Package ‘ForestFit’

September 27, 2019

Type Package

Title Statistical Modelling with Applications in Forestry

Author Mahdi Teimouri

Maintainer Mahdi Teimouri <teimouri@aut.ac.ir>

Description

Developed for the following tasks. I) Computing the probability density function, cumulative distribution function, random generation, and estimating the parameters of the eleven mixture models including mixture of Birnbaum-Saunders, BurrXII, Chen, F, Frechet, gamma, Gompertz, log-logistic, log-normal, Lomax, and Weibull. II) Point estimation of the parameters of two- and three-parameter Weibull distributions. In the case of two-parameter, twelve methods consist of generalized least square type 1, generalized least square type 2, L-moment, maximum likelihood, logarithmic moment, moment, percentile, rank correlation, least square, weighted maximum likelihood, U-statistic, weighted least square are used and investigated methods for the three-parameter case are: maximum likelihood, modified moment type 1, modified moment type 2, modified moment type 3, modified maximum likelihood type 1, modified maximum likelihood type 2, modified maximum likelihood type 3, modified maximum likelihood type 4, moment, maximum product spacing, T-L moment, and weighted maximum likelihood. III) The Bayesian estimators of the three-parameter Weibull distribution developed by Green et al. (1994) <doi:10.2307/2533217>. IV) Estimating parameters of the three-parameter Weibull distribution fitted to grouped data using three methods including approximated maximum likelihood, expectation maximization, and maximum likelihood. V) Estimating the parameters of the gamma, log-normal, and Weibull mixture models fitted to the grouped data through the EM algorithm. VI) Estimating parameters of the nonlinear growth curve fitted to the height-diameter observations.

Encoding UTF-8

License GPL (>= 2)

Depends R(>= 3.3.0), ars

Repository CRAN

Version 0.4.1

Date 2019-09-26

NeedsCompilation no

Date/Publication 2019-09-27 08:40:02 UTC

R topics documented:

dmixture . . . 2
fitbayesJSB . . . 4
fitbayesWeibull . . . 5
fitgrouped . . . 7
fitgrowth . . . 8
fitmixture . . . 10
fitmixturegrouped . . . 12
fitWeibull . . . 14
pmixture . . . 16
rmixture . . . 18

Index **20**

`dmixture` *Computing probability density function of the well-known mixture models*

Description

Computes probability density function (pdf) of the mixture model. The general form for the pdf of the mixture model is given by

$$f(x,\Theta)=\sum_{j=1}^{K}\omega_j f_j(x,\theta_j),$$

where $\Theta=(\theta_1,\dots,\theta_K)^T$, is the whole parameter vector, θ_j for $j=1,\dots,K$ is the parameter space of the j-th component, i.e. $\theta_j=(\alpha_j,\beta_j)^T$, $f_j(.,\theta_j)$ is the pdf of the j-th component, and known constant K is the number of components. The vector of mixing parameters is given by $\omega=(\omega_1,\dots,\omega_K)^T$ where ω_js sum to one, i.e., $\sum_{j=1}^{K}\omega_j=1$. Parameters α_j and β_j are the shape and scale parameters of the j-th component or both are the shape parameters. In the latter case, the parameters α and β are called the first and second shape parameters, respectively. We note that the constants ω_js sum to one, i.e. $\sum_{j=1}^{K}\omega_j=1$. The families considered for each component include Birnbaum-Saunders, Burr type XII, Chen, F, Frechet, Gamma, Gompertz, Log-normal, Log-logistic, Lomax, skew-normal, and Weibull with pdf given by the following.

- Birnbaum-Saunders

$$f(x,\theta)=\frac{\sqrt{\frac{x}{\beta}}+\sqrt{\frac{\beta}{x}}}{2\alpha x}\phi\Biggl(\frac{\sqrt{\frac{x}{\beta}}-\sqrt{\frac{\beta}{x}}}{\alpha}\Biggr),$$

- Burr XII

$$f(x,\theta)=\alpha\beta x^{\alpha-1}\Bigl(1+x^{\alpha}\Bigr)^{-\beta-1},$$

- Chen

$$f(x,\theta)=\alpha\beta x^{\alpha}\exp\bigl(x^{\alpha}\bigr)\exp\Bigl\{-\beta\exp\bigl(x^{\alpha}\bigr)+\beta\Bigr\},$$

