Package ‘MLE’

September 10, 2024

Type Package

Title Maximum Likelihood Estimation of Various Univariate and Multivariate Distributions

Version 1.1

Date 2024-09-10

Author Michail Tsagris [aut, cre],
Sofia Piperaki [aut],
Muhammad Imran [ctb]

Maintainer Michail Tsagris <mtsagris@uoc.gr>

Depends R (>= 4.0)

Imports BGFD, bivpois, Compositional, Directional, geppe, gp, MN, Rfast, Rfast2, skellam

Suggests mvcauchy

Description Several functions for maximum likelihood estimation of various univariate and multivariate distributions. The list includes more than 100 univariate continuous and discrete distributions, distributions that lie on the real line, the positive line, interval restricted, circular distributions. Further, multivariate continuous and discrete distributions, distributions for compositional and directional data, etc. Some references include Johnson N. L., Kotz S. and Balakrishnan N. (1994). ``Continuous Univariate Distributions, Volume 1'' <ISBN:978-0-471-58495-7>, Johnson, Norman L. Kemp, Adrianne W. Kotz, Samuel (2005). ``Univariate Discrete Distributions''. <ISBN:978-0-471-71580-1> and Mardia, K. V. and Jupp, P. E. (2000). ``Directional Statistics''. <ISBN:978-0-471-95333-3>.

NeedsCompilation no

Repository CRAN

Date/Publication 2024-09-10 12:20:05 UTC

Contents

`MLE-package` *Maximum Likelihood Estimation of Various Univariate and Multivariate Distributions*

Description

The package offers functions for the maximum likelihood estimation of various univariate and multivariate distributions. The list includes univariate continuous and discrete distributions, distributions that lie on the real line, the positive line, interval restricted, circular distributions. Further, multivariate continuous and discrete distributions, distributions for compositional and directional data, etc. The references are included within each set of functions.

Details

Package:	MLE
Type:	Package
Version:	1.1
Date:	2024-09-10
License:	GPL-2

Maintainers

Michail Tsagris <mtsagris@uoc.gr>.

Author(s)

Michail Tsagris <mtsagris@uoc.gr>, Sofia Piperaki <sofiapip23@gmail.com> and Muhammad Imran <imranshakoor84@yahoo.com>.

`Column-wise MLE of continuous univariate distributions defined on the positive line`

Column-wise MLE of continuous univariate distributions defined on the positive line

---

Description

Column-wise MLE of continuous univariate distributions defined on the positive line.

Usage

```
colpositive.mle(x, distr = "gamma", tol = 1e-07, maxiters = 100, parallel = FALSE)
```

Arguments

`x`	A matrix with positive valued data (zeros are not allowed).
`distr`	The distribution to fit. "gamma" stands for the gamma distribution, "weibull" for the Weibull, "pareto" for the Pareto distribution, "exp" for the exponential distribution, "exp2" I do not remember, "maxboltz" for the Maxwell-Boltzman distribution, "rayleigh" for the Rayleigh distribution and "lindley" for the Lindley distribution, "lognorm" for the log-normal distribution. "halfnorm" for the half-normal, "invgauss" for the inverse Gaussian. The "normlog" is simply the normal distribution where all values are positive. Note, this is not log-normal. It is the normal with a log link. Similarly to the inverse gaussian distribution where the mean is an exponentiated. This comes from the GLM theory. The "powerlaw" stands for the power law distribution.
`tol`	The tolerance level up to which the maximisation stops; set to 1e-07 by default.
`maxiters`	The maximum number of iterations the Newton-Raphson will perform for the Weibull distribution.
`parallel`	Do you want to calculations to take place in parallel? The default value is FALSE. This is only for the Weibull distribution.

Details

For each column, the same distribution is fitted and its parameter and log-likelihood are computed.

Value

A matrix with two, three or five (for the colnormlog.mle) columns. The first one or the first two contain the parameter(s) of the distribution and the other columns contain the log-likelihood values.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Kalimuthu Krishnamoorthy, Meesook Lee and Wang Xiao (2015). Likelihood ratio tests for comparing several gamma distributions. Environmetrics, 26(8): 571–583.

N.L. Johnson, S. Kotz and N. Balakrishnan (1994). Continuous Univariate Distributions, Volume 1 (2nd Edition).

N.L. Johnson, S. Kotz a nd N. Balakrishnan (1970). Distributions in statistics: continuous univariate distributions, Volume 2.

Tsagris M., Beneki C. and Hassani H. (2014). On the folded normal distribution. Mathematics, 2(1): 12–28.

Sharma V. K., Singh S. K., Singh U. and Agiwal V. (2015). The inverse Lindley distribution: a stress-strength reliability model with application to head and neck cancer data. Journal of Industrial and Production Engineering, 32(3): 162–173.

You can also check the relevant wikipedia pages for these distributions.

See Also

disc.mle, real.mle, prop.mle

Examples

```
x <- rgamma(100, 3, 4)
positive.mle(x, distr = "gamma")
```

`Column-wise MLE of continuous univariate distributions defined on the real line`
Column-wise MLE of continuous univariate distributions defined on the real line

---

Description

Column-wise MLE of continuous univariate distributions defined on the real line.

Usage

```
colreal.mle(x, distr = "normal", tol = 1e-07, maxiters = 100, parallel = FALSE)
```

Arguments

`x`	A numerical vector with data.
`distr`	The distribution to fit, "normal" stands for the normal distribution, "cauchy" for the Cauchy, "laplace" is the Laplace distribution.
`tol`	The tolerance level to stop the iterative process of finding the MLEs.
`maxiters`	The maximum number of iterations to implement.
`parallel`	Should the computations take place in parallel?

Details

Instead of maximising the log-likelihood via a numerical optimiser we have used a Newton-Raphson algorithm which is faster. See wikipedia for the equation to be solved. For the t distribution we need the degrees of freedom and estimate the location and scatter parameters.

The Cauchy is the t distribution with 1 degree of freedom. The Laplace distribution is also called double exponential distribution.

Value

A matrix with two, columns. The first one contains the parameters of the distribution and the second columns contains the log-likelihood values.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Johnson, Norman L. Kemp, Adrianne W. Kotz, Samuel (2005). Univariate Discrete Distributions (third edition). Hoboken, NJ: Wiley-Interscience.

https://en.wikipedia.org/wiki/Wigner_semicircle_distribution

See Also

`positive.mle`, `circ.mle`, `disc.mle`

Examples

```
x <- rnorm(1000, 10, 2)
a <- real.mle(x, distr = "normal")
```

Column-wise MLE of distributions defined in the (0, 1) interval
Column-wise MLE of distributions defined in the (0, 1) interval

---

Description

Column-wise MLE of distributions defined in the (0, 1) interval.

Usage

```
colprop.mle(x, distr = "beta", tol = 1e-07, maxiters = 100, parallel = FALSE)
```

Arguments

`x`	A numerical vector with proportions, i.e. numbers in (0, 1) (zeros and ones are not allowed).
`distr`	The distribution to fit. "beta" stands for the beta distribution, "logitnorm" is the logistic normal, "unitweibull" is the unit-Weibull and the "sp" is the standard power distribution.
`tol`	The tolerance level up to which the maximisation stops.
`maxiters`	The maximum number of iterations the Newton-Raphson will perform.
`parallel`	Should the computations take place in parallel? This is for the "spml" only.

