Package 'NovelDistns'

October 12, 2022

Type Package

Title Computes PDF, CDF, Quantile, Random Numbers and Measures of Inference for 3 General Families of Distributions

Version 0.1.0

Author Mutua Kilai, Gichuhi Waititu, Wanjoya Kibira

Maintainer Mutua Kilai <kilaimutua@gmail.com>

Description Computes the probability density function, the cumulative density function, quantile function, random numbers and measures of inference for the following families exponentiated generalized gull alpha power family, exponentiated gull alpha powerfamily, gull alpha power family.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports AdequacyModel, gsl, rootSolve, stats

Depends R (>= 2.10)

NeedsCompilation no

Repository CRAN

Date/Publication 2022-05-13 08:40:10 UTC

R topics documented:

bladderdata *Bladder Cancer data*

Description

A data set containing remission time in months of a sample of 128 bladder cancer patients

Usage

```
data("bladderdata")
```

Format

A data frame with 128 observations on the following variable.

`time` a numeric vector

Source

E. T. Lee and J. Wang, Statistical Methods for Survival Data Analysis, vol. 476, John Wiley & Sons, Hoboken, NJ, USA, 2003.

Examples

```
data(bladderdata)
## maybe str(bladderdata) ; plot(bladderdata) ...
```

egap *Exponentiated Gull Alpha Power Family of distribution*

Description

Computes the pdf, cdf, quantile, and random numbers and estimates the parameters of the exponentiated gull alpha power family of distribution specified by the cdf.

$$F(x, \Theta) = \left[\frac{\alpha G(x)}{\alpha^{G(x)}}\right]^b$$

where θ is the baseline family parameter vector. Also, b>0 are the extra parameters induced to the baseline cumulative distribution function (cdf) `G` whose pdf is g. Here, the baseline `G` refers to the cdf of: exponential, rayleigh and weibull.

Usage

```
regap(n, dist, param)
qegap(p, dist, param, log.p = FALSE, lower.tail = TRUE)
pegap(data, dist, param, log.p = FALSE, lower.tail = TRUE)
degap(data, dist, param, log = FALSE)
mlegap(data, dist,starts, method="SANN")
```

Arguments

n	number of realizations to be generated.
p	quantile value between 0 and 1.
data	Vector of observations.
param	parameter vector $\Theta = (b, \theta, \alpha)$
log	If TRUE, then log(pdf) is returned.
log.p	If TRUE, then log(cdf) is returned and quantile is computed for exp(-p).
lower.tail	If FALSE, then 1-cdf is returned and quantile is computed for 1-p.
dist	The name of family's pdf including: "exponential", "rayleigh", "weibull", "lomax"
method	the method for optimizing the log likelihood function. It can be one of "Nelder-Mead", "BFGS", "CG", "L-BFGS-B" or "SANN". The default is "BFGS". The details of these methods can be found in the manual pages for optim
starts	initial values of (theta, b, alpha)

Value

1. A vector of the same length as data, giving the pdf values computed at data.
2. A vector of the same length as data, giving the cdf values computed at data.
3. A vector of the same length as p, giving the quantile values computed at p.
4. A vector of the same length as n, giving the random numbers realizations.
5. A sequence of goodness-of-fit statistics such as: Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), Bayesian Information Criterion (BIC), Hannan-Quinn information criterion (HQIC), Cramer-von Misses statistic (CM), Anderson Darling statistic (AD), log-likelihood statistic (log). The Kolmogorov-Smirnov (KS) test statistic and corresponding p-value and the convergence status.