- F
$$f(x,\theta)=\frac{\Gamma\Bigl(\frac{\alpha+\beta}{2}\Bigr)}{\Gamma\bigl(\frac{\alpha}{2}\bigr)\Gamma\bigl(\frac{\beta}{2}\bigr)}\Bigl(\frac{\alpha}{\beta}\Bigr)^{\frac{\alpha}{2}}x^{\frac{\alpha}{2}-1}\Bigl(1+\frac{\alpha}{\beta}x\Bigr)^{-\frac{\alpha+\beta}{2}},$$
- Frechet
$$f(x,\theta)=\frac{\alpha}{\beta}\Bigl(\frac{x}{\beta}\Bigr)^{-\alpha-1}\exp\Bigl\{-\Bigl(\frac{x}{\beta}\Bigr)^{-\alpha}\Bigr\},$$
- gamma
$$f(x,\theta)=\bigl[\beta^{\alpha}\Gamma(\alpha)\bigr]^{-1}x^{\alpha-1}\exp\Bigl(-\frac{x}{\beta}\Bigr),$$
- Gompertz
$$f(x,\theta)=\beta\exp\bigl(\alpha x\bigr)\exp\Biggl\{\frac{\beta\exp\bigl(\alpha x\bigr)-1}{\alpha}\Biggr\},$$
- log-logistic
$$f(x,\theta)=\frac{\alpha}{\beta^{\alpha}}x^{\alpha-1}\left[\Bigl(\frac{x}{\beta}\Bigr)^{\alpha}+1\right]^{-2},$$
- log-normal
$$f(x,\theta)=\bigl(\sqrt{2\pi}\beta x\bigr)^{-1}\exp\Biggl\{-\frac{1}{2}\left(\frac{\log x-\alpha}{\beta}\right)^{2}\Biggr\},$$
- Lomax
$$f(x,\theta)=\frac{\alpha\beta}{(1+\alpha x)^{\beta+1}},$$
- skew-normal
$$f(x,\theta)=2\phi\Bigl(\frac{x-\alpha}{\beta}\Bigr)\Phi\Bigl(\lambda\frac{x-\alpha}{\beta}\Bigr),$$
- Weibull
$$f(x,\theta)=\frac{\alpha}{\beta}\Bigl(\frac{x}{\beta}\Bigr)^{\alpha-1}\exp\Bigl\{-\Bigl(\frac{x}{\beta}\Bigr)^{\alpha}\Bigr\},$$

where $\theta=(\alpha,\beta)$. In the skew-normal case, $\phi(.)$ and $\Phi(.)$ are the density and distribution functions of the standard normal distribution, respectively.

Usage

```
dmixture(x, g, K, param)
```

Arguments

`x`	Vector of observations.
`g`	Name of the family including "`birnbaum-saunders`", "`burrxii`", "`chen`", "`f`", "`Frechet`", "`gamma`", "`gompetrz`", "`log-normal`", "`log-logistic`", "`lomax`", "`skew-normal`", and "`weibull`".
`K`	Number of components.
`param`	Vector of the ω, α, β, and λ.

Details

For the skew-normal case, α, β, and λ are the location, scale, and skewness parameters, respectively.

Value

A vector of the same length as x, giving the pdf of the mixture model of families computed at x.

Author(s)

Mahdi Teimouri

Examples

```
x<-seq(0,20,0.1)
K<-2
weight<-c(0.6,0.4)
alpha<-c(1,2)
beta<-c(2,1)
param<-c(weight,alpha,beta)
dmixture(x, "weibull", K, param)
```

---

fitbayesJSB	*Estimating parameters of the Johnson's SB (JSB) distribution using the Bayesian approach*

Description

Suppose $y = (y_1, \ldots, y_n)^T$ denotes a vector of n independent observations coming from a four-parameter JSB distribution with pdf given by

$$f(y|\Theta) = \frac{\delta\lambda}{\sqrt{2\pi}(y-\xi)(\lambda+\xi-y)} \exp\left\{-\frac{1}{2}\left[\gamma + \delta\log\left(\frac{y-\xi}{\lambda+\xi-y}\right)\right]^2\right\},$$

where $\xi < y < \lambda + \xi$ and $\Theta = (\delta, \gamma, \lambda, \xi)^T$ with $\delta, \lambda > 0$, $-\infty < \gamma < \infty$, and $-\infty < \xi < \infty$. Using the Bayesian approach, we compute the Bayes' estimators of the JSB distribution parameters.

Usage

```
fitbayesJSB(y, n.burn=8000, n.simul=10000)
```

Arguments

y	Vector of observations.
n.burn	Length of the burn-in period, i.e., the point after which Gibbs sampler is supposed to attain convergence. By default n.burn is 8000.
n.simul	Total numbers of Gibbas sampler iterations. By default n.simul is 10,000.

Details

The Bayes' estimators are obtained by averaging on the all iterations between `n.burn` and `n.simul`.

Value

A list of objects in two parts as

1. Bayes' estimators of the parameters.
2. A sequence of four goodness-of-fit measures consist of Anderson-Darling (`AD`), Cram\'eer-von Misses (`CVM`), Kolmogorov-Smirnov (`KS`), and log-likelihood (`log-likelihood`) statistics.

