

Package ‘HGLD’

December 11, 2017

Type Package

Title Fit the Hurdle Generalized Lambda Distribution

Version 0.1.0

Author Diego Marcondes

URL <http://github.com/dmarcondes/HGLD>

BugReports <https://github.com/dmarcondes/HGLD/issues>

Maintainer Diego Marcondes <dmarcondes@ime.usp.br>

Description Functions to perform the Hurdle Generalized Lambda Distribution Regression and to fit the Hurdle Generalized Lambda Distribution to data. The functions of this package are modifications of the functions of packages gamlss, GLDEX and GLDReg, so that most of the credit for the functions of this package must go to the authors and maintainers of those three packages.

License GPL-3

Imports ggplot2 (>= 2.2.1),
gridExtra (>= 2.1.1),
grid (>= 3.3),
gamlss.dist (>= 5.0-2),
cluster (>= 2.0.6)

Depends R (>= 3.3),
gamlss (>= 5.0-2),
GLDEX (>= 2.0.0.5),
GLDreg (>= 1.0.7)

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

Reference Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) arxiv1712.02183

R topics documented:

dhgld	2
diag.hgld	3
fit.hgld	5
healthcare	6
phgld	7
plot.fit.hgld	9
plot.reg.hgld	10
qhgld	11
qhgld.reg	12
reg.hgld	13
rhgld	15
Index	17

dhgld	<i>The Hurdle Generalized Lambda Distribution Family</i>
-------	--

Description

Density of the Hurdle Generalized Lambda Distribution.

Usage

```
dhgld(x, mixture = FALSE, lambda1, lambda2 = NULL, prob = NULL,
      param = "fmkl", inverse.eps = 1e-08, max.iterations = 500)
```

Arguments

x	Vector of data.
mixture	Whether to give the density of a mixture HGLD.
lambda1	A vector of length 5 with the five parameters of the HGLD or the first HGLD if mixture = TRUE.
lambda2	A vector of length 4 with the four parameters of the second HGLD.
prob	The cluster parameter for the mixture HGLD.
param	"fmkl" or "rs".
inverse.eps	Accuracy of calculation for the numerical determination of F(x), defaults to 1e-8.
max.iterations	Maximum number of iterations in the numerical determination of F(x), defaults to 500.

Details

If the parametrization of the RS or FMKL HGLD is not acceptable, the function returns NA. This function is based on the [GLDEX](#) package.

Value

The probability density of the continuous part of the HGLD.

References

Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) *arxiv1712.02183*

Su, S.; Fitting Single and Mixture of Generalized Lambda Distributions to Data via Discretized and Maximum Likelihood Methods: GLDEX in R. (2007), Journal of Statistical Software: *21* 9.

Examples

```
library(ggplot2)
set.seed(100)
data <- healthcare[sample(1:nrow(healthcare),100),]
fit <- fit.hgld(data$log_expense)
{ggplot(data.frame(x = data$log_expense[data$log_expense!=0]),aes(x = x)) +
stat_function(fun = function(x) dhgld(x = x,lambda1 = fit$par[,2],param = "rs"))}

#mixture
set.seed(100)
data <- c(rcauchy(50,location = 10),rep(0,30),rcauchy(50))
fit <- fit.hgld(data,TRUE)
{ggplot(data.frame(x = data[data!=0]),aes(x = x)) +
stat_function(fun = function(x) dhgld(x = x,mixture = TRUE,lambda1 = fit$par[1:5,2],
lambda2 = fit$par[6:9,2],prob = fit$par[10,2],param = "rs"))}
```

diag.hgld

*Diagnostic for the Hurdle Generalized Lambda Distribution fit***Description**

Diagnostic plots and measures of goodness-of-fit for a Hurdle Generalized Lambda Distribution fit.

Usage

```
diag.hgld(fit, facet = FALSE, facet.breaks, facet.labels, facet.ncol,
  trace = TRUE, no.test = 1000, len = floor(0.9 * length(fit$data[fit$data
    != 0])), alpha = 0.05, plotKS = TRUE, KS = TRUE)
```

Arguments

fit	An object of class <code>fit.hgld</code> .
facet	Whether the plots must be faceted for better visualization.
facet.breaks	The breaks in which to facet the data. Must be the endpoints of the intervals.
facet.labels	The labels of the categories given by the facet breaks.

facet.ncol	The number of columns for the facet plot.
trace	Whether progress bar must be printed in order to trace the algorithm.
no.test	Total number of KS tests required.
len	Number of data to sample at each KS test.
alpha	Significance level of KS test.
plotKS	Whether to plot the KS resample test result within each plot.
KS	Whether the resample KS test must be performed to the nonzero values.