Author(s)

Mutua Kilai, Gichuhi A. Waititu, Wanjoya A. Kibira

Examples

```
x=runif(10,min=0,max=1)
regap(10,"exp",c(0.3,0.5,0.7))
qegap(0.6,"exp",c(0.3,0.5,0.7))
pegap(x,"exp",c(0.3,0.5,0.7))
degap(x,"exp",c(0.3,0.5,0.7))
mlegap(x,"exp",c(0.3,0.5,0.7))
```

eggap *Exponentiated Generalized Gull Alpha Power Family of distribution*

Description

Computes the pdf, cdf, quantile, and random numbers and estimates the parameters of the exponentiated G gull alpha power family of distribution due to Kilai et al. (2022) specified by the cdf.

$$F(x,\Theta) = \left[1 - \left(1 - \frac{\alpha G(x)}{\alpha^{G(x)}}\right)^{a}\right]^{b}$$

where θ is the baseline family parameter vector. Also, a>0, b>0 are the extra parameters induced to the baseline cumulative distribution function (cdf) G whose pdf is g. Here, the baseline G refers to the cdf of: exponential, rayleigh and weibull.

Usage

```
reggap(n, dist, param)
qeggap(p, dist, param, log.p = FALSE, lower.tail = TRUE)
peggap(data, dist, param, log.p = FALSE, lower.tail = TRUE)
deggap(data, dist, param, log = FALSE)
mleggap(data, dist,starts, method="SANN")
```

Arguments

n	number of realizations to be generated.
p	quantile value between 0 and 1.
data	Vector of observations.
param	parameter vector $\Theta = (a, b, \theta, \alpha)$
log	If TRUE, then log(pdf) is returned.
log.p	If TRUE, then log(cdf) is returned and quantile is computed for exp(-p).
lower.tail	If FALSE, then 1-cdf is returned and quantile is computed for 1-p.
dist	The name of family's pdf including: "exponential", "rayleigh", "weibull", "lomax"
method	the method for optimizing the log likelihood function. It can be one of "Nelder-Mead", "BFGS", "CG", "L-BFGS-B" or "SANN". The default is "BFGS". The details of these methods can be found in the manual pages for optim
starts	initial values of (theta, a, b, alpha)

Value

1. A vector of the same length as data, giving the pdf values computed at data.
2. A vector of the same length as data, giving the cdf values computed at data.
3. A vector of the same length as p, giving the quantile values computed at p.

4. A vector of the same length as n, giving the random numbers realizations.
5. A sequence of goodness-of-fit statistics such as: Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), Bayesian Information Criterion (BIC), Hannan-Quinn information criterion (HQIC), Cramer-von Misses statistic (CM), Anderson Darling statistic (AD), log-likelihood statistic (log). The Kolmogorov-Smirnov (KS) test statistic and corresponding p-value and the convergence status.

Author(s)

Mutua Kilai, Gichuhi A. Waititu, Wanjoya A. Kibira

References

Mutua Kilai et al (2022) A new generalization of Gull Alpha Power Family of distributions with application to modeling COVID-19 mortality rates, https://doi.org/10.1016/j.rinp.2022.105339.

Examples

```
x=runif(10,min=0,max=1)
reggap(10,"exp",c(0.3,0.5,0.7,0.8))
qeggap(0.6,"exp",c(0.3,0.5,0.7,0.8))
peggap(x,"exp",c(0.3,0.5,0.7,0.8))
deggap(x,"exp",c(0.3,0.5,0.7,0.8))
mleggap(x,"exp",c(0.3,0.5,0.7,0.8))
```

---

gap *Gull Alpha Power Family of distribution*

Description

Computes the pdf, cdf, quantile, and random numbers and estimates the parameters of the exponentiated gull alpha power family of distribution specified by the cdf.

$$F(x,\Theta) = \left[\frac{\alpha G(x)}{\alpha^{G(x)}}\right]$$

where θ is the baseline family parameter vector.Here, the baseline G refers to the cdf of: exponential, rayleigh and weibull.