Author(s)

Mahdi Teimouri

References

N. L. Johnson, 1949. Systems of frequency curves generated by methods of translation, *Biometrika*, 36, 149–176.

L. J. Norman, S. Kotz, and N. Balakrishnan, 1994. *Continuous Univariate Distributions*, volume I, John Wiley & Sons.

Examples

```
n<-20
xi<-0
delta<-2
gamma<-2
lambda<-20
z<-rnorm(n)
y<-xi+lambda/(1+exp(-(z-gamma)/delta))
fitbayesJSB(y, n.burn=50, n.simul=80)
```

`fitbayesWeibull` *Estimating parameters of the Weibull distribution using the Bayesian approach*

Description

Suppose $y = (y_1, \ldots, y_n)^T$ denotes a vector of n independent observations coming from a three-parameter Weibull distribution. Using the methodology given in Green et al. (1994), we compute the Bayes' estimators of the shape, scale, and location parameters.

Usage

```
fitbayesWeibull(y, n.burn=8000, n.simul=10000)
```

Arguments

`y`	Vector of observations.
`n.burn`	Length of the burn-in period, i.e., the point after which Gibbs sampler is supposed to attain convergence. By default `n.burn` is 8000.
`n.simul`	Total numbers of Gibbas sampler iterations. By default `n.simul` is 10,000.

Details

The Bayes' estimators are obtained by averaging on the all iterations between `n.burn` and `n.simul`.

Value

A list of objects in two parts as

1. Bayes' estimators of the parameters.
2. A sequence of four goodness-of-fit measures consist of Anderson-Darling (`AD`), Cram\'eer-von Misses (`CVM`), Kolmogorov-Smirnov (`KS`), and log-likelihood (`log-likelihood`) statistics.

Note

The methodology used here for computing the Bayes' estimator of the location parameter is different from that used by Green et al. (1994). This means that the location parameter is allowed to be any real value.

Author(s)

Mahdi Teimouri

References

E. J. Green, F. A. R. Jr, A. F. M. Smith, and W. E. Strawderman, 1994. Bayesian estimation for the three-parameter Weibull distribution with tree diameter data, *Biometrics*, 50(1), 254-269.

Examples

```
n<-20
alpha<-2
beta<-2
theta<-3
y<-rweibull(n,shape=alpha,scale=beta)+theta
fitbayesWeibull(y, n.burn=100, n.simul=200)
```

fitgrouped — *Estimating parameters of the Weibull distribution fitted to grouped data*

Description

Suppose a sample of n independent observations each follows a three-parameter Weibull distribution have been divided into m separate groups of the form $(r_{i-1}, r_i]$, for $i = 1, \ldots, m$. So, the likelihood function is given by

$$L(\Theta) = \frac{n!}{f_1!f_2!\ldots f_m!}\prod_{i=1}^{m}\Big[F\big(r_i|\Theta\big) - F\big(r_{i-1}|\Theta\big)\Big]^{f_i},$$

where the r_0 is the lower bound of the first group, r_m is the upper bound of the last group, and f_i is the frequency of observations within i-th group provided that $n = \sum_{i=1}^{m} f_i$. The cdf of a three-parameter Birnbaum-Saunders or Weibull distribution is given by $F(x;\Theta) = 1-\exp\Big\{-\Big(\frac{x-\mu}{\beta}\Big)^{\alpha}\Big\}$ where $\Theta = (\alpha, \beta, \mu)^T$.

Usage

```
fitgrouped(r, f, family, method1, starts, method2)
```

Arguments

`r`	A numeric vector of length $m+1$. The first element of r is lower bound of the first group and other m elements are upper bound of the m groups. We note that upper bound of the $(i-1)$-th group is the lower bound of the i-th group, for $i = 2, \ldots, m$. The lower bound of the first group and upper bound of the m-th group are chosen arbitrarily.
`f`	A numeric vector of length m containing the group's frequency.
`family`	Can be either `"weibull"` or `"birnbaum-saunders"`.
`method1`	A character string determining the method of estimation. It can be one of `"aml"`, `"em"` and `"ml"`. The short forms `"aml"`, `"em"`, and `"ml"` are described as follows. `""aml"` (for method of approximated maximum likelihood (aml)), `""em"` (for method of expectation maximization (em)), and `""ml"` (for method of maximum likelihood (ml)).
`starts`	A numeric vector of the initial values for the shape, scale, and location parameters, respectively.
`method2`	The method for optimizing the log-likelihood function. It invovles one of `"BFGS"`, `"Nelder-Mead"`, `"CG"`, `"L-BFGS-B"` or `"SANN"`.

Details

If the method is `"em"`, then the initial values (`"starts"`) and the log-likelihood optimizing method (`"method2"`) are ignored.

