Package ‘fitmix’

July 22, 2025

Title Finite Mixture Model Fitting of Lifespan Datasets

Version 0.1.0

Description Fits the lifespan datasets of biological systems such as yeast, fruit flies, and other similar biological units with well-known finite mixture models introduced by Farewell V. (1982) <doi:10.2307/2529885> and Al-Hussaini et al. (2000) <doi:10.1080/00949650008812033>. Estimates parameter space fitting of a lifespan dataset with finite mixtures of parametric distributions. Computes the following tasks; 1) Estimates parameter space of the finite mixture model by implementing the expectation maximization (EM) algorithm. 2) Finds a sequence of four goodness-of-fit measures consist of Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Kolmogorov-Smirnov (KS), and log-likelihood (log-likelihood) statistics. 3)The initial values is determined by k-means clustering.

URL `https://github.com/guven-code/fitmix/`

License GPL-3

Encoding UTF-8

RoxygenNote 7.1.1.9001

Depends R (>= 3.1.0)

Imports stats

Suggests rmarkdown, knitr

NeedsCompilation no

Author Emine Guven [aut, cre] (ORCID: <`https://orcid.org/0000-0001-9324-0879`>), Hong Qin [aut]

Maintainer Emine Guven <`emine.guven33@gmail.com`>

Repository CRAN

Date/Publication 2021-04-19 08:50:04 UTC

Contents

`dmix` *The mixture distribution*

Description

Computing probability density function for the well-known mixture models.

Usage

```
dmix(lifespan, model, K, param)
```

Arguments

`lifespan`	Vector of samples
`model`	choice of one of the mixture models; `gompertz`, `log-logistics`, `log-normal`, and `weibull`.
`K`	number of components
`param`	Vector of weight ω, shape α, and scale β parameters.

Value

A vector of the same length as lifespan data, given the pdf of the one of the mixture models computed at `lifespan`.

Examples

```
lifespan<-seq(0,30,0.2)
K<-2
weight<-c(0.6,0.4)
alpha<-c(0.5,1)
beta<-c(1,0.5)
param<-c(weight,alpha,beta)
dmix(lifespan, "log-logistic", K, param)
```

`fitmixEM` *Fits lifespan data of time units with gompertz, log-logistics, log-normal, and weibull mixture models choice of one.*

Description

Fits lifespan data of time units with gompertz, log-logistics, log-normal, and weibull mixture models choice of one.

Usage

```
fitmixEM(lifespan, model, K, initial = FALSE, starts)
```

Arguments

`lifespan`	numeric vector of lifespan dataset
`model`	model name of the one of the well-known model: `gompertz`,`log-logistics`,`log-normal`, and `weibull`.
`K`	number of well-known model components.
`initial`	logical true or false
`starts`	numeric if initial sets to true

Details

Estimates parameters of the given mixture model implementing the expectation maximization (EM) algorithm. General form for the cdf of a statistical mixture model is given by a distribution f is a mixture of `K component` distributions of $f = (f_1, f_2, ..f_K)$ if

$$f(x) = \sum_{k=1}^{K} \lambda_k f_k(x)$$

with $\lambda_k > 0$, $\sum_k \lambda_k = 1$. This equation is a stochastic model, thus it allows to generate new data points; first picks a distribution of choice, with probablities by weight, then generates another observation according to the chosen distribution. In short represenated by, Z ~ Mult($\lambda_1, \lambda_2, ...\lambda_k$) and X|Z ~ `f_Z`, where `Z` is a discrete random variable which component X is drawn from.

The families considered for the cdf of Gompertz, Log-normal, Log-logistic, and Weibull.

Value

1.The return has three values; the first value is estimate, measures, and cluster.

2. The second value includes four different measurements of goodness-of-fit tests involving: Akaike Information Criterion (`AIC`), Bayesian Information Criterion (`BIC`), Kolmogorov-Smirnov (`KS`), and log-likelihood (`log.likelihood`) statistics.

3. The last value is the output of clustering vector.

References

Farewell, V. (1982). The Use of Mixture Models for the Analysis of Survival Data with Long-Term Survivors. *Biometrics, 38(4), 1041-1046.* doi:10.2307/2529885 McLachlan, G. J. and Peel, D. (2000) *Finite Mixture Models*, John Wiley \& Sons, Inc.

Essam K. Al-Hussaini, Gannat R. Al-Dayian & Samia A. Adham (2000) On finite mixture of two-component gompertz lifetime model, *Journal of Statistical Computation and Simulation, 67:1, 20-67*, DOI: 10.1080/00949650008812033

Examples

```
lifespan<-sample(1000)
fitmixEM(lifespan, "weibull", K = 2, initial = FALSE)
```

`pmix` *The mixture cumulative distribution*

Description

Computing cumulative distribution function for the well-known mixture models.

Usage

```
pmix(lifespan, model, K, param)
```

Arguments

`lifespan`	Vector of samples
`model`	choice of one of the mixture models; `gompertz`, `log-logistics`, `log-normal`, and `weibull`.
`K`	number of components
`param`	Vector of weight ω, shape α, and scale β parameters.

Value

A vector of the same length as lifespan data, given the cdf of the one of the mixture models computed at `lifespan`.

Examples

```
lifespan<-seq(0,30,0.2)
K<-2
weight<-c(0.5,0.5)
alpha<-c(0.5,1)
beta<-c(1,0.5)
param<-c(weight,alpha,beta)
pmix(lifespan, "log-logistic", K, param)
```

`rmix` *The mixture random generation for the well-known models*

Description

Random generation for the well-known mixture models with parameters `weigth`, `shape` and `scale`.

Usage

```
rmix(N, model, K, param)
```

Arguments

`N`	Number of inputs for the mixture random generation
`model`	Choice of one of the mixture models; `gompertz`, `log-logistics`, `log-normal`, and `weibull`.
`K`	Number of components
`param`	Vector of weight ω, shape α, and scale β parameters.

Value

Outputs of random generated vector lenght of N from the given mixture model.

Examples

```
N<-100
K<-2
weight<-c(0.5,0.5)
alpha<-c(0.5,1)
beta<-c(1,0.5)
param<-c(weight,alpha,beta)
rmix(N, "weibull", K, param)
```

Index