerlandica 78:2, 264-280. doi:10.1111/stan.12316

Examples

```
x = rlnorm(1e3, 2, 2)
u = round( quantile(x, 0.98) )
lnorm_tail(x, u, confint = FALSE)
```

`lnorm_tailplot` *Plot the estimated s and the corresponding confidence intervals*

Description

This function produces a tail plot for the estimate $\hat{s}$ over a range of thresholds for a given sample, including confidence intervals computed by one of three methods (unbiased, bootstrap or jackknife). The function also allows a choice between original and log scale.

Usage

```
lnorm_tailplot(
  x,
  method = c("unbiased", "bootstrap", "jackknife"),
  R = 1000,
  conf.level = 0.95,
  ci.points = 101,
  xscale = "o"
)
```

Arguments

`x`	a vector containing the sample data.
`method`	the method used for computing the confidence intervals (options include unbiased variance estimator, jackknife, and bootstrap).
`R`	the number of the bootstrap replicates.
`conf.level`	the confidence level for the interval.
`ci.points`	the number of thresholds used in the calculation of the confidence intervals.
`xscale`	the scale of the x-axis (options include "o" = original, "l" = log scale, "b" = both).

Value

A plot showing the estimated $s(u)$ versus threshold u, optionally on a logarithmic x-axis and including confidence intervals. Note that on the right side of the plot, one can observe the corresponding sigma values, which indicate the the standard deviation on the log-scale of the lognormal distribution associated with the estimated s-values.

References

Mosimann, J. E. (1970). Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. Journal of the American Statistical Association, 65(330):930–945. doi:10.2307/2284599

Iwashita, T. & Klar, B. (2024). A gamma tail statistic and its asymptotics. Statistica Neerlandica 78:2, 264-280. doi:10.1111/stan.12316

Examples

```
x = rlnorm(2e2, 2, 2)
lnorm_tailplot(x, method="unbiased", xscale="o")
```

`pareto_tail` *Estimate of tail functional t and confidence intervals for t and alpha*

Description

This function computes the estimate of t and the associated confidence interval for t as well as $alpha$, the corresponding shape parameter under the assumption of a Pareto model according to Klar (2024). Three methods are implemented to compute the confidence intervals: a method based on the unbiased variance estimators of the underlying U-statistics and two resampling methods (jackknife and bootstrap).

Usage

```
pareto_tail(
  x,
  u,
  confint = FALSE,
  method = c("unbiased", "bootstrap", "jackknife"),
  R = 1000,
  conf.level = 0.95
)
```

Arguments

`x`	a vector containing the sample data.
`u`	the threshold for the computation of t.
`confint`	a boolean value indicating whether the confidence interval should be computed.
`method`	the method used for computing the confidence intervals (options include unbiased variance estimator, jackknife, and bootstrap).
`R`	the number of the bootstrap replicates.
`conf.level`	the confidence level for the interval.

Details

In Klar (2024) the function

$$t_X(u) \;=\; \mathbb{E}\left[\frac{|X_1 - X_2|}{X_1 + X_2} \;\middle|\; \min\{X_1, X_2\} \geq u\right]$$

is proposed as a tool for detecting Pareto-type tails, where X_1, X_2, X are $i.i.d.$ random variables from an absolutely continuous distribution supported on $[x_m, \infty)$. Theorem 1 in Klar (2024) shows that $t_X(u)$ is constant in u if and only if X has a Pareto distribution.

The estimator $\hat{t}_n\big(X_{(k)}\big)$ can be computed recursively. For $k = 2, \ldots, n-1$,

$$\hat{t}_n\big(X_{(k)}\big) \;=\; \frac{n-k+2}{n-k}\,\hat{t}_n\big(X_{(k-1)}\big) \;-\; \frac{1}{\binom{n-k+1}{2}} \sum_{j=k}^{n} \frac{X_{(j)} - X_{(k-1)}}{X_{(j)} + X_{(k-1)}}\,,$$

which can be evaluated efficiently starting from $\hat{t}_n\big(X_{(n-1)}\big) = \big(X_{(n)} - X_{(n-1)}\big)/\big(X_{(n)} + X_{(n-1)}\big)$, where $X_{(k)}$ denotes the k-th order statistic.