Details

The diagnostics techniques are applied to the nonzero data values. Returns the qq-plot and the quantile plot between the data and the theoretical fitted HGLD. Also returns a table comparing sample moments with theoretical moments. A Kolmogorov-Simornov resample test is performed and the percentage of the times that the null hypotheses, i.e., goodness-of-fit, is not rejected is displayed. All diagnostics are performed for both the RS and FMKL GLD.

Value

qqRS	ggplot qq-plot for the RS HGLD.
qqFMKL	ggplot qq-plot for the FMKL HGLD.
quantRS	ggplot quantile plot for the RS HGLD.
quantFMKL	ggplot quantile plot for the FMKL HGLD.
moments	Moments comparison for the GLD fitted to the nonzero data for both parametrizations. Those are not the moments of the HGLD.
KS	Percentage of no rejection for the KS resample test.

References

Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) *arxiv1712.02183*

Su, S. Fitting Single and Mixture of Generalized Lambda Distributions to Data via Discretized and Maximum Likelihood Methods: GLDEX in R. (2007), Journal of Statistical Software: *21* 9.

Examples

```
set.seed(100)
data <- healthcare[sample(1:nrow(healthcare),100),]
fit <- fit.hgld(data$log_expense)
d <- suppressWarnings(diag.hgld(fit, facet = FALSE, plotKS = FALSE))

#mixture
set.seed(100)
data <- c(rcauchy(50, location = 10), rep(0, 30), rcauchy(50))
fit <- fit.hgld(data = data, mixture = TRUE)
d <- suppressWarnings(diag.hgld(fit))
```

fit.hgld

*Fit the Hurdle Generalized Lambda Distribution***Description**

Fit the Hurdle Generalized Lambda Distribution to a dataset by numerical maximum likelihood.

Usage

```
fit.hgld(data, mixture = FALSE, clustering.m = clara, threshold = NULL,
  leap1 = 3, leap2 = 3, fun1 = "runif.sobol", fun2 = "runif.sobol",
  rs.leap = 3, fmk1.leap = 3, rs.init = c(-1.5, 1.5),
  fmk1.init = c(-0.25, 1.5), FUN = "runif.sobol", no = 10000)
```

Arguments

data	A vector of data.
mixture	Whether a mixture of HGLD must be fitted.
clustering.m	Clustering method used in classifying the dataset into two parts when fitting a mixture of HGLD. Valid arguments include clara, fanny and pam from the cluster library, or threshold, if the data must be divided by a threshold. Default is clara. Or a logical vector specifying how data should be split. See fun.auto.bimodal.pml for more details.
threshold	The threshold to divide the data if <i>clustering.m</i> = "threshold".
leap1	Scrambling (0,1,2,3) for the sobol sequence for the first distribution fit when fitting a mixture of HGLD. See scrambling/leap argument for runif.sobol , runif.halton or QUnif of the GLDEX package.
leap2	Scrambling (0,1,2,3) for the sobol sequence for the second distribution fit when fitting a mixture of HGLD. See scrambling/leap argument for runif.sobol , runif.halton or QUnif of the GLDEX package.
fun1	A character string of either "runif.sobol" (default), "runif.halton" or "QUnif" for the first distribution fit when fitting a mixture of HGLD. See fun.auto.bimodal.pml for more details.
fun2	A character string of either "runif.sobol" (default), "runif.halton" or "QUnif" for the second distribution fit when fitting a mixture of HGLD. See fun.auto.bimodal.pml for more details.
rs.leap	Scrambling (0,1,2,3) for the sobol sequence for the RS generalized lambda distribution fit. See scrambling/leap argument for runif.sobol , runif.halton or QUnif of the GLDEX package. See fun.data.fit.ml for more details.
fmk1.leap	Scrambling (0,1,2,3) for the sobol sequence for the FMKL generalized lambda distribution fit. See scrambling/leap argument for runif.sobol , runif.halton or QUnif of the GLDEX package. See fun.data.fit.ml for more details.
rs.init	Initial values (lambda3 and lambda4) for the RS generalized lambda distribution. See fun.data.fit.ml for more details.

<code>fmkl.init</code>	Initial values (lambda3 and lambda4) for the FMKL generalized lambda distribution. See fun.data.fit.ml for more details.
<code>FUN</code>	A character string of either "runif.sobol" (default), "runif.halton" or "QUnif". See fun.data.fit.ml for more details.
<code>no</code>	Number of initial random values to find the best initial values for optimization. See fun.data.fit.ml for more details.