Details

Maximum likelihood estimation of the parameters of the beta distribution is performed via Newton-Raphson. The distributions and hence the functions does not accept zeros. "logitnorm.mle" fits the logistic normal, hence no nwewton-Raphson is required and the "hypersecant01.mle" uses the golden ratio search as is it faster than the Newton-Raphson (less calculations).

Value

A matrix with two, columns. The first one contains the parameters of the distribution and the second columns contains the log-likelihood values.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris `<mtsagris@uoc.gr>` and Sofia Piperaki `<sofiapip23@gmail.com>`.

References

N.L. Johnson, S. Kotz and N. Balakrishnan (1994). Continuous Univariate Distributions, Volume 1 (2nd Edition).

N.L. Johnson, S. Kotz and N. Balakrishnan (1970). Distributions in statistics: continuous univariate distributions, Volume 2.

J. Mazucheli, A. F. B. Menezes, L. B. Fernandes, R. P. de Oliveira and M. E. Ghitany (2020). The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates. Journal of Applied Statistics, 47(6): 954–974.

See Also

`prop.mle`, `positive.mle`

Examples

```
x <- rbeta(1000, 1, 4)
prop.mle(x, distr = "beta")
```

`Column-wise MLE of some censored models`
Column-wise MLE of some censored models

Description

Column-wise MLE of some censored models.

Usage

```
colcens.mle(x, distr = "censweibull", di, tol = 1e-07, parallel = FALSE, cores = 0)
```

Arguments

`x`	A vector with positive valued data and zero values. If there are no zero values, a simple normal model is fitted in the end.
`distr`	The distribution to fit. "censweibull" for the censored Weibull and "censpois" for the left censored Poisson. For the "censpois" the lowest value in x is taken as the censored point and values below that number are considered to be censored.
`di`	A vector of 0s (censored) and 1s (not censored) values.
`tol`	The tolerance level up to which the maximisation stops; set to 1e-07 by default.
`parallel`	Do you want to calculations to take place in parallel? The default value is FALSE.
`cores`	In case you set parallel = TRUE, then you need to specify the number of cores.

Details

For each column, the same distribution is fitted and its parameters and log-likelihood are computed.

Value

A matrix with two or three columns. The first one or the first two contain the parameter(s) of the distribution and the second or third column the relevant log-likelihood.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Tobin James (1958). Estimation of relationships for limited dependent variables. Econometrica. 26(1): 24–36.

https://en.wikipedia.org/wiki/Tobit_model

Fritz Scholz (1996). Maximum Likelihood Estimation for Type I Censored Weibull Data Including Covariates. Technical report. ISSTECH-96-022, Boeing Information & Support Services, P.O. Box 24346, MS-7L-22.

See Also

cens.mle, colpositive.mle, colreal.mle

Examples

```
x1 <- matrix( rpois(1000 * 10, 15), ncol = 10)
x <- x1
x[x <= 10] <- 10
colMeans(x) ## simple Poisson
colcens.mle(x, distr = "censpois")
```

Column-wise MLE of some circular distributions

Column-wise MLE of some circular distributions

Description

Column-wise MLE of some circular distributions.

Usage

```
colcirc.mle(x, distr = "vm", tol = 1e-07, maxiters = 100, parallel = FALSE)
```

Arguments

x	A numerical matrix with the circular data. They must be expressed in radians.
distr	The type of distribution to fit, "vm" stands for the von Mises and "spml" is the angular Gaussian distribution.
tol	The tolerance level to stop the iterative process of finding the MLEs.
maxiters	The maximum number of iterations to implement. This is for the "spml" only.
parallel	Should the computations take place in parallel? This is for the "spml" only.

Details

The parameters of the von Mises, the bivariate angular Gaussian and wrapped Cauchy distributions are estimated. For the wrapped Cauchy, the iterative procedure described by Kent and Tyler (1988) is used. As for the von Mises distribution, we use a Newton-Raphson to estimate the concentration parameter. The angular Gaussian is described, in the regression setting in Presnell et al. (1998).

Value

A matrix with two, columns. The first one contains the parameters of the distribution and the second columns contains the log-likelihood values.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Mardia K. V. and Jupp P. E. (2000). Directional statistics. Chicester: John Wiley & Sons.

Sra S. (2012). A short note on parameter approximation for von Mises-Fisher distributions: and a fast implementation of Is(x). Computational Statistics, 27(1): 177-190.

Presnell Brett, Morrison Scott P. and Littell Ramon C. (1998). Projected multivariate linear models for directional data. Journal of the American Statistical Association, 93(443): 1068–1077.

Kent J. and Tyler D. (1988). Maximum likelihood estimation for the wrapped Cauchy distribution. Journal of Applied Statistics, 15(2): 247–254.

See Also

circ.mle,

Examples

```
x <- matrix( rnorm(100 * 10, 3, 1), ncol = 10)
x <- x / sqrt( rowSums(x^2) )
res <- colcirc.mle(x, distr = "spml")
```

`Column-wise MLE of some discrete distributions`
Column-wise MLE of some discrete distributions

Description

Column-wise MLE of some discrete distributions.

Usage

```
coldisc.mle(x, distr = "poisson", type = 1)
```

Arguments

`x`	A numerical matrix with count data, dscrete data, integers. Each column refers to a different vector of observations of the same distribution.
`distr`	The distribution to fit, "poisson" stands for the Poisson, "geom" for the geometric distribution, "borel" for the Borel distribution and "gamma" for the Gamma distribution.
`type`	This is for the geometric distribution only. Type 1 refers to the case where the minimum is zero and type 2 for the case of the minimum being 1.

Details

For each column, the same distribution is fitted and its parameter and log-likelihood are computed.

Value

A matrix with two, columns. The first one contains the parameters of the distribution and the second columns contains the log-likelihood values.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Johnson Norman L., Kotz Samuel and Balakrishnan (1997). Discrete Multivariate Distributions. Wiley.

See Also

disc.mle

Examples

```
x <- matrix(rpois(1000 * 50, 10), ncol = 50)
a <- coldisc.mle(x, distr = "poisson")
```

Column-wise MLE of the ordinal model without covariates
Column-wise MLE of the ordinal model without covariates

Description

Column-wise MLE of the ordinal model without covariates.

Usage

```
colordinal.mle(y, link = "logit")
```

Arguments

y	A numerical matrix with values 1, 2, 3,..., not zeros, or a data.frame with ordered factors.
link	This can either be "logit" or "probit". It is the link function to be used.

Details

Maximum likelihood of the ordinal model (proportional odds) is implemented. See for example the "polr" command in R or the examples.

Value

A list including:

param	A matrix with the intercepts (threshold coefficients) of the model applied to each column (or variable).
loglik	The log-likelihood values.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Agresti, A. (2002) Categorical Data. Second edition. Wiley.

See Also

ordinal.mle

Examples

```
y <- matrix( rbinom(100 * 10, 2, 0.5) + 1, ncol = 10 )
res <- colordinal.mle(y, link = "probit")
```

`MLE for multivariate discrete data`
MLE for multivariate discrete data

Description

MLE for multivariate discrete data.

Usage

```
mvdisc.mle(x, distr = "multinom", tol = 1e-07)
```

Arguments

`x`	A matrix with discrete valued non negative data.
`distr`	The distribution to fit. "multinom" stands for the multinomial distribution, "dirimultinom" stands for the Dirichlet-multinomial distribution. "bp.mle" and "bp.mle2" stand for the bivariate Poisson distribution. The The "bp.mle" returns a lot of information and is slower than "bp.mle2", which returns fewer information, but is faster.
`tol`	The tolerance level to terminate the Newton-Raphson algorithm for the Dirichlet multinomial distribution.