Value

A two-part list of objects given by the following:

1. Estimated parameters of the three-parameter Birnbaum-Saunders or Weibull distribution fitted to the gropued data.
2. A sequence of goodness-of-fit measures consist of Akaike Information Criterion (`AIC`), Consistent Akaike Information Criterion (`CAIC`), Bayesian Information Criterion (`BIC`), Hannan-Quinn information criterion (`HQIC`), Anderson-Darling (`AD`), Chi-square (`Chi-square`), Cram\'eer-von Misses (`CVM`), Kolmogorov-Smirnov (`KS`), and log-likelihood (`log-likelihood`) statistics.

Author(s)

Mahdi Teimouri

References

G. J. McLachlan and T. Krishnan, 2007. *The EM Algorithm and Extensions*, John Wiley & Sons.

A. P. Dempster, N. M. Laird, and D. B. Rubin, 1977. Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B (methodological)*, 1-38.

M. Teimouri and A. K. Gupta, 2012. Estimation Methods for the Gompertz–Makeham Distribution Under Progressively Type-I Interval Censoring Scheme, *National Academy Science Letters*, 35(3).

Examples

```
r<-c(0,1,2,3,4,10)
f<-c(2,8,12,15,4)
starts<-c(2,2,0)
fitgrouped(r,f,"birnbaum-saunders","em")
fitgrouped(r,f,"weibull","ml",starts,"CG")
```

---

fitgrowth — *Estimatinng the parametersof the fitted non-linear growth curve to the height-diameter(H-D) observations*

Description

Estimates the parameters of the nine well-known three-parameter non-linear curves fitted to the height-diameter observations. These nine models are given by the following.

- Richards (Richards(1959))
$$H = 1.3 + \beta_1 + \frac{\beta_2}{D + \beta_3},$$
- Gompertz (Winsor(1992))
$$H = 1.3 + \beta_1 e^{-\beta_2 e^{-\beta_3 D}},$$

- Hossfeld IV (Zeide(1993))
$$H = 1.3 + \frac{\beta_1}{1 + \frac{1}{\beta_2 D^{\beta_3}}},$$
- Korf (Flewelling and De Jong(1994))
$$H = 1.3 + \beta_1 e^{-\beta_2 D^{-\beta_3}},$$
- logistic (Pearl and Reed (1920))
$$H = 1.3 + \frac{\beta_1}{1 + \beta_2 e^{-\beta_3 D}},$$
- Prodan (Prodan(1968))
$$H = 1.3 + \frac{D^2}{\beta_1 D^2 + \beta_2 D + \beta_3},$$
- Ratkowsky (Ratkowsky(1990))
$$H = 1.3 + \beta_1 e^{-\frac{\beta_2}{D+\beta_3}},$$
- Sibbesen (Huang et al. (1992))
$$H = 1.3 + \beta_1 D^{\beta_2 D^{-\beta_3}},$$
- Weibull (Yang et al. (1978))
$$H = 1.3 + \beta_1 \left(1 - e^{-\beta_2 D^{\beta_3}}\right),$$

Usage

```
fitgrowth(h,d,model,starts)
```

Arguments

`h`	Vector of height observations.
`d`	Vector of diameter observations.
`model`	The name of the fitted model including `"chapman-richards"`, `"gompertz"`, `"hossfeldiv"`, `"korf"`, `"logistic"`, `"prodan"` , `"ratkowsky"`, `"Sibbesen"`, and `"weibull"`.
`starts`	A list of starting values for the parameters β_1, β_2, and β_3.

Value

A list of objects in four parts as

1. Estimated parameters and corresponding summaries including standard errors, computed t-statistics, and p-values.
2. Residuals.
3. Covariance matrix of the estimated model parameters (coefficients) $\hat{\beta}_1$, $\hat{\beta}_1$, and $\hat{\beta}_1$.
4. Residual standard error, i.e., $\hat{\sigma}$.
5. The hieght-diameter scatterplot superimposed by the fitted model.

Author(s)

Mahdi Teimouri

References

J. W. Flewelling and R. De Jong. (1994). Considerations in simultaneous curve fitting for repeated height-diameter measurements, *Canadian Journal of Forest Research*, 24(7), 1408-1414.

S. Huang, S. J. Titus, and D. P. Wiens. 1992. Comparison of nonlinear height±diameter functions for major Alberta tree species. *Canadian Journal of Forest Research*, 22, 1297-1304.

R. Pearl and L. J. Reed. (1920). On the rate of growth of the population of the United States since 1790 and its mathematical representation, *Proceedings of the National Academy of Sciences of the United States of America*, 6(6), 275.

M. Prodan. 1968. The spatial distribution of trees in an area. *Allg. Forst Jagdztg*, 139, 214-217.

D. A. Ratkowsky. 1990. *Handbook of nonlinear regression*, New York, Marcel Dekker, Inc.

F. J. Richards. 1959. A flexible growth function for empirical use. *Journal of Experimental Botany*, 10, 290-300.

S. B. Winsor. 1992. The Gompertz curve as a growth curve. *Proceedings of National Academic Science*, USA, 18, 1-8.

R. C. Yang, A. Kozak, J. H. G. Smith. 1978. The potential of Weibull-type functions as a flexible growth curves. *Canadian Journal of Forest Research*, 8, 424-431.

