Package 'weibull4'

January 25, 2021

Type Package

Title Fits Data into 4-Parameters Weibull Distribution

Version 1.0.0

Author Vitor Hugo Moreau

Maintainer Vitor Hugo Moreau `<vitorhmc@ufba.br>`

Description Performs a curve fit to 4-parameters Weibull distribution using Metropolis algorithm -
Markov chain-Monte Carlo method. Special usage for fitting COVID-
19 epidemic data on daily new cases and deaths. Also, builds the 4-
parameters Weibull distribution curve using given parameters (shape, scale, location and area).

License Creative Commons Attribution 4.0 International License

Encoding UTF-8

LazyData true

Depends R (>= 3.5.0)

RoxygenNote 7.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2021-01-25 20:40:02 UTC

R topics documented:

Brazil_COVID   *Brazil's COVID-19 number of daily new cases and deaths*

Description

Dataset for weibull4 package. This dataset contain the number of the Brazil's new daily cases and deaths for COVID-19 to be fitted to the 4-parameters Weibull distribution in weibull4 package.

Usage

```
data("Brazil_COVID")
```

Format

A data frame with 349 observations on the following 3 variables.

date Brazil_COVID$date

new_cases Brazil_COVID$new_cases

new_deaths Brazil_COVID$new_deaths

Details

x values are in Date format. So, xmax must be as.Date() too.

Source

"https://covid.ourworldindata.org/data/owid-covid-data.csv"

References

Roser M, Ritchie H, Ortiz-Ospina E, Hasel J. Coronavirus Pandemic (COVID-19) [Internet]. 2020 [cited 2020 May 15]. Available from: https://ourworldindata.org/coronavirus

Examples

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

Canada_COVID   *Canada's COVID-19 number of daily new cases and deaths*

Description

Dataset for weibull4 package. This dataset contain the number of the Canada's new daily cases and deaths for COVID-19 to be fitted to the 4-parameters Weibull distribution in weibull4 package.

Usage

```
data("Canada_COVID")
```

Format

A data frame with 349 observations on the following 3 variables.

date Canada_COVID$date

new_cases Canada_COVID$new_cases

new_deaths Canada_COVID$new_deaths

Details

x values are in Date format. So, xmax must be as.Date() too.

Source

"https://covid.ourworldindata.org/data/owid-covid-data.csv"

References

Roser M, Ritchie H, Ortiz-Ospina E, Hasel J. Coronavirus Pandemic (COVID-19) [Internet]. 2020 [cited 2020 May 15]. Available from: https://ourworldindata.org/coronavirus

Examples

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

---

likelihood   *Calculates the Log Likelihood*

Description

Likelihood is an Internal Function of the Weibull4 Package

Usage

```
likelihood(x, y, param, modes)
```

Arguments

x	Vector with the x values
y	Vector with the y values
param	Vector with shape, scale, location, area and SD parameters for calculating the log of Likelihood for the weibull4 package
modes	Sets unimodal (modes=1) or bimodal (modes=2) Weibull's distribution

Value

Unitary vector with the sum of the likelihood

Author(s)

Florian Hartig - Theoretical Ecology

References

https://theoreticalecology.wordpress.com/2010/09/17/metropolis-hastings-mcmc-in-r/

Examples

```
function (param)
{
    shape <- param[1]
    scale <- param[2]
    loc <- param[3]
    area <- param[4]
    sd <- param[5]
    pred <- weibull4(xi, shape, scale, loc, area)
    singlelikelihoods <- dnorm(yi, mean = pred, sd = sd, log = T)
    sumll <- sum(singlelikelihoods, na.rm = T)
    return(sumll)
  }
```

---

posterior   *Calculates the posterior distribution for Metropolis-MCMC*

Description

This is an internal function of the weibull4 package

Usage

```
posterior(x, y, param, modes)
```

Arguments

x	Vector with the x values
y	Vector with the y values
param	Vector containning shape, scale, location, area and SD parameters
modes	Sets unimodal (modes=1) or bimodal (modes=2) Weibull's distribution

Value

Vector containning the posterior distribution for Metropolis-MCMC

Author(s)

