Package 'sepro'

May 16, 2022

Type Package

Title Separation Process

Version 0.1

Date 2022-05-08

Description Estimating of parameters in statistical distributions

URL https://github.com/hdrbv/sepro

BugReports https://github.com/hdrbv/sepro/issues

Depends R (>= 4.0.0)

Imports devtools, roxygen2, mixtools, readxl, MASS, survival, npsurv, lsei, fitdistrplus, foreach, miscTools, dplyr, ggplot2, stargazer, testthat, usethis, available, cachem, pkgdown

Encoding UTF-8

Language en-US

License Copyright (c) 2022 hdrbv. All rights reserved.

RoxygenNote 7.2.0

NeedsCompilation no

Author hdr bv [cre, aut]

Maintainer hdr bv <for_qa@ro.ru>

R topics documented:

`sepro-package` *Separation Process*

Description

Estimating of parameters in statistical distributions

Details

The DESCRIPTION file:

Package:	sepro
Type:	Package
Title:	Separation Process
Version:	0.1
Date:	2022-05-08
Authors@R:	person("hdr", "bv", email = "for_qa@ro.ru", role = c("cre","aut"))
Description:	Estimating of parameters in statistical distributions
URL:	https://github.com/hdrbv/sepro
BugReports:	https://github.com/hdrbv/sepro/issues
Depends:	R (>= 4.0.0)
Imports:	devtools, roxygen2, mixtools, readxl, MASS, survival, npsurv, lsei, fitdistrplus, foreach, miscTools, dply
Encoding:	UTF-8
Language:	en-US
License:	Copyright (c) 2022 hdrbv. All rights reserved.
RoxygenNote:	7.2.0
Author:	hdr bv [cre, aut]
Maintainer:	hdr bv <for_qa@ro.ru>

EM function from package Separation Process (sepro) - realisation of EM algorithm (Expectation–maximization algorithm)

Author(s)

NA

References

Dellaert, Frank (2002). "The Expectation Maximization Algorithm"

`EM` *EM Function*

Description

This function realized EM algorithm (Expectation-Maximization algorithm) for data clustering. In this case, for mixture models.

Usage

```
EM(x0, k)
```

Arguments

x0	input data (vector)
k	amount of clusters (mixture components)

Details

In initial step (Step_0) explorer must determine the initial gaussian mixture model parameters. Explorer can take any random parameters or (using surface analysis and some assumptions) take anothers.

In the first step (called E_Step) with function E_Step explorer calculate the posterior probabilities (which named "posterior.df") for each item of initial dataset.

In the second step (called M_Step) with function M_Step explorer update component parameters (using likelihood function).

After that, explorer repeat M_Step (100 times or until difference between current value of logarithm of likelihood function and previous value of logarithm of likelihood function will be less than 10^(-6) in other words: logarithm of likelihood function didn't change much).

In the end explorer has updated probabilities for each item and updated parameters for each distribution (explorer should look at E.Step and M.Step)

Value

Parameters for each distribution

Author(s)

hdrbv

Examples

```
For example we want to separate 2 Gaussian distributions
and estimate parameters of each one.
Let us assume that vector x1 - mixture of these distributions.
So we can use EM algorithm here:
EM1 <- sepro::EM(x0 = x0, k = 2).
```

---

plot_em *plot_em Function*

Description

This function created plot of mixture distributions as a result of separation process via EM algorithm (Expectation-Maximization algorithm)

Usage

```
plot_em(x0, EM)
```

Arguments

x0	input data (vector)
EM	application of EM function to input data

Value

Plot of mixture

Author(s)

hdrbv

Examples

```
For example we want to separate 2 Gaussian distributions,
estimate parameters of each one and plot distributions.
Let us assume that vector x1 - mixture of these distributions.
So we can use EM algorithm here:
EM1 <- sepro::EM(x0 = x0, k = 2).
And create plot:
plot_em(x0, EM1)
```

Index