B. Zeide. 1993. Analysis of growth equation. *Forest Science*, 39, 594-616.

Examples

```
h<-c( 1.8,  2.7,  3.3,  2.4,  2.9,  3.4,  2.5,  3.2,  4.1,  2.7,
      2.7,  2.2,  4.1,  1.5,  2.6, 17.1,  3.0,  2.5,  3.8,  2.1,
      3.2,  2.5,  2.8,  2.2,  2.0,  2.4,  4.2,  2.6,  2.5,  3.7,
      2.2,  3.0,  3.2,  2.5,  3.1,  3.0,  9.7, 12.1,  2.0,  2.4,
      3.2,  2.4,  2.9,  2.4,  3.2,  2.0,  2.5, 12.8, 18.2)
d<-c(13.7, 16.8, 20.8, 13.5, 17.0, 16.5, 15.0, 40.9, 20.8, 18.0,
     16.0, 12.7, 26.4, 11.4, 16.8, 66.0, 18.3, 9.70, 19.8, 16.5,
     22.9, 15.2, 29.0, 22.4, 11.4, 22.9, 26.7, 19.3, 24.1, 22.4,
     13.5, 11.7, 19.3, 18.0, 19.6, 26.4, 72.1, 66.0, 11.7, 16.0,
     13.5, 15.2, 17.0, 12.4, 16.0, 11.4, 14.5, 63.0, 55.6)
starts<-c(18,0.0005,2)
fitgrowth(h,d,"weibull",starts=starts)
```

`fitmixture` *Estimating parameters of the well-known mixture models*

Description

Estimates parameters of the mixture model using the expectation maximization (EM) algorithm. General form for the cdf of a statistical mixture model is given by

$$F(x, \Theta) = \sum_{j=1}^{K} \omega_j F_j(x, \theta_j),$$

where $\Theta = (\theta_1, \dots, \theta_K)^T$, is the whole parameter vector, θ_j for $j = 1, \dots, K$ is the parameter space of the j-th component, i.e. $\theta_j = (\alpha_j, \beta_j)^T$, $F_j(.,\theta_j)$ is the cdf of the j-th component, and known constant K is the number of components. Parameters α and β are the shape and scale parameters or both are the shape parameters. In the latter case, the parameters α and β are called the first and second shape parameters, respectively. We note that the constants ω_js sum to one, i.e. $\sum_{j=1}^{K} \omega_j = 1$. The families considered for the cdf F include Birnbaum-Saunders, Burr type XII, Chen, F, Frechet, Gamma, Gompertz, Log-normal, Log-logistic, Lomax, skew-normal, and Weibull.

Usage

```
fitmixture(x, family, K, initial="FALSE", starts)
```

Arguments

`x`	Vector of observations.
`family`	Name of the family including: "`birnbaum-saunders`", "`burrxii`", "`chen`", "`f`", "`Frechet`", "`gamma`", "`gompetrz`", "`log-normal`", "`log-logistic`", "`lomax`", "`skew-normal`", and "`weibull`".
`K`	Number of components.
`initial`	The sequence of initial values including $\omega_1, \dots, \omega_K, \alpha_1, \dots, \alpha_K, \beta_1, \dots, \beta_K$. For skew normal case the vector of initial values of skewness parameters will be added. By default the initial values automatically is determind by k-means method of clustering.
`starts`	If "`initial=TRUE`", then sequence of the initial values must be given.

Details

It is worth noting that identifiability of the mixture models supposed to be held. For skew-normal case we have $\theta_j = (\alpha_j, \beta_j, \lambda_j)^T$ in which $-\infty < \alpha_j < \infty$, $\beta_j > 0$, and $-\infty < \lambda_j < \infty$, respectively, are the location, scale, and skewness parameters of the j-th component, see Azzalini (1985).

Value

1. The output has three parts, The first part includes vector of estimated weight, shape, and scale parameters.
2. The second part involves a sequence of goodness-of-fit measures consist of Akaike Information Criterion (`AIC`), Consistent Akaike Information Criterion (`CAIC`), Bayesian Information

Criterion (`BIC`), Hannan-Quinn information criterion (`HQIC`), Anderson-Darling (`AD`), Cram\'eer-von Misses (`CVM`), Kolmogorov-Smirnov (`KS`), and log-likelihood (`log-likelihood`) statistics.

3. The last part of the output contains clustering vector.

Author(s)

Mahdi Teimouri

References

A. Azzalini, 1985. A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, 12, 171-178.

A. P. Dempster, N. M. Laird, and D. B. Rubin, 1977. Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society Series B*, 39, 1-38.