Florian Hartig - Theoretical Ecology

References

https://theoreticalecology.wordpress.com/2010/09/17/metropolis-hastings-mcmc-in-r/

Examples

```
function (param)
{
    return(likelihood(param) + prior(param))
  }
```

---

prior   *Calculates the Prior Distribution for Metropolis-MCMC*

Description

This is an internal function of the Weibull4 package

Usage

```
prior(param)
```

Arguments

param   A vector with shape, scale, location, area and SD parameters

Value

A vector with prior distribution for Metropolis-MCMC

Author(s)

Florian Hartig - Theoretical Ecology

References

https://theoreticalecology.wordpress.com/2010/09/17/metropolis-hastings-mcmc-in-r/

Examples

```
function (param)
{
    shape <- param[1]
    scale <- param[2]
    loc <- param[3]
    area <- param[4]
    sd <- param[5]
    shapeprior <- dunif(shape, min = 1, max = 5, log = T)
    scaleprior <- dnorm(scale, sd = scale/2, log = T)
    locprior <- dunif(loc, min = 1, max = loc * 2, log = T)
    areaprior <- dunif(area, min = area/2, max = area * 2, log = T)
    sdprior <- dunif(sd, min = 1, max = sd * 2, log = T)
    return(shapeprior + scaleprior + locprior + areaprior + sdprior)
  }
```

---

proposalfunction   *Proposal Distribution for Metropolis-MCMC*

Description

This is an internal function of the weibull4 package

Usage

```
proposalfunction(param)
```

Arguments

param   Vector containing shape, scale, location, area and SD parameters

Value

Vector containing proposal values for shape, scale, location, area and SD

Author(s)

Florian Hartig - Theoretical Ecology

References

https://theoreticalecology.wordpress.com/2010/09/17/metropolis-hastings-mcmc-in-r/

Examples

```
function (param)
{
    return(rnorm(5, mean = param, sd = param * 0.015))
  }
```

---

run_metropolis_MCMC   *Runs the Metropolis-MCMC algorithm for weibull4 package*

Description

This is an internal function of the weibull4 package

Usage

```
run_metropolis_MCMC(x, y, startvalue, iterations, modes)
```

Arguments

x	Vector with the x values
y	Vector with the y values
startvalue	Vector with starting shape, scale, location, area and SD values for Metropolis-MCMC calculations
iterations	Number of iterations to be performed in MCMC simulation
modes	Sets unimodal (modes=1) or bimodal (modes=2) Weibull's distribution

Value

Matrix with 5 columns and iterations rows with Markov chains for shape, scale, location, area and SD parameters)

Author(s)

Florian Hartig - Theoretical Ecology

References

https://theoreticalecology.wordpress.com/2010/09/17/metropolis-hastings-mcmc-in-r/

Examples

```
function (startvalue, iterations)
{
    chain <- array(dim = c(iterations + 1, 5))
    chain[1, ] <- startvalue
    for (i in 1:iterations) {
        proposal <- proposalfunction(chain[i, ])
        probab <- exp(posterior(proposal) - posterior(chain[i,
            ]))
        if (runif(1) < probab) {
            chain[i + 1, ] <- proposal
        }
        else {
            chain[i + 1, ] <- chain[i, ]
        }
    }
    return(chain)
  }
```

US_COVID   *US number of new COVID-19 cases*

Description

Dataset for weibull4 package. This dataset contain the number of the US' new daily cases of COVID-19 to be fitted to the 4-parameters Weibull distribution in weibull4 package with modes=2, in order to fit to the second wave of COVID-19 infections.

Usage

```
data("US_COVID")
```

Format

A data frame with 336 observations on the following 2 variables.

date US_COVID$date

new_cases US_COVID$new_cases

new_deaths US_COVID$new_deaths

Details

x values are in Date format. So, xmax must be as.Date()

Source

"https://covid.ourworldindata.org/data/owid-covid-data.csv"

References

Roser M, Ritchie H, Ortiz-Ospina E, Hasel J. Coronavirus Pandemic (COVID-19) [Internet]. 2020 [cited 2020 May 15]. Available from: https://ourworldindata.org/coronavirus

Examples

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

---

weibull4	*non-linear regression data fitter to the 4-parameters Weibull distribution*