Value

A list including:

`iters`	The number of iterations required by the Newton-Raphson algortihm.
`loglik`	A vector with the value of the maximised log-likelihood.
`param`	A vector of the parameters.
`prob`	A vector with the estimated probabilities.

For the "bp.mle" a list including:

`lambda`	A vector with the estimated values of (λ_1, λ_2) and λ_3. Note that $\hat{\lambda}_1 = \bar{x}_1 - \lambda_3$ and $\hat{\lambda}_1 = \bar{x}_1 - \lambda_3$, where $\bar{x}_1$ and $\bar{x}_2$ are the two sample means.
`rho`	The estimated correlation coefficient, that is: $\frac{\hat{\lambda}_3}{\sqrt{\left(\hat{\lambda}_1 + \hat{\lambda}_3\right)\left(\hat{\lambda}_2 + \hat{\lambda}_3\right)}}$.
`ci`	The 95% Confidence intervals using the observed and the asymptotic information matrix.

`loglik` The log-likelihood values assuming independence ($\lambda_3 = 0$) and assuming the bivariate Poisson distribution.

`pvalue` Three p-values for testing $\lambda_3 = 0$. These are based on the log-likelihood ratio and two Wald tests using the observed and the asymptotic information matrix.

For the "bp.mle2" a list including:

`lambda` A vector with the estimated values of (λ_1, λ_2) and λ_3. Note that $\hat{\lambda}_1 = \bar{x}_1 - \lambda_3$ and $\hat{\lambda}_1 = \bar{x}_1 - \lambda_3$, where $\bar{x}_1$ and $\bar{x}_2$ are the two sample means.

`loglik` The log-likelihood values assuming independence ($\lambda_3 = 0$) and assuming the bivariate Poisson distribution.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris `<mtsagris@uoc.gr>` and Sofia Piperaki `<sofiapip23@gmail.com>`.

References

Johnson Norman L., Kotz Samuel and Balakrishnan (1997). Discrete Multivariate Distributions. Wiley.

Kawamura K. (1984). Direct calculation of maximum likelihood estimator for the bivariate Poisson distribution. Kodai Mathematical Journal, 7(2): 211–221.

Kocherlakota S. and Kocherlakota K. (1992). Bivariate discrete distributions. CRC Press.

Karlis D. and Ntzoufras I. (2003). Analysis of sports data by using bivariate poisson models. Journal of the Royal Statistical Society: Series D (The Statistician), 52(3): 381–393.

See Also

`disc.mle`, `coldisc.mle`

Examples

```
x <- t( rmultinom(1000, 20, c(0.4, 0.5, 0.1) ) )
mvdisc.mle(x, distr = "multinom")
```

`MLE of (hyper-)spherical distributions`

MLE of (hyper-)spherical distributions

Description

MLE of (hyper-)spherical distributions.

Usage

```
hspher.mle(x, distr = "vmf", ina, full = FALSE, ell = FALSE, tol = 1e-07)
```

Arguments

`x`	A matrix with directional data, i.e. unit vectors.
`distr`	The distribution to fit. Spherical distributions: "purka" is the Purkayastha distribution, and "sipc" is the spherical isotropic projected Cauchy. These are rotationally symmetric distributions. The "wood" is the Wood distribution, a bimodal distribution. The next three are elliptically symmetric distributions. The "kent" is the Kent distribution, "esag" is the elliptically symmetric angular Gaussian and the "sespc" is the spherical elliptically symmetric projected Cauchy distribution. Spherical and hyper-spherical distributions: "vmf" stands for the von Mises-Fisher distribution, "multivmf" is for the vMF with multiple groups, "acg" is the angular central Gaussian and and the "pkbd" is the Poisson kernel based distribution. The "spcauchy" and "spcauchy2" are the spherical Cauchy (2 different methods of estimation) ,"pkbd" and "pkbd2" is the Poisson-kernel based distribution, "iag" stands for the independent angular Gaussian distribution, and works for spherical and hyper-spherical data, and "ESAGd" is the generalization to the hyper-sphere.
`ina`	A numerical vector with discrete numbers starting from 1, i.e. 1, 2, 3, 4,... or a factor variable. Each number denotes a sample or group. If you supply a continuous valued vector the function will obviously provide wrong results.
`full`	If you want some extra information, the inverse of the covariance matrix, set this equal to TRUE. Otherwise leave it FALSE.
`ell`	This is for the multivmf.mle only. Do you want the log-likelihood returned? The default value is TRUE.
`tol`	The tolerance value at which to terminate the iterations.

Details

For the von Mises-Fisher, the normalised mean is the mean direction. For the concentration parameter, a Newton-Raphson is implemented. For the angular central Gaussian distribution there is a constraint on the estimated covariance matrix; its trace is equal to the number of variables. An iterative algorithm takes place and convergence is guaranteed. Newton-Raphson for the projected normal distribution, on the sphere, is implemented as well.

The "vmf" estimates the mean direction and concentration of a fitted von Mises-Fisher distribution. The von Mises-Fisher distribution for groups of data is also implemented. The "acg"" fits the angular central Gaussian distribution. There is a constraint on the estimated covariance matrix; its trace is equal to the number of variables. An iterative algorithm takes place and convergence is guaranteed. The "iag" implements MLE of the (hyper-)spherical projected normal distribution. The "esag" is for spherical data, while "ESAGd" is for hyper-spherical data. The "spcauchy" is faster than the "spcacuhy2" because it employs the Newton-Raphson algortihm, but for high dimensions the latter is preferred. Both functions estimate the parameters of the spherical Cauchy distribution, for any dimension. Despite the name sounds confusing, it is implemented for arbitrary dimensions, not only the sphere. The function employs a combination of the fixed points iteration algorithm and the Brent algorithm. The "pkbd" is faster than "pkbd2" beacuse it employs the Newton-Raphson algorithm, but for high dimensions the latter is preferred. Both estimate the parameters of the Poisson kernel based distribution (PKBD), for any dimension. The "sipc" implements MLE of the spherical independent projected Cauchy distribution, for spherical data only.

Value

For the von Mises-Fisher a list including:

- `loglik` The maximum log-likelihood value.
- `mu` The mean direction.
- `kappa` The concentration parameter.

For the multi von Mises-Fisher a list including:

- `loglik` A vector with the maximum log-likelihood values if ell is set to TRUE. Otherwise NULL is returned.
- `mi` A matrix with the group mean directions.
- `ki` A vector with the group concentration parameters.

For the angular central Gaussian a list including:

- `iter` The number if iterations required by the algorithm to converge to the solution.
- `cova` The estimated covariance matrix.

For the spherical projected normal a list including:

- `iters` The number of iteration required by the Newton-Raphson.
- `mesi` A matrix with two rows. The first row is the mean direction and the second is the mean vector. The first comes from the second by normalising to have unit length.
- `param` A vector with the elements, the norm of mean vector, the log-likelihood and the log-likelihood of the spherical uniform distribution. The third value helps in case you want to do a log-likleihood ratio test for uniformity.