Details

Given a dataset, estimate by the numerical maximum likelihood method the five parameters of the HGLD. Fit both the RS and the FMKL parametrizations. Also fit mixture of HGLD.

Value

<code>par</code>	The estimate of the HGLD five parameters for both the RS and FMKL parametrizations if <i>mixture</i> = <i>FALSE</i> . Otherwise present the estimation of ten parameters: the zero probability mass, the four parameters of each GLD and the clustering parameter <i>p</i> .
<code>data</code>	The data used in the fit.
<code>mixture</code>	Whether a mixture of HGLD was fitted.

References

- Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) *arxiv1712.02183*
- Su, S.; Fitting Single and Mixture of Generalized Lambda Distributions to Data via Discretized and Maximum Likelihood Methods: GLDEX in R.(2007), Journal of Statistical Software: *21* 9.

Examples

```
set.seed(100)
data <- healthcare[sample(1:nrow(healthcare),100),]
fit <- fit.hgld(data$log_expense)

#mixture
set.seed(100)
data <- c(rcauchy(50,location = 10),rep(0,30),rcauchy(50))
fit <- fit.hgld(data = data,mixture = TRUE)
```

healthcare

Healthcare data set

Description

Data set containing the healthcare expense of 129,257 customers of a Brazilian healthcare company between 2006 and 2009.

Usage

```
healthcare
```

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 289456 rows and 8 columns.

Details

The expenses are in Reais (Brazilian currency) and were deflated to the January 2006 value.

Value

<code>ID</code>	The ID of the customer.
<code>sex</code>	The sex of the customer.
<code>age</code>	The age of the customer on the considered year.
<code>expense</code>	The healthcare expense of the customer on the considered year.
<code>log_expense</code>	The logarithm of the healthcare expense of the customer on the considered year.
<code>year</code>	The considered year.
<code>previous_expense</code>	The healthcare expense of the customer on the previous year.
<code>log_previous_expense</code>	The logarithm of the healthcare expense of the customer on the previous year.

References

Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) *arxiv1712.02183*

phgld

The Hurdle Generalized Lambda Distribution Family

Description

Probability function of the Hurdle Generalized Lambda Distribution.

Usage

```
phgld(q, mixture = FALSE, lambda1, lambda2 = NULL, prob = NULL,
      param = "fmkl", inverse.eps = 1e-08, max.iterations = 500)
```

Arguments

<code>q</code>	Vector of quantiles.
<code>mixture</code>	Whether to give the density of a mixture HGLD.
<code>lambda1</code>	A vector of length 5 with the five parameters of the HGLD or the first HGLD if <code>mixture = TRUE</code> .
<code>lambda2</code>	A vector of length 4 with the four parameters of the second HGLD.
<code>prob</code>	The cluster parameter for the mixture HGLD.
<code>param</code>	"fmkl" or "rs".
<code>inverse.eps</code>	Accuracy of calculation for the numerical determination of $F(x)$, defaults to $1e-8$.
<code>max.iterations</code>	Maximum number of iterations in the numerical determination of $F(x)$, defaults to 500.

Details

If the parametrization of the RS or FMKL GLD is not acceptable, the function returns NA. This function is based on the [GLDEX](#) package.

Value

The probability function of the HGLD.

References

Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a zero-inflated Generalized Lambda Distribution to healthcare claims. (2017) *arxiv1712.02183*

Su, S.; Fitting Single and Mixture of Generalized Lambda Distributions to Data via Discretized and Maximum Likelihood Methods: GLDEX in R. (2007), Journal of Statistical Software: *21* 9.

Examples

```
library(ggplot2)
set.seed(100)
data <- healthcare[sample(1:nrow(healthcare),100),]
fit <- fit.hgld(data$log_expense)
{ggplot(data.frame(x = data$log_expense),aes(x = x)) +
  stat_function(fun = function(x) phgld(q = x,lambda1 = fit$par[,2],param = "rs"))}

#mixture
set.seed(100)
data <- c(rcauchy(50,location = 10),rep(0,30),rcauchy(50))
fit <- fit.hgld(data,TRUE)
{ggplot(data.frame(x = data[data!=0]),aes(x = x)) +
  stat_function(fun = function(x) phgld(q = x,mixture = TRUE,lambda1 = fit$par[1:5,2],
    lambda2 = fit$par[6:9,2],prob = fit$par[10,2],param = "rs"))}
```


plot.fit.hgld

*Plot the Hurdle Generalized Lambda Distribution fit***Description**

Density plot of the fit given by an object of class [fit.hgld](#).