M. Teimouri, S. Rezakhah, and A. Mohammdpour, 2018. EM algorithm for symmetric stable mixture model, *Communications in Statistics-Simulation and Computation*, 47(2), 582-604.

Examples

```
n<-50
K<-2
m<-10
weight<-c(0.3,0.7)
alpha<-c(1,2)
beta<-c(2,1)
param<-c(weight,alpha,beta)
x<-rmixture(n, "weibull", K, param)
fitmixture(x,"weibull", K, initial="FALSE")
```

`fitmixturegrouped` *Estimating parameters of the well-known mixture models fitted to the grouped data*

Description

Estimates parameters of the gamma, log-normal, and Weibull mixture models fitted to the grouped data using the expectation maximization (EM) algorithm. General form for the cdf of a statistical mixture model is given by

$$F(x,\Theta)=\sum_{k=1}^{K}\omega_k F_k(x,\theta_k),$$

where $\Theta=(\theta_1,\dots,\theta_K)^T$, is the whole parameter vector, θ_k for $k=1,\dots,K$ is the parameter space of the j-th component, i.e. $\theta_k=(\alpha_k,\beta_k)^T$, $F_j(.,\theta_j)$ is the cdf of the k-th component, and known constant K is the number of components. Parameters α and β are the shape and scale parameters. The constants ω_ks sum to one, i.e. $\sum_{k=1}^{K}\omega_k=1$. The families considered for the cdf F include Gamma, Log-normal, and Weibull. If a sample of n independent observations each follows a distribution with cdf F have been divided into m separate groups of the form $(r_{i-1}, r_i]$, for $i = 1, \dots, m$. So, the likelihood function of the observed data is given by

$$L(\Theta|f_1,\dots,f_m) = \frac{n!}{f_1!f_2!\dots f_m!}\prod_{i=1}^{m}\Big[\frac{F_i(\Theta)}{F(\Theta)}\Big]^{f_i},$$

where

$$F_i(\Theta) = \sum_{k=1}^{K}\omega_k\int_{r_{i-1}}^{r_i} f(x|\theta_k)dx,$$

$$F(\Theta) = \sum_{k=1}^{K}\omega_k f(x|\theta_k)dx,$$

in which $f(x|\theta_k)$ denotes the pdf of the j-th component. Using the the EM algorithm proposed by Dempster et al. (1977), we can solve $\partial L(\Theta|f_1,\dots,f_m)/\partial\Theta = 0$ by introducing two new missing variables.

Usage

```
fitmixturegrouped(family, r, f, K, initial="FALSE", starts)
```

Arguments

`family`	Name of the family including: "gamma", "log-normal", "skew-normal", and "weibull".
`r`	A numeric vector of length $m+1$. The first element of r is lower bound of the first group and other m elements are upper bound of the m groups. We note that upper bound of the $(i-1)$-th group is the lower bound of the i-th group, for $i = 2, \dots, m$. The lower bound of the first group and upper bound of the m-th group are chosen arbitrarily. If raw data are available, the smallest and largest observations are chosen for lower bound of the first group and upper bound of the m-th group, respectively.
`f`	A numeric vector of length m containing the group's frequency.
`K`	Number of components.
`initial`	The sequence of initial values including $\omega_1,\dots,\omega_K,\alpha_1,\dots,\alpha_K,\beta_1,\dots,\beta_K$. For skew normal case the vector of initial values of skewness parameters will be added. By default the initial values automatically is determind by k-means method of clustering.
`starts`	If "initial=TRUE", then sequence of the initial values must be given.

Details

Identifiability of the mixture models supposed to be held. For skew-normal mixture model the parameter vector of k-th component gets the form $\theta_k = (\alpha_k, \beta_k, \lambda_k)^T$ where α_k, β_k, and λ_k denote the location, scale, and skewness parameters, respectively.

Value

1. The output has two parts, The first part includes vector of estimated weight, shape, and scale parameters.
2. A sequence of goodness-of-fit measures consist of Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), Bayesian Information Criterion (BIC), Hannan-Quinn information criterion (HQIC), Anderson-Darling (AD), Cram\'eer-von Misses (CVM), Kolmogorov-Smirnov (KS), and log-likelihood (log-likelihood) statistics.