For the spherical Cauchy and the PKBD a list including:

- `mesos` The mean in R^{d+1}. See Tsagris and Alenazy (2023) for a re-parametrization that applies in the spherical Cauchy also.
- `mu` The mean direction.
- `gamma` The norm of the mean in R^{d+1}. See Tsagris and Alenazy (2023) for a re-parametrization that applies in the spherical Cauchy also.
- `rho` The concetration parameter, this takes values in [0, 1).
- `loglik` The log-likelihood value.

For the SIPC a list including:

- `mu` The mean direction.

loglik	The log-likelihood value.

For the Kent a list including:

runtime	The run time of the procedure.
G	A 3 x 3 matrix whose first column is the mean direction. The second and third columns are the major and minor axes respectively.
param	A vector with the concentration κ and ovalness β parameters and the angle ψ used to rotate **H** and hence estimate **G** as in Kent (1982).
logcon	The logarithm of the normalising constant, using the third type approximation (Kume and Wood, 2005).
loglik	The value of the log-likelihood.

For the ESAG a list including:

mu	The mean vector in R^3.
gam	The two γ parameters.
loglik	The log-likelihood value.
vinv	The inverse of the covariance matrix. It is returned if the argument "full" is TRUE.
rho	The rho parameter (smallest eigenvalue of the covariance matrix). It is returned if the argument "full" is TRUE.
psi	The angle of rotation ψ set this equal to TRUE. It is returned if the argument "full" is TRUE.
iag.loglik	The log-likelihood value of the isotropic angular Gaussian distribution. That is, the projected normal distribution which is rotationally symmetric.

For the SESPC a list including:

mu	The mean vector in R^3.
theta	The two θ parameters.
loglik	The log-likelihood value.
vinv	The inverse of the covariance matrix. It is returned if the argument "full" is TRUE.
lambda	The λ_2 parameter (smallest eigenvalue of the covariance matrix). It is returned if the argument "full" is TRUE.
psi	The angle of rotation ψ set this equal to TRUE. It is returned if the argument "full" is TRUE.
sipc.loglik	The log-likelihood value of the isotropic prohected Cuchy distribution, which is rotationally symmetric.

For the Wood distribution a list including:

info	A 5 x 3 matrix containing the 5 parameters, γ, δ, α, β and κ along with their corresponding 95% confidence intervals all expressed in degrees.
modes	The two axis of the modes of the distribution expressed in degrees.

unitvectors	A 3 x 3 matrix with the 3 unit vectors associated with the γ and δ parameters.
loglik	The value of the log-likelihood.

For the Purkayastha a list including:

theta	The median direction.
alpha	The concentration parameter.
loglik	The log-likelihood.
alpha.sd	The standard error of the concentration parameter.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Mardia, K. V. and Jupp, P. E. (2000). Directional statistics. Chicester: John Wiley & Sons.

Sra, S. (2012). A short note on parameter approximation for von Mises-Fisher distributions: and a fast implementation of Is(x). Computational Statistics, 27(1): 177–190.

Tyler D. E. (1987). Statistical analysis for the angular central Gaussian distribution on the sphere. Biometrika 74(3): 579-589.

Zehao Yu and Xianzheng Huang (2024). A new parameterization for elliptically symmetric angular Gaussian distributions of arbitrary dimension. Electronic Journal of Statististics, 18(1): 301–334.

Tsagris M. (2024). Directional data analysis using the spherical Cauchy and the Poisson-kernel based distribution. https://arxiv.org/pdf/2409.03292

Paine P.J., Preston S.P., Tsagris M and Wood A.T.A. (2018). An Elliptically Symmetric Angular Gaussian Distribution. Statistics and Computing, 28, 689–697.

Tsagris M. and Alzeley O. (2023). Circular and spherical projected Cauchy distributions: A Novel Framework for Circular and Directional Data Modeling. https://arxiv.org/pdf/2302.02468.pdf

Kato S. and McCullagh P. (2020). Some properties of a Cauchy family on the sphere derived from the Mobius transformations. Bernoulli, 26(4): 3224–3248. https://arxiv.org/pdf/1510.07679.pdf

Golzy M. and Markatou M. (2020). Poisson kernel-based clustering on the sphere: convergence properties, identifiability, and a method of sampling. Journal of Computational and Graphical Statistics, 29(4): 758–770.

Sablica L., Hornik K. and Leydold J. (2023). Efficient sampling from the PKBD distribution. Electronic Journal of Statistics, 17(2): 2180–2209.

Wood A.T.A. (1982). A bimodal distribution on the sphere. Journal of the Royal Statistical Society, Series C, 31(1): 52–58.

Purkayastha S. (1991). A Rotationally Symmetric Directional Distribution: Obtained through Maximum Likelihood Characterization. The Indian Journal of Statistics, Series A, 53(1): 70–83

Cabrera J. and Watson G. S. (1990). On a spherical median related distribution. Communications in Statistics-Theory and Methods, 19(6): 1973–1986.

See Also

circ.mle

Examples

```
m <- c(0, 0, 0, 0)
s <- cov(iris[, 1:4])
x <- matrix( rnorm(100 * 3), ncol = 3 )
x <- x / sqrt( rowSums(x^2) )
hspher.mle(x, distr = "iag")
```

MLE of Bell type (univariate continuous) distributions

MLE of Bell type (univariate continuous) distributions

Description

MLE of Bell type (univariate continuous) distributions.

Usage

```
bell.mle(x, a, b, k, lambda, distr = "BB12", method = "B")
```

Arguments

x	A vector with continuous valued data.
a	Initial value for the strictly positive scale parameter of the baseline distribution.
b	Initial value for the strictly positive shape parameter of the baseline distribution.
k	Initial value for the strictly positive shape parameter of the baseline distribution.
lambda	Initial value for the strictly positive parameter of the Bell distribution.
distr	The distribution to fit, "BB12" stands for the Bell Burr-12, "BBX" for the Bell Burr-10, "BE" for the Bell exponential, "BEW" for the Bell exponentiated Weibull, "BEE" for the Bell exponentiated exponential, "BF" for the Bell Fisk distribution, "BL" for the Bell Lomax, "BW" for the Bell Weibull distribution, "CBB12" for the complementary Bell Burr-12, "CBBX" for the complementary Bell Burr-X distribution, "CBE" for the complementary Bell exponential distribution, "CBEW" for the complementary Bell exponentiated Weibul distribution, "CBEE" for the complementary Bell extended exponentia distribution, "CBF" for the complementary Bell Fisk distribution, "CBL" for the complementary Bell Lomax distribution, and "CBW" for the complementary Bell Weibull distribution.
method	The procedure for optimising the log-likelihood function after setting the initial values of the parameters and data vector for which the Bell-based distributions are fitted. It could be "Nelder-Mead," "BFGS," "CG," "L-BFGS-B," or "SANN." "BFGS" is set as the default.

Details

These functions facilitate the fitting of Bell-based extended distributions, including the Bell Burr-12(a, b, k, lambda), Bell Burr-10(a, lambda), Bell exponential(a, lambda), Bell exponentiated Weibull(a, b, k, lambda), Bell extended exponential(a, b, lambda), Bell Fisk(a, b, lambda), Bell Lomax(a, b, lambda), Bell Weibull(a, b, lambda), complementary Bell Burr-12(a, b, k, lambda), complementary Bell Burr-10(a, lambda), complementary Bell exponential(a, lambda), complementary Bell exponentiated Weibull(a, b, k, lambda), complementary Bell extended exponential(a, b, lambda), complementary Bell Fisk(a, b, lambda), complementary Bell Lomax(a, b, lambda), and complementary Bell Weibull(a, b, lambda).