Confidence intervals for $t(u)$ based on the following methods for variance estimation are also provided:

- Unbiased variance estimator
- Bootstrap resampling
- Jackknife resampling

A two-sided $(1-\gamma)$ confidence interval for the estimator $\hat{t}_n(u)$ is :

$$\left[\max\left\{\hat{t}_n(u) \;-\; z_{1-\frac{\gamma}{2}}\,\frac{\hat{\sigma}_u}{\sqrt{n\,U_n^{(2)}(u)}},\; 0\right\},\, \min\left\{\hat{t}_n(u) \;+\; z_{1-\frac{\gamma}{2}}\,\frac{\hat{\sigma}_u}{\sqrt{n\,U_n^{(2)}(u)}},\; 1\right\}\right],$$

where $z_{1-\frac{\gamma}{2}} = \Phi^{-1}(1-\frac{\gamma}{2})$ is the appropriate quantile of the standard normal distribution, $\hat{\sigma}_u$ is an estimator of the standard deviation of $c\,\hat{t}_n(u)$, for a constant c specified in section 4.1. of Klar (2024), and $U_n^{(2)}(u)$ is a U-statistic given by

$$U_n^{(2)}(u) \;=\; \frac{2}{n\,(n-1)} \sum_{i=1}^{n} (n-i) 1\{X_{(i)} \;\geq\; u\}.$$

Value

A matrix containing:

`threshold`	The value of the threshold u.
`t.estimate`	Estimate of the tail functional t.
`t.ci1`	The lower bound of the confidence interval for t (if `confint = TRUE`).
`t.ci2`	The upper bound of the confidence interval for t (if `confint = TRUE`).
`alpha`	Estimate of the shape parameter under a Pareto model.
`alpha.ci1`	The lower bound of the confidence interval for alpha (if `confint = TRUE`).
`alpha.ci2`	The upper bound of the confidence interval for alpha (if `confint = TRUE`).

References

Klar, B. (2024). A Pareto tail plot without moment restrictions. *The American Statistician*. doi:10.1080/00031305.2024.2413081

Examples

```
x <- actuar::rpareto1(1e3, shape=1, min=1)
pareto_tail(x, round( quantile(x, c(0.1, 0.5, 0.75, 0.9, 0.95, 0.99)) ), confint = FALSE)
```

`pareto_tailplot` *Plot the estimated t and the corresponding confidence intervals*

Description

This function produces a tail plot for the estimate $\hat{t}$ over a range of thresholds for a given sample, including confidence intervals computed by one of three methods (unbiased, bootstrap or jackknife). The function also allows a choice between original and log scale.

Usage

```
pareto_tailplot(
  x,
  method = c("unbiased", "bootstrap", "jackknife"),
  R = 1000,
  conf.level = 0.95,
  ci.points = 101,
  xscale = "b"
)
```

Arguments

`x`	a vector containing the sample data.
`method`	the method used for computing the confidence intervals (options include unbiased variance estimator, jackknife, and bootstrap).
`R`	the number of the bootstrap replicates.
`conf.level`	the confidence level for the interval.
`ci.points`	the number of thresholds used in the calculation of the confidence intervals.
`xscale`	the scale of the x-axis (options include "o" = original, "l" = log scale, "b" = both).

Details

For more details about the estimator $\hat{t}$ and the computation of the confidence intervals see pareto_tail.

Value

A plot showing the estimated $t(u)$ versus threshold u, optionally on a logarithmic x-axis and including confidence intervals. Note that on the right side of the plot, one can observe the corresponding alpha values, which indicate the shape parameter of the Pareto distribution associated with the estimated t-values.

References

Klar, B. (2024). A Pareto tail plot without moment restrictions. *The American Statistician*. doi:10.1080/00031305.2024.2413081

Examples

```
x <- actuar::rpareto1(1e3, shape=1, min=1)
pareto_tailplot(x, method="unbiased", xscale="o")
```

Index