Package: logcumulant (via r-universe)

June 13, 2026

Type Package

Title Goodness-of-Fit Tests and Diagrams Based on Mellin Log-Cumulants

Version 0.1.0

Description A family of three complementary goodness-of-fit tests based on an adaptation of Hotelling's T-squared statistic applied to vectors of sample log-cumulants (Mellin statistics) for positive-support reliability data. The package provides the asymptotic chi-squared reference and parametric bootstrap p-values for reliable finite-sample inference, covering the Weibull, Frechet, Gamma, Inverse-Gamma, Log-Normal, and Log-Logistic families. It also provides three diagnostic diagrams (log-cumulant, kurtosis-skewness, and coefficient-of-variation) with bootstrap concentration ellipses, in the spirit of moment-ratio diagrams. Methods are described in Santos, Ospina, Espinheira and Oliveira (2025).

License GPL-3

Encoding UTF-8

LazyData true

Depends R (>= 4.0)

Imports Rcpp, MASS, VGAM, actuar, numDeriv, goftest, ggplot2, gridExtra

LinkingTo Rcpp, RcppArmadillo

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

URL https://github.com/raydonal/logcumulant, https://raydonal.github.io/logcumulant/

BugReports https://github.com/raydonal/logcumulant/issues

Roxygen list(markdown = FALSE)

Config/roxygen2/version 8.0.0

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

Date/Publication 2026-06-06 15:26:48 UTC

RemoteUrl https://github.com/raydonal/logcumulant

RemoteRef HEAD

RemoteSha de9cc60ed9ef167e00b9e2603a730d6c7cbec298

Contents

`ad_cvm_test` *Anderson-Darling and Cramer-von Mises tests*

Description

Computes the Anderson-Darling (AD) and Cramer-von Mises (CvM) statistics and their p-values for a fitted distribution, based on the probability integral transform.

Usage

```
ad_cvm_test(x, dist, theta)
```

Arguments

x	Numeric vector of positive observations.
dist	Character; distribution name.
theta	Numeric length-2 parameter vector (typically MLE).

Value

A list with AD, AD_p, CvM, CvM_p.

Examples

```
set.seed(1); x <- rdist(100, "Weibull", c(2, 1))
ad_cvm_test(x, "Weibull", c(2, 1))
```

---

aic_value *Akaike information criterion for a fitted family*

Description

Akaike information criterion for a fitted family

Usage

```
aic_value(x, dist, fit)
```

Arguments

x	Numeric vector of positive observations.
dist	Character; distribution name.
fit	A mle_fit object for dist.

Value

Numeric AIC value.

Examples

```
set.seed(1); x <- rdist(100, "Gamma", c(3, 0.5))
aic_value(x, "Gamma", mle_fit(x, "Gamma"))
```

cv_diagram *Coefficient-of-variation diagram*

Description

Draws the coefficient-of-variation diagnostic diagram on the original scale (CV γ_2 versus skewness γ_3) with theoretical loci, bootstrap cloud, and 95% concentration ellipse.

Usage

```
cv_diagram(
  data,
  data_name = "Dataset",
  B = NULL,
  seed = 42,
  level = 0.95,
  xlim = c(0, 2.2),
  ylim = c(-0.2, 5)
)
```

Arguments

data	Numeric vector of positive observations.
data_name	Character; label used in the title.
B	Integer; bootstrap replicates (default chosen adaptively from n).
seed	Integer random seed.
level	Numeric; ellipse confidence level (default 0.95).
xlim, ylim	Numeric length-2 axis limits.

Value

A ggplot object.

Examples

```
data(reliability_datasets)
cv_diagram(reliability_datasets$Yarn, "Yarn", B = 200)
```

distribution_dispatchers *Distribution dispatchers for the six reliability families*

Description

Unified density, random-number, and distribution-function interfaces for the six positive-support families supported by the package: Weibull, Frechet, Gamma, Inverse-Gamma, Log-Normal, and Log-Logistic. The two-parameter vector `theta = c(par1, par2)` is interpreted as `(shape, scale)` for all families except Log-Normal, where it is `(meanlog, sdlog)`.

Usage

```
ldist(x, dist, theta, log = TRUE)

rdist(n, dist, theta)

pdist(q, dist, theta)
```

Arguments

`x`, `q`	Numeric vector of quantiles (positive support).
`dist`	Character; one of `"Weibull"`, `"Frechet"`, `"Gamma"`, `"InvGamma"`, `"LogNormal"`, `"LogLogistic"`.
`theta`	Numeric length-2 parameter vector (see Details).
`log`	Logical; if `TRUE` (default) `ldist` returns log-density.
`n`	Integer; number of random values to draw.