Value

A list including:

`param`	The parameters of the distribution.
`loglik`	The log-likelihood value.

Author(s)

Muhammad Imran.

R implementation and documentation: Muhammad Imran <`imranshakoor84@yahoo.com`>.

References

Fayomi A., Tahir M. H., Algarni A., Imran M. and Jamal F. (2022). A new useful exponential model with applications to quality control and actuarial data. Computational Intelligence and Neuroscience, 2022.

Alanzi, A. R., Imran M., Tahir M. H., Chesneau C., Jamal F. Shakoor S. and Sami, W. (2023). Simulation analysis, properties and applications on a new Burr XII model based on the Bell-X functionalities. AIMS Mathematics, 8(3): 6970–7004.

Algarni A. (2022). Group Acceptance Sampling Plan Based on New Compounded Three- Parameter Weibull Model. Axioms, 11(9): 438.

Kleiber, C. and Kotz, S. (2003). Statistical size distributions in economics and actuarial sciences. John Wiley & Sons.

Zimmer W. J., Keats J. B. and Wang F. K. (1998). The Burr XII distribution in reliability analysis. Journal of Quality Technology, 30(4): 386–394.

Nadarajah S., Cordeiro G. M. and Ortega E. M. (2013). The exponentiated Weibull distribution: a survey. Statistical Papers, 54: 839–877.

Nadarajah S. (2011). The exponentiated exponential distribution: a survey. Advances in Statistical Analysis, 95: 219–251.

See Also

`disc.mle`

Examples

```
x <- rgamma(1000, 3, 5)
# Fitting of the Bell Burr-12 (BB12) distribution
bell.mle(x, a = 2.1, b = 1.3, k = 0.02, lambda = 1.2, distr = "BB12", method = "B")
# Fitting of the Bell exponential (BE) distribution
bell.mle(x, a = 2.1, lambda = 0.5 ,distr = "BE", method = "B")
```

MLE of continuous univariate distributions defined on the positive line

MLE of continuous univariate distributions defined on the positive line

---

Description

MLE of continuous univariate distributions defined on the positive line.

Usage

```
positive.mle(x, distr = "gamma", tol = 1e-07, maxiters = 100)
```

Arguments

x	A vector with positive valued data (zeros are not allowed).
distr	The distribution to fit. "gamma" stands for the gamma distribution, "chisq" for the χ^2 distribution, "weibull" for the Weibull, "lomax" for the Lomax, "foldnorm" for the folded normal, "betaprime" for the beta-prime distribution, "lognorm" for the log-normal, "logcauchy" for the log-Cauchy, "loglogictic" for the log-logistic distribution. "halfnorm" for the half-normal, "invgauss" for the inverse Gaussian, "pareto" for the Pareto distribution, "exp" for the exponential distribution, "exp2" I do not remember, "maxboltz" for the Maxwell-Boltzman distribution, "rayleigh" is the Rayleigh distribution, "lindley" is the Lindley distribution, "halfcauchy" is the half-Cauchy distribution and "powerlaw" is the power law distribution. The "normlog" is simply the normal distribution where all values are positive. Note, this is not log-normal. It is the normal with a log link. Similarly to the inverse gaussian distribution where the mean is an exponentiated. This comes from the GLM theory. The "epois" stands for the exponential-Poisson, the "gep" for the generalized exponential-Poisson and the "pe" for the Poisson-exponential distribution, the "zigamma" and "ziweibull" stand for the zero inflated gamma and Weibull distributions, respectively, and they accept zeros.
tol	The tolerance level up to which the maximisation stops; set to 1e-07 by default.
maxiters	The maximum number of iterations the Newton-Raphson will perform.

Details

Instead of maximising the log-likelihood via a numerical optimiser we have used a Newton-Raphson algorithm which is faster. See wikipedia for the equations to be solved. For the t distribution we need the degrees of freedom and estimate the location and scatter parameters.

Value

Usually a list with three elements, but this is not for all cases.

`iters`	The number of iterations required for the Newton-Raphson to converge.
`loglik`	The value of the maximised log-likelihood.
`param`	The vector of the parameters.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Kalimuthu Krishnamoorthy, Meesook Lee and Wang Xiao (2015). Likelihood ratio tests for comparing several gamma distributions. Environmetrics, 26(8):571–583.

N.L. Johnson, S. Kotz and N. Balakrishnan (1994). Continuous Univariate Distributions, Volume 1 (2nd Edition).

N.L. Johnson, S. Kotz a nd N. Balakrishnan (1970). Distributions in statistics: continuous univariate distributions, Volume 2.

Tsagris M., Beneki C. and Hassani H. (2014). On the folded normal distribution. Mathematics, 2(1):12–28.

Sharma V. K., Singh S. K., Singh U. and Agiwal V. (2015). The inverse Lindley distribution: a stress-strength reliability model with application to head and neck cancer data. Journal of Industrial and Production Engineering, 32(3): 162–173.

Barreto-Souza W. and Cribari-Neto F. (2009). A generalization of the exponential-Poisson distribution. Statistics and Probability Letters, 79(24): 2493–2500.

Louzada F., Ramos P. L. and Ferreira H. P. (2020). Exponential-Poisson distribution: estimation and applications to rainfall and aircraft data with zero occurrence. Communications in Statistics–Simulation and Computation, 49(4): 1024–1043.

Rodrigues G. C., Louzada F. and Ramos P. L. (2018). Poisson-exponential distribution: different methods of estimation. Journal of Applied Statistics, 45(1): 128–144.

Taylor S. and Pollard K. (2009). Hypothesis Tests for Point-Mass Mixture Data with Application to Omics Data with Many Zero Values. Statistical Applications in Genetics and Molecular Biology, 8(1): 1–43.

You can also check the relevant wikipedia pages for these distributions.

See Also

`disc.mle`, `real.mle`, `prop.mle`

Examples

```
x <- rgamma(100, 3, 4)
positive.mle(x, distr = "gamma")
```

`MLE of continuous univariate distributions defined on the real line`
MLE of continuous univariate distributions defined on the real line

Description

MLE of continuous univariate distributions defined on the real line.

Usage

```
real.mle(x, distr = "normal", v = 5, tol = 1e-7)
```

Arguments

`x`	A numerical vector with data.
`distr`	The distribution to fit, "normal" stands for the normal distribution, "gumbel" for the Gumbel, "cauchy" for the Cauchy, "logistic" for the logistic distribution, "ct" for the (central) t distribution, "t" for the (non-central) t distribution, "wigner" is the Wigner semicircle distribution and "laplace" is the Laplace distribution. "cauchy0" and "gnormal0" are the Cauchy and generalised normal distributions, respectively, with zero location. The generalised normal distribution is also known as the exponential power distribution or the generalized error distribution.
`v`	The degrees of freedom of the t distribution.
`tol`	The tolerance level up to which the maximisation stops set to 1e-07 by default.

Details

Instead of maximising the log-likelihood via a numerical optimiser we have used a Newton-Raphson algorithm which is faster. See wikipedia for the equation to be solved. For the t distribution we need the degrees of freedom and estimate the location and scatter parameters.

The Cauchy is the t distribution with 1 degree of freedom. The Laplace distribution is also called double exponential distribution.

Value

Usually a list with three elements, but this is not for all cases.