Description

Input script for using the weibull4 module. Weibull4 fit data from daily new cases and deaths of an epidemic peak. It was firstly designed to model data from COVID-19

Usage

```
weibull4(x, y, shape=NA, scale=NA, loc=NA, area=NA, shape2=NA, scale2=NA, loc2=NA,
area2=NA, iter=1000, xmax=0, modes=1, modes2=1, split=NA)
```

Arguments

x	Vector: time data (may be date/time or numeric)
y	Vector: observed/measure event
shape	Starting value for Weibull's shape parameters. If it is NA, weibull4.fit will try to calculate it from x and y data.
scale	Starting value for Weibull's scale parameters. If it is NA, weibull4.fit will try to calculate it from x and y data.
loc	Starting value for Weibull's location parameters. If it is NA, weibull4.fit will try to calculate it from x and y data.
area	Starting value for Weibull's area parameters or the area under the PDF curve. If it is NA, weibull4.fit will try to calculate it from x and y data.
shape2	Starting value for the shape parameters of the second mode of the Weibull's distribution. If it is NA, weibull4.fit will try to calculate it from x and y data. It works only if modes=2.
scale2	Starting value for the scale parameters of the second mode of the Weibull's distribution. If it is NA, weibull4.fit will try to calculate it from x and y data. It works only if modes=2.
loc2	Starting value for the location parameters of the second mode of the Weibull's distribution. If it is NA, weibull4.fit will try to calculate it from x and y data. It works only if modes=2.
area2	Starting value for the area parameters of the second mode of the Weibull's distribution. If it is NA, weibull4.fit will try to calculate it from x and y data. It works only if modes=2.
iter	Number of iterations to perform Metropolis-MCMC.
split	The date (x axis) in which the data will be split to be calculated by two distinct distribution. This option was implemented to fit data feom the second wave of infections and deaths for COVID-19. With split set, it is possible to analyze curve pattern with up to 4 waves of infection, since both modes and modes2 are set to 2. In this case, the data will be analyzed as two bimodal Weibull distribution.

modes	Sets whether data defore the split date (if the split is not NA) may be fit with uni-modal or bimodal Weibull distribution. If split is not set, modes is the number of modes of the unique distribution to be used. This option was implemented to fit data feom the second wave of infections and deaths for COVID-19. Use modes=1 for unimodal distribution (single peak) and modes=2 for bimodal distribution (two peaks, with no or small valley between them).
modes2	Sets whether data after the split date (if split is not NA) may be fit with unimodal or bimodal Weibull distribution. If split is not set, modes2 will not be used. This option was implemented to fit data feom the second wave of infections and deaths for COVID-19. Use modes=1 for unimodal distribution (single peak) and modes=2 for bimodal distribution (two peaks).
xmax	Forecast date to be calculated after x data. It must be in the same format than x.

Details

This package was specially built to fit COVID-19 data on the number of daily new cases and deaths in countries. So x must be integer. Alternatively, Date format is allowed

Value

LIST containning:

Fit data	Matrix with x and y fitted data
Estimates	Matrix containning shape, scale, location, area and SD of the MEtropolis-MCMC in the row 1; and standard deviation for each parameter in the row 2
Markov chains	Matrix containning the Markov chains for shape, scale, location, area and SD parameters

Warning

This package is a secondary product of the reffered science paper. Please, note that there is no warrants or professional support on its use

Note

Comments, suggestions and doubts must be sent to vitorhmc@ufba.br

Author(s)

Vitor Hugo Moreau, Ph.D.

References

MOREAU, V. H. (2021) Using the Weibull distribution to model COVID-19 epidemic data, Model Assisted Statistic Applications, in press.