Usage

```
## S3 method for class 'fit.hgld'
plot(x, histogram = TRUE, hcolor = "white",
     hfill = "black", bins = 50, dcolor = c("black", "red"),
     dtype = c("solid", "solid"), xlab = "Data", ylab = "Density", ...)
```

Arguments

x	A fit.hgld object.
histogram	Logical. Whether the estimated density must be superimposed by the data histogram.
hcolor	Color of the histogram.
hfill	Color to fill the histogram.
bins	Number of histogram bins.
dcolor	Color of the RS and FMKL density plot.
dtype	Type of the RS and FMKL density line.
xlab	Label of the x-axis.
ylab	Label of the y-axis.
...	Arguments to be passed to methods.

Details

The density may be plotted by itself or superimposed by the data histogram. Plot only the nonzero data values.

Value

RS	ggplot plot of the fitted RS GLD.
FMKL	ggplot plot of the fitted FMKL GLD.
RF	ggplot plot of the fitted RS and FMKL GLD on the same plot.

References

Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) *arxiv1712.02183*

Examples

```
set.seed(100)
data <- healthcare[sample(1:nrow(healthcare),100),]
fit <- fit.hgld(data$log_expense)
plot(fit)

#mixture
set.seed(100)
data <- c(rcauchy(50,location = 10),rep(0,30),rcauchy(50))
fit <- fit.hgld(data = data,mixture = TRUE)
plot(fit)
```

plot.reg.hgld

Predict density plot of a Hurdle Generalized Lambda Regression

Description

Predict density plot of a Hurdle Generalized Lambda Regression.

Usage

```
## S3 method for class 'reg.hgld'
plot(x, newvalues, name = row.names(newvalues),
     color = NULL, xlab = "Data",
     xlim = c(min(x$NZdata[[all.vars(x$loc.formula)[1]]]),
               max(x$NZdata[[all.vars(x$loc.formula)[1]]])), title = "", ...)
```

Arguments

<code>x</code>	An object of class reg.hgld .
<code>newvalues</code>	A data frame with the new values of the covariates. Column names must match the ones given in formulas <i>loc.formula</i> and <i>zero.formula</i> .
<code>name</code>	A vector with the names of the new values rows. Default is the row names of the new values data frame.
<code>color</code>	The color of each density curve. Must have one color for each row of new values.
<code>xlab</code>	Label of the x-axis.
<code>xlim</code>	A vector with the limits of the x-axis.
<code>title</code>	Legend title.
<code>...</code>	Arguments to be passed to methods.

Details

Given an object of class [reg.hgld](#) and new values for the covariates, returns the density function of the fitted HGLD, given the covariates. The contrast on the *newvalues* data frame must be same of the *data* used in the [reg.hgld](#) object. All density curves are plotted in the same plot.

Value

plot [ggplot](#) density plot for the given new values

References

Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) *arxiv1712.02183*

Examples

```
set.seed(10)
tmp <- na.omit(healthcare)
data <- tmp[sample(1:nrow(tmp),100),]
formula <- log_expense ~ age + sex + log_previous_expense
reg <- suppressWarnings(reg.hgld(data,formula,formula,TRUE,n.simu = 10,param = "rs",plotKS = TRUE))
newvalues <- tmp[sample(1:nrow(tmp),5),c(2,3,5,8)]
plot(reg,newvalues)
```

qhglD

*The Hurdle Generalized Lambda Distribution Family***Description**

Quantile function of a Hurdle Generalized Lambda Distribution.

Usage

```
qhglD(p, mixture = FALSE, lambda1, lambda2 = NULL, prob = NULL,
      param = "fmkl", trace = FALSE, inverse.eps = 1e-08,
      max.iterations = 500)
```

Arguments

p	Vector of probabilities.
mixture	Whether to give the density of a mixture HGLD.
lambda1	A vector of length 5 with the five parameters of the HGLD or the first HGLD if mixture = TRUE.
lambda2	A vector of length 4 with the four parameters of the second HGLD.
prob	The cluster parameter for the mixture HGLD.
param	"fmkl" or "rs".
trace	Whether progress bar must be printed in order to trace the algorithm.
inverse.eps	Accuracy of calculation for the numerical determination of F(x), defaults to 1e-8.
max.iterations	Maximum number of iterations in the numerical determination of F(x), defaults to 500.