`iters`	The number of iterations required for the Newton-Raphson to converge.
`scale`	The estimated scale parameter of the Cauchy distribution.
`loglik`	The value of the maximised log-likelihood.
`param`	The vector of the parameters.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Johnson, Norman L. Kemp, Adrianne W. Kotz, Samuel (2005). Univariate Discrete Distributions (third edition). Hoboken, NJ: Wiley-Interscience.

https://en.wikipedia.org/wiki/Wigner_semicircle_distribution

Do M.N. and Vetterli M. (2002). Wavelet-based Texture Retrieval Using Generalised Gaussian Density and Kullback-Leibler Distance. Transaction on Image Processing. 11(2): 146–158.

See Also

positive.mle, circ.mle, disc.mle

Examples

```
x <- rnorm(1000, 10, 2)
a <- real.mle(x, distr = "normal")
```

MLE of count data (univariate discrete distributions)

MLE of count data

Description

MLE of count data.

Usage

```
disc.mle(x, distr = "poisson", N = NULL, type = 1, tol = 1e-07)
```

Arguments

x	A vector with discrete valued data.
distr	The distribution to fit, "poisson" stands for the Poisson, "zip" for the zero-inflated Poisson, "ztp" for the zero-truncated Poisson, "negbin" for the negative binomial, "binom" for the binomial, "borel" for the Borel distribution, "geom" for the geometric, "logseries" for the log-series distribution, "betageom" for the beta-geometric, "betabinom" for the beta-binomial distribution and "skellam" for the Skellam distribution, "gp" for the generalised Poisson distribution and "gammapois" for the gamma-Poisson distribution.

`type`	This argument is for the negative binomial and the geometric distribution. In the negative binomial you can choose which way your prefer. Type 1 is for smal sample sizes, whereas type 2 is for larger ones as is faster. For the geometric it is related to its two forms. Type 1 refers to the case where the minimum is zero and type 2 for the case of the minimum being 1.
`N`	This is for the binomial distribution only, specifying the total number of successes. If NULL, it is sestimated by the data. It can also be a vector of successes.
`tol`	The tolerance level up to which the maximisation stops set to 1e-07 by default.

Details

Instead of maximising the log-likelihood via a numerical optimiser we used a Newton-Raphson algorithm which is faster.

See wikipedia for the equation to be solved in the case of the zero inflated distribution. https://en.wikipedia.org/wiki/Zero-inflated_model. In order to avoid negative values we have used link functions, log for the $lambda$ and logit for the π as suggested by Lambert (1992). As for the zero truncated Poisson see https://en.wikipedia.org/wiki/Zero-truncated_Poisson_distribution.

Value

The following list is not inclusive of all cases. Different functions have different names. In general a list including:

`mess`	This is for the negbin.mle only. If there is no reason to use the negative binomial distribution a message will appear, otherwise this is NULL.
`iters`	The number of iterations required for the Newton-Raphson to converge.
`loglik`	The value of the maximised log-likelihood.
`prob`	The probability parameter of the distribution. In some distributions this argument might have a different name. For example, param in the zero inflated Poisson.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris `<mtsagris@uoc.gr>` and Sofia Piperaki `<sofiapip23@gmail.com>`.

References

Lambert D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. Technometrics. 34 (1): 1–14

Johnson N. L., Kotz S. and Kemp A. W. (1992). Univariate Discrete Distributions (2nd ed.). Wiley.

Skellam J. G. (1946) The frequency distribution of the difference between two Poisson variates belonging to different populations. Journal of the Royal Statistical Society, series A 109/3, 26.

Nikoloulopoulos A.K. and Karlis D. (2008). On modeling count data: a comparison of some well-known discrete distributions. Journal of Statistical Computation and Simulation, 78(3): 437–457.

See Also

real.mle

Examples

```
x <- rpois(100, 2)
disc.mle(x, type = "poisson")
```

`MLE of distributions defined in the (0, 1) interval`

MLE of distributions defined in the (0, 1) interval

Description

MLE of distributions defined in the (0, 1) interval.

Usage

```
prop.mle(x, distr = "beta", tol = 1e-07, maxiters = 50)
```

Arguments

`x`	A numerical vector with proportions, i.e. numbers in (0, 1) (zeros and ones are not allowed).
`distr`	The distribution to fit. "beta" stands for the beta distribution, "ibeta" for the inflated beta, (0-inflated or 1-inflated, depending on the data), "logitnorm" is the logistic normal and "hsecant01" stands for the hyper-secant.
`tol`	The tolerance level up to which the maximisation stops.
`maxiters`	The maximum number of iterations to implement.

Details

Maximum likelihood estimation of the parameters of the beta distribution is performed via Newton-Raphson. The distributions and hence the functions does not accept zeros. "logitnorm" fits the logistic normal, hence no nwewton-Raphson is required and the "hypersecant01" uses the golden ratio search as is it faster than the Newton-Raphson (less calculations). The distributions included are the Kumaraswamy, zero inflated logistic normal, simplex, unit Weibull and continuous Bernoulli and standard power. Instead of maximising the log-likelihood via a numerical optimiser we have used a Newton-Raphson algorithm which is faster. See wikipedia for the equations to be solved.

Value

A list including:

`iters`	The number of iterations required by the Newton-Raphson.
`loglik`	The value of the log-likelihood.
`param`	The estimated parameters. In the case of "hypersecant01.mle" this is called "theta" as there is only one parameter.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Kumaraswamy P. (1980). A generalized probability density function for double-bounded random processes. Journal of Hydrology 46(1-2): 79–88.

Jones M.C. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. Statistical Methodology, 6(1): 70–81.

J. Mazucheli, A. F. B. Menezes, L. B. Fernandes, R. P. de Oliveira and M. E. Ghitany (2020). The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates. Journal of Applied Statistics, 47(6): 954–974.

Leemis L.M. and McQueston J.T. (2008). Univariate Distribution Relationships. The American Statistician, 62(1): 45–53.

You can also check the relevant wikipedia pages.

See Also

colprop.mle, comp.mle

Examples

```
x <- rbeta(1000, 1, 4)
prop.mle(x, distr = "beta")
```

MLE of distributions for compositional data

MLE of distributions for compositional data

Description

MLE of distributions for compositional data.

Usage

```
comp.mle(x, distr = "diri", type = 1, a = NULL, tol = 1e-07)
```

Arguments

`x`	A matrix containing the compositional data. Zero values are not allowed except for the case of the ZAD which is designed for the case of zero values present.
`distr`	The distribution to fit. "diri" stands for the Dirichlet distribution, "zad" is the Zero Adjusted Dirichlet distribution and "afolded" for the α-folded model (Tsagris and Stewart, 2020).
`type`	This is for the Dirichlet distribution ("diri"). Type 1 uses a vectorised version of the Newton-Raphson (Minka, 2012). In high dimensions this is to be preferred. If the data are too concentrated, regardless of the dimensions, this is also to be preferrred. Type 2 uses the regular Newton-Raphson, with matrix multiplications. In small dimensions this can be considerably faster.
`a`	The value of α. If this is NULL, the function will estimate it internally.
`tol`	The tolerance level idicating no further increase in the log-likelihood.

Details

Maximum likelihood estimation of the parameters of a Dirichlet distribution is performed via Newton-Raphson. Initial values suggested by Minka (2012) are used.