Examples

```
## Perform non-linear curve fitting with US' data for daily new deaths of COVID-19, with split date
## on Aug, 15th, bimodal distribution before and unimodal distribution after the split date.
## Examples below are with low number of iterations (1000), because of CRAN rules limitations.
## They will possibily give poor results. For best results set iter=10000.
fit <- weibull4(US_COVID$date, US_COVID$new_deaths, split=as.Date("2020-09-15"),
modes=2, modes2=1,iter=1000)
plot(US_COVID$date, US_COVID$new_deaths, ylab="US' daily new deaths", xlab="Date")
lines(fit[[1]][,1], fit[[1]][,2], col="red")
## Perform non-linear curve fitting with Canada's data for daily new cases of COVID-19 using two
## unimodal Weibull distribution, with split date on Aug,  1st
fit <- weibull4(Canada_COVID$date, Canada_COVID$new_cases, split=as.Date("2020-08-01"),
modes=1, modes2=1, iter=1000)
plot(Canada_COVID$date, Canada_COVID$new_cases, ylab="Canada's daily new cases", xlab="Date")
lines(fit[[1]][,1], fit[[1]][,2], col="red")
## Perform non-linear curve fitting with Brazil's data for daily new deaths of COVID-19 using a
## single bimodal Weibull distribution
fit <- weibull4(Brazil_COVID$date, Brazil_COVID$new_deaths, modes=2, iter=1000)
plot(Brazil_COVID$date, Brazil_COVID$new_deaths, ylab="Brazil's daily new deaths", xlab="Date")
lines(fit[[1]][,1], fit[[1]][,2], col="red")
```

---

weibull4.build	*weibull 4-parameters distribution building function*

Description

Builds data for a 4-parameters Weibull distribution of a given x data

Usage

```
weibull4.build(x=seq(0,1,length.out=10), shape=2.5, scale=1, loc=0, area=20,
shape2=5, scale2=2, loc2=6, area2=1, modes=1)
```

Arguments

x	Vector: data range for calculation of the Weibull distribution. If it is NULL, it will be set to seq(0,1,0.1)
shape	Weibull's shape parameter
scale	Weibull's scale parameter
loc	Weibull's location parameter
area	Weibull's area parameter: area under the PDF curve
shape2	second mode Weibull's shape parameter. It works only if modes=2.
scale2	second mode Weibull's scale parameter. It works only if modes=2.
loc2	second mode Weibull's location parameter. It works only if modes=2.
area2	second mode Weibull's area parameter. It works only if modes=2.

modes Sets whether data may be fit with unimodal or bimodal Weibull distribution. This option was implemented to fit data feom the second wave of infections and deaths for COVID-19. Use modes=1 for unimodal distribution (single peak) and modes=2 for bimodal distribution (two peaks).

Details

This package was specially built to fit COVID-19 data on the number of daily new cases and deaths in countries. So x must be integer. Alternatively, Date format is allowed

Value

Vector: f(x) Weibull distribution's ordinate

Warning

This package is a secondary product of the reffered science paper. Please, note that there is no warrants or professional support on its use

Note

Comments, suggestions and doubts must be sent to vitorhmc@ufba.br

Author(s)

Vitor Hugo Moreau, Ph.D.

References

MOREAU, V. H. (2021) Using the Weibull distribution to model COVID-19 epidemic data. Modeling Assited Statistics Applications, 2021, in press.

Examples

```
## Build a 4-parameters Weibull distribution with given parameters
weibull4.build(seq(1,100,1), shape=2.5, scale=30, loc=10, area=1000, modes=1)
## Build and plot 4-parameters Weibull distribution with given parameters
plot(seq(1,100,1), weibull4.build(seq(1,100,1), 2, 30, 10, 1), type="l")
## Build and plot 4-parameters Weibull distribution with a time series in the abscissa
Date <- seq(Sys.Date(), as.Date("2022-12-31"),1)
plot(Date, weibull4.build(Date, 1.6, 100, 100, 100), type="l")
## Build and plot a bimodal, 4-parameters Weibull distribution with given parameters
weibull4.build(seq(1,100,1), shape=2.5, scale=30, loc=10, area=1000,
shape2=2.5, scale2=60, loc2=40, area2=1000, modes=2)
```

weibull4.fit	*Weibull 4-parameters Metropolis-MCMC non-linear curve fitting function*

Description

Package to perform non-linear regression in data on the number of daily new cases and daily new deaths of COVID-19 and other epidemics to the 4-parameters Weibull distribution using Metropolis-Markov Chain-Monte Carlo Simulations (MCMC), as described in Moreau, 2021

Usage

```
weibull4.fit(x, y, shape=NA, scale=NA, loc=NA, area=NA, shape2=NA, scale2=NA, loc2=NA,
area2=NA, iter=1000, xmax=0, modes=1)
```