Details

The HGLD must be non-negative, otherwise the quantile is only approximate. If the parametrization of the RS or FMKL GLD is not acceptable, the function returns NA. This function is based on the [GLDEX](#) package.

Value

The quantile function of a HGLD.

References

Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) *arxiv1712.02183*

Su, S.; Fitting Single and Mixture of Generalized Lambda Distributions to Data via Discretized and Maximum Likelihood Methods: GLDEX in R. (2007), Journal of Statistical Software: *21* 9.

Examples

```
set.seed(100)
data <- healthcare[sample(1:nrow(healthcare),100),]
fit <- fit.hglid(data$log_expense)
qhglid(p = seq(0,1,0.05),lambda1 = fit$par$RS,param = "rs")

#mixture
qhglid(p = seq(0,1,0.1),mixture = TRUE,lambda1 = c(0.05,0,1,3,6),lambda2 = c(0.3,2,3,6),prob = 0.5,
      param = "fmkl")
```

qhglid.reg	<i>Predict quantiles of a Hurdle Generalized Lambda regression</i>
------------	--

Description

Quantile function of a Hurdle Generalized Lambda Regression.

Usage

```
qhglid.reg(p, reg, newvalues, l0 = TRUE, location = TRUE)
```

Arguments

- p Vector of probabilities.
- reg An object of class [reg.hglid](#).
- newvalues A data frame with the new values of the covariates. Column names must match the ones given in formulas *loc.formula* and *zero.formula*.
- l0 Whether to return the lambda0 of each profile.
- location Whether to return the location of each profile.

Details

Given an object of class `reg.hgld`, a percentile and new values for the covariates, it returns the quantile of the fitted HGLD, given the covariates. The fitted HGLD must be non-negative. The contrast on the *newvalues* data frame must be same of the *data* used in the `reg.hgld` object.

Value

The quantile of the new values, according to the fitted HGLD regression.

References

Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) *arxiv1712.02183*

Examples

```
set.seed(10)
tmp <- na.omit(healthcare)
data <- tmp[sample(1:nrow(tmp),100),]
formula <- log_expense ~ age + sex + log_previous_expense
reg <- suppressWarnings(reg.hgld(data,formula,formula,TRUE,n.simu = 10,param = "rs",plotKS = TRUE))
newvalues <- tmp[sample(1:nrow(tmp),5),c(2,3,8)]
suppressWarnings(qhgld.reg(p = seq(0.05,0.95,0.1),reg,newvalues))
```

reg.hgld

Fit a Hurdle Generalized Lambda Distribution Regression model

Description

Fit a Hurdle Generalized Lambda Distribution Regression model to a dataset.

Usage

```
reg.hgld(data, zero.formula, loc.formula, full = FALSE, param = "fmkl",
  maxit = 20000, init = NULL, alpha = 0.05, n.simu = 1000,
  plotKS = TRUE, h.bins = 50)
```

Arguments

<code>data</code>	A dataset containing the variables of the model.
<code>zero.formula</code>	A symbolic expression of the model to be fitted to the inflation parameter of the distribution.
<code>loc.formula</code>	A symbolic expression of the model to be fitted to the location of the distribution.
<code>full</code>	Whether a simulation method must be applied to derive a confidence interval for the location regression coefficients.
<code>param</code>	"fmkl" or "rs".

maxit	Maximum number of iterations for numerical optimization.
init	Choose a different set of initial values to start the optimization process. This can either be full set of parameters including GLD parameter estimates, or it can just be the coefficient estimates of the regression model.
alpha	Significant level of the Confidence Interval for the GLD regression.
n.simu	Number of times to repeat the simulation runs, defaults to 1000.
plotKS	Whether to plot the KS resample test result within each plot.
h.bins	Number of bins for the GLD Regression normalized quantiles residuals histogram.

Details

Given a dataset, estimate by the numerical maximum likelihood method the regression coefficients of the model and the five parameters of the error GLD. The regression coefficients that model the location of the distribution are estimated by the functions [GLD.lm](#) and [GLD.lm.full](#) of the package [GLDreg](#). The regression coefficients that model the hurdle parameter of the distribution are estimated by the function [gamlss](#).