Value

A list including:

`loglik`	The value of the log-likelihood.
`param`	The estimated parameters.
`phi`	The precision parameter. If covariates are linked with it (function "diri.reg2"), this will be a vector.
`mu`	The mean vector of the distribution.
`runtime`	The time required by the MLE.
`best`	The estimated optimal α of the folded model.
`p`	The estimated probability inside the simplex of the folded model.
`mu`	The estimated mean vector of the folded model.
`su`	The estimated covariance matrix of the folded model.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Minka Thomas (2012). Estimating a Dirichlet distribution. Technical report.

Ng Kai Wang, Guo-Liang Tian, and Man-Lai Tang (2011). Dirichlet and related distributions: Theory, methods and applications. John Wiley & Sons.

Tsagris M. and Stewart C. (2018). A Dirichlet regression model for compositional data with zeros. Lobachevskii Journal of Mathematics, 39(3): 398–412. Preprint available from https://arxiv.org/pdf/1410.5011.pdf

Tsagris M. and Stewart C. (2022). A Review of Flexible Transformations for Modeling Compositional Data. In Advances and Innovations in Statistics and Data Science, pp. 225–234. https://link.springer.com/chapter/10.10 3-031-08329-7_10

Tsagris M. and Stewart C. (2020). A folded model for compositional data analysis. Australian and New Zealand Journal of Statistics, 62(2): 249–277. https://arxiv.org/pdf/1802.07330.pdf

See Also

prop.mle

Examples

```
x <- matrix( rgamma(100 * 4, c(5, 6, 7, 8), 1), ncol = 4)
x <- x / rowSums(x)
res <- comp.mle(x)
```

MLE of some censored models

MLE of some censored models

Description

MLE of some censored models.

Usage

```
cens.mle(x, distr = "tobit", di, tol = 1e-07)
```

Arguments

x	A vector with positive valued data and zero values. If there are no zero values, a simple normal model is fitted in the end.
distr	The distribution to fit. "tobit" stands for the tobit model, "censweibull" for the censored Weibull and "censpois" for the left censored Poisson. For the "censpois" the lowest value in x is taken as the censored point and values below that number are considered to be censored.
di	A vector of 0s (censored) and 1s (not censored) values.
tol	The tolerance level up to which the maximisation stops; set to 1e-07 by default.

Details

The tobin model is useful for (univariate) positive data with left censoring at zero. There is the assumption of a latent variable. Tthe values of that variable which are positive concide with the observed values. If some values are negative, they are left censored and the observed values are zero. Instead of maximising the log-likelihood via a numerical optimiser we have used a Newton-Raphson algorithm which is faster.

Value

A list including:

`iters`	The number of iterations required for the Newton-Raphson to converge.
`loglik`	The value of the maximised log-likelihood.
`param`	The vector of the parameters.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Tobin James (1958). Estimation of relationships for limited dependent variables. Econometrica. 26(1):24–36.

https://en.wikipedia.org/wiki/Tobit_model

Fritz Scholz (1996). Maximum Likelihood Estimation for Type I Censored Weibull Data Including Covariates. Technical report. ISSTECH-96-022, Boeing Information & Support Services, P.O. Box 24346, MS-7L-22.

See Also

`colcens.mle`, `positive.mle`, `truncmle`

Examples

```
x <- rnorm(300, 3, 5)
x[ x < 0 ] <- 0   ## left censoring. Values below zero become zero
cens.mle(x, distr = "tobit")

x1 <- rpois(10000, 15)
x <- x1
x[x <= 10] <- 10
mean(x) ## simple Poisson
cens.mle(x, distr = "censpois")$lambda
```

MLE of some circular distributions

MLE of some circular distributions

Description

MLE of some circular distributions.

Usage

```
circ.mle(x, rads = FALSE, distr = "vm", N = 2, ina, tol = 1e-07, maxiters = 100)
```

Arguments

x	A numerical vector with the circular data. They must be expressed in radians. If distr is "spml" or "purka" this can also be a matrix with two columns, the cosinus and the sinus of the circular data.
rads	If the data are in radians set this to TRUE.
distr	The type of distribution to fit, "vm" stands for the von Mises, "spml" is the angular Gaussian, "purka" is the Purkayastha, and "wrapcauchy" is the wrapped Cauchy distribution, "circexp" and "circbeta" stand for the circular exponential and the circular beta distributions, respectively. "cardio" is the cardioid distribution and "ggvm" is the generalized von Mises distribution, "cipc" is the circular independent projected Cauchy, "gcpc" is the generalised circular projected Cauchy distribution and "mmvm" is the multi-modal von Mises distribution. "multivm" and "multispml" denote the von Mises and the angular Gaussian but for multiple samples.
N	The number of modes to consider in the multi-modal von Mises distribution.
ina	A numerical vector with discrete numbers starting from 1, i.e. 1, 2, 3, 4,... or a factor variable. Each number denotes a sample or group. If you supply a continuous valued vector the function will obviously provide wrong results. This is only for "multivm" and "multispml".
tol	The tolerance level to stop the iterative process of finding the MLEs.
maxiters	The maximum number of iterations to implement.

Details

The parameters of the bivariate angular Gaussian, wrapped Cauchy, circular exponential, cardioid, circular beta, geometrically generalised von Mises, CIPC (reparametrised version of the wrapped Cauchy), GCPC (generalisation of the CIPC) and multi-modal von Mises distributions are estimated. For the Wrapped Cauchy, the iterative procedure described by Kent and Tyler (1988) is used. The Newton-Raphson algortihm for the angular Gaussian is described in the regression setting in Presnell et al. (1998). The circular exponential is also known as wrapped exponential distribution.

Value

A list including:

- `iters` The iterations required until convergence. This is returned in the wrapped Cauchy distribution only.
- `param` A vector consisting of the estimates of the two parameters, the mean direction for both distributions and the concentration parameter κ and the ρ for the von Mises (and the multi-modal von Mises) and wrapped Cauchy respectively. For the circular beta this contains the mean angle and the α and β parameters. For the cardioid distribution this contains the μ and ρ parameters. For the generalised von Mises this is a vector consisting of the ζ, κ, μ and α parameters of the generalised von Mises distribution as described in Equation (2.7) of Dietrich and Richter (2017).
- `gamma` The norm of the mean vector of the angular Gaussian, the CIPC and the GCPC distributions.
- `mu` The mean vector of the angular Gaussian, the CIPC and the GCPC distributions.
- `mumu` In the case of "angular Gaussian distribution this is the mean angle in radians.
- `circmu` In the case of the CIPC and the GCPC this is the mean angle in radians.
- `rho` For the GCPC distribution this is the eigenvalue of the covariance matrix, or the covariance determinant.
- `lambda` The lambda parameter of the circular exponential distribution.
- `theta` The median direction of the Purkayastha distribution.
- `alpha` The concentration parameter of the Purkayastha distribution.
- `alpha.sd` The standard error of the concentration parameter of the Purkayastha distribution.
- `loglik` The log-likelihood.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Mardia K. V. and Jupp P. E. (2000). Directional statistics. Chicester: John Wiley & Sons.

Sra S. (2012). A short note on parameter approximation for von Mises-Fisher distributions: and a fast implementation of $I_s(x)$. Computational Statistics, 27(1): 177–190.

Presnell Brett, Morrison Scott P. and Littell Ramon C. (1998). Projected multivariate linear models for directional data. Journal of the American Statistical Association, 93(443): 1068–1077.

Kent J. and Tyler D. (1988). Maximum likelihood estimation for the wrapped Cauchy distribution. Journal of Applied Statistics, 15(2): 247–254.