Value

`ldist` returns the (log-)density, `rdist` a random sample of length `n`, and `pdist` the cumulative distribution function, each evaluated at the supplied points.

Examples

```
set.seed(1)
x <- rdist(100, "Weibull", c(2, 1))
head(ldist(x, "Weibull", c(2, 1)))
pdist(1, "Gamma", c(3, 0.5))
```

fisher_closed *Closed-form Fisher information matrix*

Description

Per-sample (unit) Fisher information matrix for the supported families. The Weibull/Frechet matrix uses the corrected positive-definite form derived in the methodology.

Usage

```
fisher_closed(dist, theta)
```

Arguments

dist	Character; distribution name.
theta	Numeric length-2 parameter vector.

Value

A 2 by 2 Fisher information matrix.

Examples

```
fisher_closed("Weibull", c(2, 1))
```

gof_analyze *Full goodness-of-fit analysis for one distribution*

Description

Fits a single family and returns the three T^2 statistics (with asymptotic and, optionally, bootstrap p-values), the AD and CvM tests, and the AIC, in a single row.

Usage

```
gof_analyze(x, dist, use_bootstrap = FALSE, B = NULL, seed = NULL)
```

Arguments

x	Numeric vector of positive observations.
dist	Character; distribution name.
use_bootstrap	Logical; compute bootstrap p-values.
B	Integer; bootstrap replicates.
seed	Optional integer random seed.

Value

A one-row `data.frame` of statistics and p-values.

Examples

```
set.seed(1); x <- rdist(100, "Weibull", c(2, 1))
gof_analyze(x, "Weibull")
```

`gof_compare_all` *Compare all candidate distributions*

Description

Runs `gof_analyze` across all six (or a chosen subset of) families and returns a comparison table, the natural entry point for model selection.

Usage

```
gof_compare_all(
  x,
  dists = .LC_DISTS,
  use_bootstrap = FALSE,
  B = NULL,
  seed = NULL
)
```

Arguments

`x`	Numeric vector of positive observations.
`dists`	Character vector of distribution names to compare.
`use_bootstrap`	Logical; compute bootstrap p-values.
`B`	Integer; bootstrap replicates.
`seed`	Optional integer random seed.

Value

A `data.frame` with one row per distribution.

Examples

```
set.seed(1); x <- rdist(100, "Weibull", c(2, 1))
gof_compare_all(x)
```

jacobian_J *Analytic Jacobian of log-cumulants with respect to parameters*

Description

Returns the analytic Jacobian $\mathbf{J}_{\mathcal{V}} = \partial\kappa_{\mathcal{V}}/\partial\theta$ for the selected set of cumulant orders, used in the construction of the T^2 statistics.

Usage

```
jacobian_J(dist, theta, V)
```

Arguments

`dist`	Character; distribution name.
`theta`	Numeric length-2 parameter vector.
`V`	Integer vector of cumulant orders (e.g. `c(2,3)`).

Value

A `length(V)` by 2 numeric matrix.

Examples

```
jacobian_J("Weibull", c(2, 1), V = c(2, 3))
```

kurtosis_diagram *Kurtosis-skewness diagram*

Description

Draws the kurtosis-skewness diagnostic diagram on the original scale (skewness γ_3 versus excess kurtosis γ_4), including the feasible-region boundary $\gamma_4 = \gamma_3^2 - 2$, theoretical loci, bootstrap cloud, and 95% concentration ellipse.

Usage

```
kurtosis_diagram(
  data,
  data_name = "Dataset",
  B = NULL,
  seed = 42,
  level = 0.95,
  xlim = c(-1.5, 4),
  ylim = c(-3, 16)
)
```

Arguments

data	Numeric vector of positive observations.
data_name	Character; label used in the title.
B	Integer; bootstrap replicates (default chosen adaptively from n).
seed	Integer random seed.
level	Numeric; ellipse confidence level (default 0.95).
xlim, ylim	Numeric length-2 axis limits.

Value

A ggplot object.

Examples

```
data(reliability_datasets)
kurtosis_diagram(reliability_datasets$Yarn, "Yarn", B = 200)
```

---

log_cumulant_diagram *Log-cumulant diagram*

Description

Draws the log-cumulant diagnostic diagram (κ_3 versus κ_2) with the theoretical loci of the six reference distributions, a bootstrap cloud of the sample estimate, and a 95% concentration ellipse.