Usage

```
rgap(n, dist, param)
qgap(p, dist, param, log.p = FALSE, lower.tail = TRUE)
pgap(data, dist, param, log.p = FALSE, lower.tail = TRUE)
dgap(data, dist, param, log = FALSE)
mlgap(data, dist,starts, method="SANN")
```

Arguments

n	number of realizations to be generated.
p	quantile value between 0 and 1.
data	Vector of observations.
param	parameter vector $\Theta = (\theta, \alpha)$
log	If TRUE, then log(pdf) is returned.
log.p	If TRUE, then log(cdf) is returned and quantile is computed for exp(-p).
lower.tail	If FALSE, then 1-cdf is returned and quantile is computed for 1-p.
dist	The name of family's pdf including: "exponential", "rayleigh", "weibull", "lomax"
method	the method for optimizing the log likelihood function. It can be one of "Nelder-Mead", "BFGS", "CG", "L-BFGS-B" or "SANN". The default is "BFGS". The details of these methods can be found in the manual pages for optim
starts	initial values of (theta, alpha)

Value

1. A vector of the same length as data, giving the pdf values computed at data.
2. A vector of the same length as data, giving the cdf values computed at data.
3. A vector of the same length as p, giving the quantile values computed at p.
4. A vector of the same length as n, giving the random numbers realizations.
5. A sequence of goodness-of-fit statistics such as: Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), Bayesian Information Criterion (BIC), Hannan-Quinn information criterion (HQIC), Cramer-von Misses statistic (CM), Anderson Darling statistic (AD), log-likelihood statistic (log). The Kolmogorov-Smirnov (KS) test statistic and corresponding p-value and the convergence status.

Author(s)

Mutua Kilai, Gichuhi A. Waititu, Wanjoya A. Kibira

References

Muhammad et al (2020) A Gull Alpha Power Weibull distribution with applications to real and simulated data. https://doi.org/10.1371/journal.pone.0233080

Examples

```
x=runif(10,min=0,max=1)
rgap(10,"exp",c(0.3,0.5))
qgap(0.6,"exp",c(0.3,0.5))
pgap(x,"exp",c(0.3,0.5))
dgap(x,"exp",c(0.3,0.5))
mlgap(x,"exp",c(0.3,0.5))
```

italydata *COVID-19 Mortality Rates for Italy*

Description

A data set containing COVID-19 mortality rates for Italy for a period of 59 days from 27 Feb 2020 to 27 April 2020.

Usage

```
data("italydata")
```

Format

A data frame with 59 observations on the following 2 variables.

`date` a character vector

`rate` a numeric vector

Source

https://covid19.who.int/

Examples

```
data(italydata)
## maybe str(italydata) ; plot(italydata) ...
```

jetairplane *Number of failures of Boeing Jets*

Description

A data set containing number of failures for air conditioning systems of jet airplane data.

Usage

```
data("jetairplane")
```

Format

A data frame with 212 observations on the following variable.

`failures` a numeric vector

Source

Exponentiated Kumaraswamy-Dagum distribution with applications to income and lifetime data

Examples

```
data(jetairplane)
## maybe str(jetairplane) ; plot(jetairplane) ...
```

---

kenyadata *COVID-19 daily cases for Kenya*

Description

A data set containing COVID-19 daily cases for Kenya for a period of 56 days from 28 March 2020 to 24 May 2020

Usage

```
data("kenyadata")
```

Format

A data frame with 58 observations on the following 2 variables.

`date` a character vector

`cases` a numeric vector

Source

https://covid19.who.int/

Examples

```
data(kenyadata)
## maybe str(kenyadata) ; plot(kenyadata) ...
```

ukdata *COVID-19 Mortality Rates for United Kingdom*

Description

A data set containing COVID-19 mortality rates for United Kingdom for a period of 76 days from 15 April 2020 to 30 June 2020

Usage

```
data("ukdata")
```

Format

A data frame with 76 observations on the following 2 variables.

`date` a character vector

`rate` a numeric vector

Source

https://covid19.who.int/

Examples

```
data(ukdata)
## maybe str(ukdata) ; plot(ukdata) ...
```

Index