Dietrich T. and Richter W. D. (2017). Classes of geometrically generalized von Mises distributions. Sankhya B, 79(1): 21–59.

https://en.wikipedia.org/wiki/Wrapped_exponential_distribution

Jammalamadaka S. R. and Kozubowski T. J. (2003). A new family of circular models: The wrapped Laplace distributions. Advances and Applications in Statistics, 3(1), 77–103.

Tsagris M. and Alzeley O. (2023). Circular and spherical projected Cauchy distributions: A Novel Framework for Circular and Directional Data Modeling. https://arxiv.org/pdf/2302.02468.pdf

Barnett M. J. and Kingston R. L. (2024). A note on the Hendrickson-Lattman phase probability distribution and its equivalence to the generalized von Mises distribution. Journal of Applied Crystallography, 57(2).

Purkayastha S. (1991). A Rotationally Symmetric Directional Distribution: Obtained through Maximum Likelihood Characterization. The Indian Journal of Statistics, Series A, 53(1): 70–83.

Cabrera J. and Watson G. S. (1990). On a spherical median related distribution. Communications in Statistics-Theory and Methods, 19(6): 1973–1986

See Also

colcirc.mle,

Examples

```
y <- rcauchy(100, 3, 1)
x <- y
res <- circ.mle(x, distr = "wrapcauchy")
```

MLE of some continuous multivariate distributions

MLE of some continuous multivariate distributions

Description

MLE of some continuous multivariate distributions.

Usage

```
mv.mle(x, distr = "mvnorm", v = 1, tol = 1e-7)
```

Arguments

x	A matrix with numerical data.
distr	The distribution to fit. "mvnorm" stands for the multivariate normal distribution, "mvlnorm" for the multivariate log-normal, "mvt" is the multivariate t distribution and "invdir" stands for the inverse Dirichlet distribution. If you want the multivariate Cauchy distribution, simply choose "mvt" and set the v argument equal to 1.
v	The degrees of freedom. Must be a positive number, greater than zero.
tol	The tolerance value to terminate the EM algorithm.

Details

The mean vector, covariance matrix and the value of the log-likelihood of the multivariate normal or log-normal distribution is calculated. For the log-normal distribution we also provide the expected value and the covariance matrix. The location vector, scatter matrix and the value of the log-likelihood for the multivariate t distribution is calculated. Maximum likelihood estimation of the parameters of the inverted is performed via Newton-Raphson.

Value

A list including:

`loglik`	The log-likelihood multivariate distribution.
`mu`	The mean vector.
`sigma`	The covariance matrix.
`m`	The expected mean vector of the multivariate log-normal distribution.
`s`	The expected covariance matrix of the multivariate log-normal distribution.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris `<mtsagris@uoc.gr>` and Sofia Piperaki `<sofiapip23@gmail.com>`.

References

Kotz S., Balakrishnan N. and Johnson N. L. (2004). Continuous multivariate distributions, Volume 1: Models and applications (Vol. 1). John wiley & sons.

Nadarajah S. and Kotz S. (2008). Estimation methods for the multivariate t distribution. Acta Applicandae Mathematicae, 102(1): 99–118.

Bdiri T. and Bouguila N. (2012). Positive vectors clustering using inverted Dirichlet finite mixture models. Expert Systems with Applications, 39(2): 1869–1882.

http://isi.cbs.nl/iamamember/CD2/pdf/329.PDF

https://en.wikipedia.org/wiki/Log-normal_distribution#Multivariate_log-normal

See Also

`real.mle`

Examples

```
x <- matrix( rnorm(100 * 5), ncol = 5)
res <- mv.mle(x)
```

MLE of some matrix distributions
MLE of some matrix distributions

Description

MLE of some matrix distributions some matrix distributions.

Usage

```
matrix.mle(X, distr = "MN")
```

Arguments

X	For the matrix normal, a list with k elements (k is the sample size), k matrices of dimension $n \times p$ each. For the matrix Fisher an array containing rotation matrices in SO(3).
distr	The distribution to fit. "MN" stands for the matrix normal, while "mfisher" stands for the matrix Fisher distribution (defined in SO(3)).

Value

For the matrix normal a list including:

runtime	The runtime required for the whole fitting procedure.
iters	The number of iterations required for the estimation of the U and V matrices.
M	The estimated mean matrix of the distribution, a numerical matrix of dimensions $n \times p$.
U	The estimated covariance matrix associated with the rows, a numerical matrix of dimensions $n \times n$.
V	The estimated covariance matrix associated with the columns, a numerical matrix of dimensions $p \times p$.

For the matrix Fisher the components of $svd(\bar{X})$.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Pocuca N., Gallaugher M. P., Clark K. M. and McNicholas P. D. (2019). Assessing and Visualizing Matrix Variate Normality. arXiv:1910.02859.

https://en.wikipedia.org/wiki/Matrix_normal_distribution#Definition

Prentice M. J. (1986). Orientation statistics without parametric assumptions. Journal of the Royal Statistical Society. Series B (Methodological), 48(2): 214–222.

See Also

mv.mle, hspher.mle

Examples

```
## silly example
n <- 8  ;  p <- 4
X <- list()
for ( i in 1:200 )
X[[ i ]] <- matrix( rnorm(n * p), ncol = p )
mod <- matrix.mle(X)
```

MLE of some truncated distributions

MLE of some truncated distributions

Description

MLE of some truncated distributions.

Usage

```
truncmle(x, distr = "trunccauchy", a, b, tol = 1e-07)
```

Arguments

x	A numerical vector with continuous data. For the Cauchy distribnution, they can be anywhere on the real line. For the exponential distribution they must be strictly positive.
distr	The type of distribution to fit, "trunccauchy" and "truncexpmle" stand for the truncated Cauchy and truncated exponential distributions, respectively.
a	The lower value at which the Cauchy distribution is truncated.
b	The upper value at which the Cauchy or the exponential distribution is truncated. For the exponential this must be greater than zero.
tol	The tolerance value to terminate the fitting algorithm.

Details

Maximum likelihood of some truncated distributions is performed.

Value

A list including:

iters	The number of iterations reuired by the Newton-Raphson algorithm.
loglik	The log-likelihood.
lambda	The λ parameter in the exponential distribution.
param	The location and scale parameters in the Cauchy distribution.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

David Olive (2018). Applied Robust Statistics (Chapter 4).

http://lagrange.math.siu.edu/Olive/ol-bookp.htm

See Also

cens.mle

Examples

```
x <- rnorm(500)
truncmle(x, a = -1, b = 1)
```

MLE of the ordinal model without covariates

MLE of the ordinal model without covariates

Description

MLE of the ordinal model without covariates.

Usage

```
ordinal.mle(y, link = "logit")
```

Arguments

y	A numerical vector with values 1, 2, 3,..., not zeros, or an ordered factor.
link	This can either be "logit" or "probit". It is the link function to be used.

Details

Maximum likelihood of the ordinal model (proportional odds) is implemented. See for example the "polr" command in R or the examples.

Value

A list including:

loglik	The log-likelihood of the model.
a	The intercepts (threshold coefficients) of the model.

Author(s)

Michail Tsagris and Sofia Piperaki.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr> and Sofia Piperaki <sofiapip23@gmail.com>.

References

Agresti, A. (2002) Categorical Data. Second edition. Wiley.

See Also

colordinal.mle

Examples

```
y <- factor( rbinom(100,3,0.5), ordered = TRUE )
res <- ordinal.mle(y)
res <- ordinal.mle(y, link = "probit")
```

Index