Value

coefficients	The estimated coefficients of the HGLD regression.
Zplot	A function that generates the four diagnostic plots of the logistic regression.
Zres.sumarry	Summary of the quantile residuals of the logistic regression.
Zfit	Quantile residuals versus fitted values for the logistic regression.
Zindex	Quantile residuals versus index for the logistic regression.
Zdensity	Quantile residuals density for the logistic regression.
Zqq	QQ-norm plot of the quantile residuals for the logistic regression.
NZqq	QQ-plot for the GLD regression.
NZquant	Quantile plot for the GLD regression.
NZhistogram	The histogram of the residuals.
NZplot	The quantile residuals plots for the GLD regression.
NZres.sumarry	Summary of the quantile residuals of the GLD regression.
NZfit	Quantile residuals versus fitted values for the GLD regression.
NZindex	Quantile residuals versus index for the GLD regression.
NZdensity	Quantile residuals density for the GLD regression.
NZqqQuant	QQ-norm plot of the quantile residuals for the GLD regression.
KS	KS test p-value for the GLD regression.
gamlss	The gamlss object of the fitted logistic regression.
GLDreg	The GLDreg object of the fitted GLD regression.
Zdata	The data used to fit the logistic regression
NZdata	The data used to fit the GLD regression.

zero.formula	A symbolic expression of the model to be fitted to the inflation parameter of the distribution.
loc.formula	A symbolic expression of the model to be fitted to the location of the distribution.
param	"fmkl" or "rs".
full	Whether a simulation method must be applied to derive a confidence interval for the location regression coefficients.

References

Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) *arxiv1712.02183*

Su, S.; Flexible Parametric Quantile Regression Model. (2015), *Statistics & Computing* May 2015, Volume 25, Issue 3, pp 635-650

Examples

```
set.seed(100)
tmp <- na.omit(healthcare)
data <- tmp[sample(1:nrow(tmp),100),]
formula <- log_expense ~ age + sex + log_previous_expense
r <- suppressWarnings(reg.hgld(data,formula,formula,TRUE,n.simu = 10,param = "rs",plotKS = TRUE))
```

rhgld

The Hurdle Generalized Lambda Distribution Family

Description

Sample from a Hurdle Generalized Lambda Distribution.

Usage

```
rhgld(n, mixture = FALSE, lambda1, lambda2 = NULL, prob = NULL,
      param = "fmkl")
```

Arguments

n	Number of observations to be generated.
mixture	Whether to give the density of a mixture HGLD.
lambda1	A vector of length 5 with the five parameters of the HGLD or the first HGLD if mixture = TRUE.
lambda2	A vector of length 4 with the four parameters of the second HGLD.
prob	The cluster parameter for the mixture HGLD.
param	"fmkl" or "rs".

Details

If the parametrization of the RS or FMKL GLD is not acceptable, the function returns NA. This function is based on the [GLDEX](#) package.

Value

A sample of a HGLD.

References

Marcondes, D.; Peixoto, C.; Maia, A. C.; Fitting a Hurdle Generalized Lambda Distribution to healthcare expenses. (2017) *arxiv1712.02183*

Su, S.; Fitting Single and Mixture of Generalized Lambda Distributions to Data via Discretized and Maximum Likelihood Methods: GLDEX in R. (2007), Journal of Statistical Software: *21* 9.

Examples

```
library(ggplot2)
set.seed(100)
qplot(rhgld(n = 1000, lambda1 = c(0.05, 0, 1, 3, 6), param = "fmkl"), ..density.., geom = "histogram",
      bins = 30)
qplot(rhgld(n = 1000, mixture = TRUE, lambda1 = c(0.05, 0, 1, 3, 6), lambda2 = c(0.3, 2, 3, 6), prob = 0.5,
      param = "fmkl"), geom = "histogram", bins = 50)
```


Index

*Topic **datasets**

healthcare, [6](#)

dhgld, [2](#)

diag.hgld, [3](#)

fit.hgld, [3](#), [5](#), [9](#)

fun.auto.bimodal.pml, [5](#)

fun.data.fit.ml, [5](#), [6](#)

gamlss, [14](#)

ggplot, [4](#), [9](#), [11](#)

GLD.lm, [14](#)

GLD.lm.full, [14](#)

GLDEX, [2](#), [5](#), [8](#), [12](#), [16](#)

GLDreg, [14](#)

healthcare, [6](#)

phgld, [7](#)

plot.fit.hgld, [9](#)

plot.reg.hgld, [10](#)

qhgld, [11](#)

qhgld.reg, [12](#)

QUnif, [5](#)

reg.hgld, [10](#), [12](#), [13](#), [13](#)

rhgld, [15](#)

runif.halton, [5](#)

runif.sobol, [5](#)