Usage

```
log_cumulant_diagram(
  data,
  data_name = "Dataset",
  B = NULL,
  seed = 42,
  level = 0.95,
  xlim = c(-2, 2),
  ylim = c(0, 2)
)
```

Arguments

data	Numeric vector of positive observations.
data_name	Character; label used in the title.
B	Integer; bootstrap replicates (default chosen adaptively from n).
seed	Integer random seed.
level	Numeric; ellipse confidence level (default 0.95).
xlim, ylim	Numeric length-2 axis limits.

Value

A ggplot object.

Examples

```
data(reliability_datasets)
log_cumulant_diagram(reliability_datasets$Yarn, "Yarn", B = 200)
```

---

mle_fit *Maximum-likelihood fit of a reliability distribution*

Description

Fits one of the six supported families by maximum likelihood, optimizing on the log-scale of the parameters for numerical stability, and returns the estimates together with the observed-information-based covariance.

Usage

```
mle_fit(x, dist, init = NULL)
```

Arguments

x	Numeric vector of positive observations.
dist	Character; distribution name.
init	Optional numeric length-2 vector of starting values.

Value

A list with elements theta (estimates), Sigma (covariance of $\sqrt{n}(\hat{\theta} - \theta)$), loglik, and conv (convergence flag).

Examples

```
set.seed(1)
x <- rdist(200, "Gamma", c(3, 0.5))
fit <- mle_fit(x, "Gamma")
fit$theta
```

`multi_lc_diagram` *Multi-dataset log-cumulant diagram*

Description

Overlays bootstrap clouds for several datasets on the log-cumulant diagram, distinguishing datasets by colour and plotting symbol (the "empirical data" legend) while the theoretical loci keep the "theoretical curve" legend.

Usage

```
multi_lc_diagram(
  datasets_list,
  dataset_names = NULL,
  B = 1000,
  seed = 42,
  xlim = c(-2, 2),
  ylim = c(0, 2),
  alpha_points = 0.35,
  point_size = 2.6
)
```

Arguments

`datasets_list`	Named list of numeric vectors.
`dataset_names`	Optional character vector of names to use.
`B`	Integer; bootstrap replicates per dataset.
`seed`	Integer random seed.
`xlim`, `ylim`	Numeric length-2 axis limits.
`alpha_points`	Numeric; point transparency.
`point_size`	Numeric; point size.

Value

A `ggplot` object.

Examples

```
data(reliability_datasets)
multi_lc_diagram(reliability_datasets[c("Airplane","BallBearing","Yarn")], B = 300)
```

plot_lc *Quick log-cumulant plot*

Description

Convenience wrapper around log_cumulant_diagram providing the compact plot_lc(data = x, B = 100) interface requested for quick diagnostics. plot.lc is kept as an alias for backward compatibility.

Usage

```
plot_lc(data, B = 100, data_name = "Sample", seed = 42, ...)
```

Arguments

data	Numeric vector of positive observations.
B	Integer; bootstrap replicates.
data_name	Character; label used in the title.
seed	Integer random seed.
...	Further arguments passed to log_cumulant_diagram.

Value

A ggplot object.

Examples

```
data(reliability_datasets)
plot_lc(reliability_datasets$BallBearing, B = 100)
```

power_study *Empirical power study*

Description

Monte Carlo study of the power of the three T^2 tests and the AD/CvM tests against a set of alternative distributions, with optional size-correction.

Usage

```
power_study(
  n = 100,
  Nsim = 1000,
  eta = 0.05,
  alternatives = names(.ALT_CONFIGS),
  use_bootstrap = FALSE,
  B = NULL,
  seed = 2025,
  verbose = TRUE
)
```

Arguments

n	Integer; sample size.
Nsim	Integer; number of Monte Carlo replications.
eta	Numeric; nominal significance level.
alternatives	Character vector of alternative names to evaluate.
use_bootstrap	Logical; use bootstrap calibration.
B	Integer; bootstrap replicates.
seed	Integer random seed.
verbose	Logical; print progress.

Value

A `data.frame` of empirical power by test and alternative.

Examples

```
power_study(n = 100, Nsim = 100)
```

---

reliability_datasets *Nine reliability datasets used in the applications*

Description

A named list with nine positive-valued reliability datasets analyzed in the paper: Kevlar, Resistors, Tensile, Airplane, BallBearing, Airborne, Failure, Yarn, AirCon.

Usage

```
data(reliability_datasets)
```

Format

A named list of numeric vectors.

`size_study` *Empirical size (Type I error) study*

Description

Monte Carlo study of the empirical size of the three T^2 tests (asymptotic and, optionally, bootstrap) and the AD/CvM tests under a true null model, across several sample sizes.