Arguments

x	Vector: time data (may be date/time or numeric)
y	Vector: observed/measure event
shape	Starting value for Weibull's shape parameters. If it is NA, weibull4.fit will try to calculate it from x and y data.
scale	Starting value for Weibull's scale parameters. If it is NA, weibull4.fit will try to calculate it from x and y data.
loc	Starting value for Weibull's location parameters. If it is NA, weibull4.fit will try to calculate it from x and y data.
area	Starting value for Weibull's area parameters or the area under the PDF curve. If it is NA, weibull4.fit will try to calculate it from x and y data.
shape2	Starting value for the shape parameters of the second mode of the Weibull's distribution. If it is NA, weibull4.fit will try to calculate it from x and y data. It works only if modes=2.
scale2	Starting value for the scale parameters of the second mode of the Weibull's distribution. If it is NA, weibull4.fit will try to calculate it from x and y data. It works only if modes=2.
loc2	Starting value for the location parameters of the second mode of the Weibull's distribution. If it is NA, weibull4.fit will try to calculate it from x and y data. It works only if modes=2.
area2	Starting value for the area parameters of the second mode of the Weibull's distribution. If it is NA, weibull4.fit will try to calculate it from x and y data. It works only if modes=2.
iter	Number of iterations to perform Metropolis-MCMC.
modes	Sets whether data may be fit with unimodal or bimodal Weibull distribution. This option was implemented to fit data feom the second wave of infections and deaths for COVID-19. Use modes=1 for unimodal distribution (single peak) and modes=2 for bimodal distribution (two peaks).
xmax	Forecast date to be calculated after x data. It must be in the same format than x.

Details

This package was specially built to fit COVID-19 data on the number of daily new cases and deaths in countries. So x must be integer. Alternatively, Date format is allowed

Value

LIST containning:

`Fit data`	Matrix with x and y fitted data
`Estimates`	Matrix containning shape, scale, location, area and SD of the MEtropolis-MCMC in the row 1; and standard deviation for each parameter in the row 2
`Markov chains`	Matrix containning the Markov chains for shape, scale, location, area and SD parameters

Warning

This package is a secondary product of the reffered science paper. Please, note that there is no warrants or professional support on its use

Note

Comments, suggestions and doubts must be sent to vitorhmc@ufba.br

Author(s)

Vitor Hugo Moreau, Ph.D

References

MOREAU, V. H. (2021) Forecast projections of COVID-19 pandemic by Weibull distribution from daily new cases and deaths data, Model Assisted Statistics and Applications. in press. Part of the MCMC code came from the Florian Hartig blog: "https://theoreticalecology.wordpress.com/2010/09/17/metropolis-hastings-mcmc-in-r/"

Examples

```
## Perform non-linear curve fitting with World's data for daily new deaths
## of COVID-19
## Examples below are with low number of iterations (1000), because of CRAN rules limitations.
## They will possibily give poor results. For best results set iter=10000.
fit <- weibull4(US_COVID$date, US_COVID$new_deaths, iter=1000)
plot(US_COVID$date, US_COVID$new_deaths, ylab="US' daily new deaths", xlab="Date")
lines(fit[[1]][,1], fit[[1]][,2], col="red")
## Perform non-linear curve fitting with Canada's data for daily new cases of
## COVID-19 using two unimodal Weibull distribution, with split date on Aug,  ## 1st
fit <- weibull4(Canada_COVID$date, Canada_COVID$new_cases, modes=1, iter=1000)
plot(Canada_COVID$date, Canada_COVID$new_cases, ylab="Canada's daily new cases", xlab="Date")
lines(fit[[1]][,1], fit[[1]][,2], col="red")
## Perform non-linear curve fitting with Brazil's data for daily new deaths of
## COVID-19 using a single bimodal Weibull distribution
```

```
fit <- weibull4(Brazil_COVID$date, Brazil_COVID$new_deaths, modes=2, iter=1000)
plot(Brazil_COVID$date, Brazil_COVID$new_deaths, ylab="Brazil's daily new deaths", xlab="Date")
lines(fit[[1]][,1], fit[[1]][,2], col="red")
```

Index