Author(s)

Mahdi Teimouri

References

G. J. McLachlan and P. N. Jones, 1988. Fitting mixture models to grouped and truncated data via the EM algorithm, *Biometrics*, 44, 571-578

Examples

```
n<-50
K<-2
m<-10
weight<-c(0.3,0.7)
alpha<-c(1,2)
beta<-c(2,1)
param<-c(weight,alpha,beta)
x<-rmixture(n, "weibull", K, param)
r<-seq(min(x),max(x),length=m+1)
D<-data.frame(table(cut(x,r,labels=NULL,include.lowest=TRUE,right=FALSE,dig.lab=4)))
f<-D$Freq
fitmixturegrouped("weibull",r,f,K,initial="FALSE")
```

---

fitWeibull — *Estimating parameters of the Weibull distribution through classical methods*

Description

Estimates the parameters of the two- and three-parameter Weibull model with pdf and cdf given by

$$f(x;\alpha,\beta,\theta)=\frac{\alpha}{\beta}\left(\frac{x-\theta}{\beta}\right)^{\alpha-1}\exp\biggl\{-\left(\frac{x-\theta}{\beta}\right)^{\alpha}\biggr\},$$

and

$$F(x;\alpha,\beta,\theta)=1-\exp\biggl\{-\left(\frac{x-\theta}{\beta}\right)^{\alpha}\biggr\},$$

where $x>\theta$, $\alpha>0$, $\beta>0$ and $-\infty<\theta<\infty$. Here, the parameters α, β, and θ are known in the literature as the shape, scale, and location, respectively. If $\theta=0$, then $f(x;\alpha,\beta)$ and $F(x;\alpha,\beta)$ in above are the pdf and cdf of a two-parameter Weibull distribution, respectively.

Usage

```
fitWeibull(mydata, location, method, starts)
```

Arguments

`mydata`	Vector of observations
`starts`	Initial values for starting the iterative procedures such as Newton-Raphson.
`location`	Either "TRUE" or "FALSE". If location="TRUE", then shift parameter will be considered; otherwise the shift parameter omitted.
`method`	Used method for estimating the parameters. In the two-parameter case, methods are "`greg1`" (for the method of generalized regression type 1), "`greg2`" (for the method of generalized regression type 2), "`lmoment`" (for the method of L-moment), "`ml`" (for the method of maximum likelihood (ML)), "`mlm`" (for the method of logarithmic moment), "`moment`" (for the method of moment), "`pm`" (for the method of percentile), "`rank`" (for the method of rank correlation), "`reg`" (for the method of least square), "`ustat`" (for the method of U-statistic), "`wml`" (for the method of weighted ML), and "`wreg`" (for the method of weighted least square). In three-parameter case the methods are "`mle`" (for the method of ML), "`modifiedmoment1`" (for the method of modified moment (MM) type 1), "`modifiedmoment2`" (for the method of MM type 2), "`modifiedmoment3`" (for the method of MM type 3), "`modifiedml1`" (for the method of modified ML type 1), "`modifiedml2`" (for the method of modified ML type 2), "`modifiedml3`" (for the method of modified ML type 3), "`modifiedml4`" (for the method of modified ML type 4), "`moment`" (for the method of moment), "`mps`" (for the method of maximum product spacing), "`tlmoment`" (for the method of T-L moment), and "`wml`" (for the method of weighted ML).

Details

For the method `wml`, all weights have been provided for sample size less that or equal to 100. This means that both methods `ml` and `wml` give the same estimates for samples of size larger than 100.

Value

A list of objects in two parts given by the following:

1. Estimated parameters for two- or three-parameter Weibull distribution.
2. A sequence of goodness-of-fit measures consist of Akaike Information Criterion (`AIC`), Consistent Akaike Information Criterion (`CAIC`), Bayesian Information Criterion (`BIC`), Hannan-Quinn information criterion (`HQIC`), Anderson-Darling (`AD`), Cram\'eer-von Misses (`CVM`), Kolmogorov-Smirnov (`KS`), and log-likelihood (`log-likelihood`) statistics.

Author(s)

Mahdi Teimouri

References

R. C. H. Cheng and M. A. Stephens, 1989. A goodness-of-fit test using Moran's statistic with estimated parameters, *Biometrika*, 76(2), 385-392.

C. A. Clifford and B. Whitten, 1982. Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution, *Communication in Statistics-Theory and Methods*, 11(23), 2631-2656.

D. Cousineau, 2009. Nearly unbiased estimators for the three-parameter Weibull distribution with greater efficiency than the iterative likelihood method, *British Journal of Mathematical and Statistical Psychology*, 62, 167-191.

G. Cran, 1988. Moment estimators for the 3-parameter Weibull distribution, *IEEE Transactions on Reliability*, 37(4), 360-363.

J. R. Hosking, 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics, *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1), 105-124.

Y. M. Kantar, 2015. Generalized least squares and weighted least squares estimation methods for distributional parameters, *REVSTAT-Statistical Journal*, 13(3), 263-282.

M. Teimouri and S. Nadarajah, 2012. A simple estimator for the Weibull shape parameter, *International Journal of Structural Stability and Dynamics*, 12(2), 2395-402.

M. Teimouri, S. M. Hoseini, and S. Nadarajah, 2013. Comparison of estimation methods for the Weibull distribution, *Statistics*, 47(1), 93-109.

F. Wang and J. B. Keats, 1995. Improved percentile estimation for the two-parameter Weibull distribution, *Microelectronics Reliability*, 35(6), 883-892.