Usage

```
size_study(
  sample_sizes = c(30, 50, 100, 200),
  Nsim = 1000,
  eta = 0.05,
  use_bootstrap = FALSE,
  B = NULL,
  seed = 2025,
  verbose = TRUE
)
```

Arguments

`sample_sizes`	Integer vector of sample sizes.
`Nsim`	Integer; number of Monte Carlo replications.
`eta`	Numeric; nominal significance level.
`use_bootstrap`	Logical; include bootstrap calibration.
`B`	Integer; bootstrap replicates.
`seed`	Integer random seed.
`verbose`	Logical; print progress.

Value

A `data.frame` of empirical rejection rates.

Examples

```
size_study(sample_sizes = c(30, 50), Nsim = 100)
```

T2_all *All three nested T^2 statistics*

Description

Convenience wrapper returning the three nested versions $T^2_{(2,3)}$, $T^2_{(1,2,3)}$, and $T^2_{(1,\ldots,6)}$ for a single fitted model.

Usage

```
T2_all(x, dist, fit = NULL)
```

Arguments

x	Numeric vector of positive observations.
dist	Character; null distribution name.
fit	Optional precomputed mle_fit object.

Value

A named list with components T2_23, T2_123, T2_123456, each as returned by T2_one.

Examples

```
set.seed(1); x <- rdist(100, "Weibull", c(2, 1))
T2_all(x, "Weibull")
```

T2_bootstrap *Parametric bootstrap p-values for the T^2 statistics*

Description

Computes parametric-bootstrap p-values for the three nested T^2 statistics. The bootstrap calibrates the ill-conditioned reference distribution and is the recommended mode of inference in finite samples.

Usage

```
T2_bootstrap(x, dist, B = NULL, fit = NULL, seed = NULL)
```

Arguments

x	Numeric vector of positive observations.
dist	Character; null distribution name.
B	Integer; number of bootstrap replicates (default chosen adaptively).
fit	Optional precomputed mle_fit object.
seed	Optional integer random seed for reproducibility.

Value

A list with the observed statistics and the bootstrap p-values `p_boot` for the three versions.

Examples

```
set.seed(1); x <- rdist(80, "Weibull", c(2, 1))
T2_bootstrap(x, "Weibull", B = 199, seed = 1)
```

`T2_one` *Hotelling-type T^2 statistic for one cumulant set*

Description

Computes the log-cumulant T^2 goodness-of-fit statistic for a single choice of cumulant orders `V`, with the asymptotic chi-squared reference using the corrected (full-rank) degrees of freedom.

Usage

```
T2_one(x, dist, V, fit = NULL)
```

Arguments

`x`	Numeric vector of positive observations.
`dist`	Character; null distribution name.
`V`	Integer vector of cumulant orders (e.g. `c(2,3)`).
`fit`	Optional precomputed `mle_fit` object.

Value

A list with the statistic `T2`, degrees of freedom `df`, and asymptotic p-value `p_asym`.

Examples

```
set.seed(1); x <- rdist(100, "Weibull", c(2, 1))
T2_one(x, "Weibull", V = c(2, 3))
```

theoretical_lc *Theoretical log-cumulants*

Description

Closed-form theoretical log-cumulants $\kappa_1, \ldots, \kappa_{order}$ (Mellin cumulants of the second kind) for a given family and parameter vector, as tabulated in the methodology.

Usage

```
theoretical_lc(dist, theta, order = 6)
```

Arguments

dist	Character; distribution name (see distribution_dispatchers).
theta	Numeric length-2 parameter vector.
order	Integer; highest cumulant order to return (default 6).

Value

Numeric vector of length order with the log-cumulants.

Examples

```
theoretical_lc("Weibull", c(2, 1))
theoretical_lc("Gamma", c(3, 0.5), order = 4)
```

three_diagrams *Combined three-panel diagnostic figure*

Description

Arranges the log-cumulant, kurtosis-skewness, and coefficient-of-variation diagrams side by side for a single dataset.

Usage

```
three_diagrams(data, data_name = "Dataset", B = NULL, seed = 42)
```

Arguments

data	Numeric vector of positive observations.
data_name	Character; label used in the title.
B	Integer; bootstrap replicates (default chosen adaptively from n).
seed	Integer random seed.

Value

A gtable drawn via gridExtra::grid.arrange.

Examples

```
data(reliability_datasets)
three_diagrams(reliability_datasets$Yarn, "Yarn", B = 200)
```

Index