Package 'BivKLD'

September 2, 2026

Type Package

Title Bivariate Kullback-Leibler Divergence

Version 0.1.0

Date 2026-08-22

Description Estimates the directed Kullback-Leibler divergence between two bivariate continuous distributions by numerical integration of kernel density estimates. Also computes pairwise divergences among groups and exact divergences for discrete, bivariate normal, bivariate Pareto type II, and independent bivariate Weibull models. The kernel estimator follows Chackochan, Sankaran and Nair (2026) <doi:10.1080/03610926.2025.2496687>.

License GPL-3

Encoding UTF-8

Depends R (>= 4.1.0)

Imports ks

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Config/roxygen2/version 8.1.0

NeedsCompilation no

Author Saurav Singla [aut, cre],
Rafflesia Chackochan [aut]

Maintainer Saurav Singla <s00singla@gmail.com>

Repository CRAN

Date/Publication 2026-09-02 21:30:09 UTC

Contents

`biv_kld` *Estimate Bivariate Kullback-Leibler Divergence*

Description

Estimates the directed divergence from the distribution represented by `x` to that represented by `y`. Bivariate Gaussian kernel densities are evaluated on a common rectangular grid and numerically integrated.

Usage

```
biv_kld(
  x,
  y,
  Hx = NULL,
  Hy = NULL,
  bandwidth = c("scv", "normal"),
  grid_size = 100L,
  range = NULL,
  standardize = c("none", "pooled", "separate"),
  details = FALSE
)
```

Arguments

`x`, `y`	Numeric matrices or data frames with exactly two columns.
`Hx`, `Hy`	Optional symmetric positive-definite bandwidth matrices. When omitted, bandwidths are selected according to `bandwidth`.
`bandwidth`	Either `"scv"` for unconstrained smoothed cross-validation or `"normal"` for the normal-scale selector.
`grid_size`	One integer or two integers giving the number of grid points in each coordinate.
`range`	Optional integration limits. See Details.
`standardize`	Standardization applied before estimation: `"none"`, `"pooled"`, or `"separate"`.
`details`	If `TRUE`, return densities, grid, and bandwidth matrices in addition to the estimate.

Details

`range` may be a list of two increasing numeric ranges, a two-by-two matrix whose rows contain coordinate limits, or `c(x_min, x_max, y_min, y_max)`. By default the observed pooled range is padded using the selected bandwidths.

Pooled standardization applies the same affine transformation to both samples and therefore leaves the population KL divergence unchanged. Separate standardization, used in the motivating analyses, compares the shapes after removing each sample's own location and scale; it does not estimate the divergence between the original distributions.

Value

A non-negative numeric estimate when `details = FALSE`. Otherwise, an object of class `bivkld_estimate` containing the estimate, bandwidths, evaluation grid, densities, and settings.

References

Chackochan, R., Sankaran, P. G., & Unnikrishnan Nair, N. (2026). Bivariate Kullback-Leibler divergence. *Communications in Statistics - Theory and Methods*, 55(1), 292-312. doi:10.1080/03610926.2025.2496687

Examples

```
set.seed(2026)
x <- cbind(rnorm(60), rnorm(60))
y <- cbind(rnorm(60, 0.4), rnorm(60, -0.2))
biv_kld(x, y, bandwidth = "normal", grid_size = 40)
```

`biv_kld_discrete` *Exact Bivariate KL Divergence for Discrete Distributions*

Description

Computes the directed KL divergence between two probability tables. The inputs can be matrices, arrays, or vectors, but must have identical shapes.

Usage

```
biv_kld_discrete(
  p,
  q,
  normalize = FALSE,
  tolerance = sqrt(.Machine$double.eps)
)
```

Arguments

`p, q`	Probability tables for the first and second distributions.
`normalize`	If `TRUE`, normalize each input to sum to one.
`tolerance`	Allowed absolute error in a probability sum.

Value

A non-negative numeric scalar, possibly `Inf` when `q` is zero at a cell to which `p` assigns positive probability.

Examples

```
p <- matrix(c(0.2, 0.3, 0.1, 0.4), 2)
q <- matrix(c(0.1, 0.4, 0.2, 0.3), 2)
biv_kld_discrete(p, q)
```

`biv_kld_matrix` *Pairwise Bivariate Kullback-Leibler Divergence*

Description

Computes all directed kernel KL divergences among two or more groups on one common integration grid. Each group density and bandwidth is evaluated only once.

Usage

```
biv_kld_matrix(
  x,
  group = NULL,
  H = NULL,
  bandwidth = c("scv", "normal"),
  grid_size = 100L,
  range = NULL,
  standardize = c("none", "pooled", "separate")
)
```

Arguments

`x`	Either a list of bivariate samples or a single numeric matrix or data frame with exactly two columns.
`group`	A grouping vector when `x` is a single sample. Must be `NULL` when `x` is a list.
`H`	Optional list of bandwidth matrices, one per group.
`bandwidth`	Either `"scv"` for unconstrained smoothed cross-validation or `"normal"` for the normal-scale selector.
`grid_size`	One integer or two integers giving the number of grid points in each coordinate.
`range`	Optional integration limits. See Details.
`standardize`	Standardization applied before estimation: `"none"`, `"pooled"`, or `"separate"`.

Value

A numeric matrix. Row `i`, column `j` is the estimated directed divergence from group `i` to group `j`. Bandwidths, grid information, and standardization are stored as attributes.

Examples

```
measurements <- iris[, c("Sepal.Length", "Sepal.Width")]
kld <- biv_kld_matrix(measurements, iris$Species,
                      bandwidth = "normal", grid_size = 40)
round(kld, 3)
```

---

biv_kld_normal *Exact KL Divergence Between Bivariate Normal Distributions*

Description

Computes `D(N(mean1, sigma1) || N(mean2, sigma2))` analytically.

Usage

```
biv_kld_normal(mean1, sigma1, mean2, sigma2)
```

Arguments

`mean1`, `mean2`	Numeric vectors of length two.
`sigma1`, `sigma2`	Symmetric positive-definite two-by-two covariance matrices.

Value

A non-negative numeric scalar.

Examples

```
biv_kld_normal(c(-2, 2), matrix(c(1, 1, 1, 2), 2),
               c(-2, 2), matrix(c(3, 3, 3, 6), 2))
```

---

model_kld *Exact KL Divergence for Two Models from the Reference Paper*

Description

`biv_kld_pareto2()` evaluates the divergence between bivariate Pareto type II models with survival functions `(1 + x1 + x2)^(-alpha)` and `(1 + x1 + x2)^(-beta)`. `biv_kld_independent_weibull()` evaluates the divergence between models with survival functions `exp(-alpha[1] * x1 - alpha[2] * x2)` and the corresponding expression in `beta`.

Usage

```
biv_kld_pareto2(alpha, beta)

biv_kld_independent_weibull(alpha, beta)
```

Arguments

alpha, beta	Positive model parameters. Scalars for `biv_kld_pareto2()` and vectors of length two for `biv_kld_independent_weibull()`.

Value

A non-negative numeric scalar.

References

Chackochan, R., Sankaran, P. G., & Unnikrishnan Nair, N. (2026). Bivariate Kullback-Leibler divergence. *Communications in Statistics - Theory and Methods*, 55(1), 292-312. doi:10.1080/03610926.2025.2496687

Examples

```
biv_kld_pareto2(1, 2)
biv_kld_independent_weibull(c(1, 2), c(2, 3))
```

Index