L. Zhang, M. Xie, and L. Tang, 2008. On Weighted Least Squares Estimation for the Parameters of Weibull Distribution. *In: Pham H. (eds) Recent Advances in Reliability and Quality in Design*. Springer Series in Reliability Engineering. Springer, London.

Examples

```
n<-100
alpha<-2
beta<-2
theta<-3
mydata<-rweibull(n,shape=alpha,scale=beta)+theta
starts<-c(2,2,3)
fitWeibull(mydata, "TRUE", "mps", starts)
fitWeibull(mydata, "TRUE", "wml", starts)
fitWeibull(mydata, "FALSE", "mlm", starts)
fitWeibull(mydata, "FALSE", "ustat", starts)
```

---

pmixture *Computing cumulative distribution function of the well-known mixture models*

Description

Computes cumulative distribution function (cdf) of the mixture model. The general form for the cdf of the mixture model is given by

$$F(x,\Theta)=\sum_{j=1}^{K}\omega_j F(x,\theta_j),$$

where $\Theta=(\theta_1,\dots,\theta_K)^T$, is the whole parameter vector, θ_j for $j=1,\dots,K$ is the parameter space of the j-th component, i.e. $\theta_j=(\alpha_j,\beta_j)^T$, $F_j(.,\theta_j)$ is the cdf of the j-th component, and known constant K is the number of components. The vector of mixing parameters is given by $\omega=(\omega_1,\dots,\omega_K)^T$ where ω_js sum to one, i.e., $\sum_{j=1}^{K}\omega_j=1$. Parameters α and β are the shape and scale parameters or both are the shape parameters. In the latter case, the parameters α and β are called the first and second shape parameters, respectively. The families considered for each component include Birnbaum-Saunders, Burr type XII, Chen, F, Frechet, Gamma, Gompertz, Log-normal, Log-logistic, Lomax, skew-normal, and Weibull.

Usage

```
pmixture(x, g, K, param)
```

Arguments

x	Vector of observations.
g	Name of the family including: "birnbaum-saunders", "burrxii", "chen", "f", "frechet", "gamma", "gompetrz", "log-normal", "log-logistic", "lomax", "skew-normal", and "weibull".
K	Number of components.
param	Vector of the ω, α, β, and λ.

Details

For the skew-normal case, α, β, and λ are the location, scale, and skewness parameters, respectively.

Value

A vector of the same length as x, giving the cdf of the mixture model computed at x.

Author(s)

Mahdi Teimouri

Examples

```
x<-seq(0,20,0.1)
K<-2
weight<-c(0.6,0.4)
alpha<-c(1,2)
beta<-c(2,1)
param<-c(weight,alpha,beta)
pmixture(x, "weibull", K, param)
```

`rmixture` *Generating random realizations from the well-known mixture models*

Description

Generates iid realizations from the mixture model with pdf given by

$$f(x, \Theta) = \sum_{j=1}^{K} \omega_j f(x, \theta_j),$$

where K is the number of components, θ_j, for $j = 1, \ldots, K$ is parameter space of the j-th component, i.e. $\theta_j = (\alpha_j, \beta_j)^T$, and Θ is the whole parameter vector $\Theta = (\theta_1, \ldots, \theta_K)^T$. Parameters α and β are the shape and scale parameters or both are the shape parameters. In the latter case, parameters α and β are called the first and second shape parameters, respectively. We note that the constants ω_js sum to one, i.e., $\sum_{j=1}^{K} \omega_j = 1$. The families considered for the cdf f include Birnbaum-Saunders, Burr type XII, Chen, F, Fr\'echet, Gamma, Gompertz, Log-normal, Log-logistic, Lomax, skew-normal, and Weibull.

Usage

```
rmixture(n, g, K, param)
```

Arguments

`n`	Number of requested random realizations.
`g`	Name of the family including "`birnbaum-saunders`", "`burrxii`", "`chen`", "`f`", "`frechet`", "`gamma`", "`gompetrz`", "`log-normal`", "`log-logistic`", "`lomax`", "`skew-normal`", and "`weibull`".
`K`	Number of components.
`param`	Vector of the ω, α, β, and λ.

Details

For the skew-normal case, α, β, and λ are the location, scale, and skewness parameters, respectively.

Value

A vector of length n, giving a sequence of random realizations from given mixture model.

Author(s)

Mahdi Teimouri

Examples

```
n<-50
K<-2
weight<-c(0.3,0.7)
alpha<-c(1,2)
beta<-c(2,1)
param<-c(weight,alpha,beta)
rmixture(n, "weibull", K, param)
```

Index

dmixture, 2

fitbayesJSB, 4
fitbayesWeibull, 5
fitgrouped, 7
fitgrowth, 8
fitmixture, 10
fitmixturegrouped, 12
fitWeibull, 14

pmixture, 16

rmixture, 18