Package 'metaquant'

February 11, 2025

Type Package

Title Estimating Means, Standard Deviations and Visualising
Distributions using Quantiles

Version 0.1.1

Maintainer Udara Kumaranathunga <U.Kumaranathunga@latrobe.edu.au>

Author Udara Kumaranathunga [cre],
Alysha De Livera [aut],
Luke Prendergast [aut]

Description Implements a novel density-
based approach for estimating unknown means, visualizing distributions,
and meta-analyses of quantiles. A detailed vignettes with example datasets and code to pre-
pare data and analyses is available at <https://bookdown.org/a2delivera/metaquant/>.
The methods are described in the pre-
print by De Livera, Prendergast and Kumaranathunga (2024, <doi:10.48550/arXiv.2411.10971>).

License GPL-3

Encoding UTF-8

Depends R (>= 3.5.0)

RoxygenNote 7.3.2

Imports gld, sld, stats, ggplot2, plotly, magrittr, dplyr, estmeansd

NeedsCompilation no

Repository CRAN

Date/Publication 2025-02-11 17:00:02 UTC

Contents

---

est.mean   *Estimating Sample Mean using Quantiles*

Description

This function estimates the sample mean from a study presenting quantile summary measures with the sample size (n). The quantile summaries can fall into one of the following categories:

- S_1: { minimum, median, maximum }
- S_2: { first quartile, median, third quartile }
- S_3: { minimum, first quartile, median, third quartile, maximum }

The est.mean function implements newly proposed flexible quantile-based distribution methods for estimating sample mean (De Livera et al., 2024). It also incorporates existing methods for estimating sample means as described by Luo et al. (2018) and McGrath et al. (2020).

Usage

```
est.mean(
    min = NULL,
    q1 = NULL,
    med = NULL,
    q3 = NULL,
    max = NULL,
    n = NULL,
    method = "gld/sld",
    opt = TRUE
    )
```

Arguments

min	numeric value representing the sample minimum.
q1	numeric value representing the first quartile of the sample.
med	numeric value representing the median of the sample.
q3	numeric value representing the third quartile of the sample.
max	numeric value representing the sample maximum.
n	numeric value specifying the sample size.
method	character string specifying the approach used to estimate the sample means. The options are the following:

> 'gld/sld' The default option. The method proposed by De Livera et al. (2024). Estimation using the Generalized Lambda Distribution (GLD) for 5-number summaries (S_3), and the Skew Logistic Distribution (SLD) for 3-number summaries (S_1 and S_2).
>
> 'luo' Method of Luo et al. (2018).

> 'hozo/wan/bland' The method proposed by Wan et al. (2014). i.e., the method of Hozo et al. (2005) for S_1, method of Wan et al. (2014) for S_2, and method of Bland (2015) for S_3.
>
> 'bc' Box-Cox method proposed by McGrath et al. (2020).
>
> 'qe' Quantile Matching Estimation method proposed by McGrath et al. (2020).

opt   logical value indicating whether to apply the optimization step of 'gld/sld' method, in estimating their parameters using theoretical quantiles. The default value is TRUE.

Details

The 'gld/sld' method (i.e., the method of De Livera et al., (2024)) of est.mean uses the following quantile based distributions:

- Generalized Lambda Distribution (GLD) for estimating the sample mean using 5-number summaries (S_3).

- Skew Logistic Distribution (SLD) for estimating the sample mean using 3-number summaries (S_1 and S_2).

The generalised lambda distribution (GLD) is a four parameter family of distributions defined by its quantile function under the FKML parameterisation (Freimer et al., 1988). De Livera et al. propose that the GLD quantlie function can be used to approximate a sample's distribution using 5-point summaries. The four parameters of GLD quantile function include: a location parameter (λ_1), an inverse scale parameter ($\lambda_2>0$), and two shape parameters (λ_3 and λ_4).

The quantile-based skew logistic distribution (SLD), introduced by Gilchrist (2000) and further modified by van Staden and King (2015) is used to approximate the sample's distribution using 3-point summaries. The SLD quantile function is defined using three parameters: a location parameter (λ), a scale parameter (η), and a skewing parameter (δ).

For 'gld/sld' method, the parameters of the generalized lambda distribution (GLD) and skew logistic distribution (SLD) are estimated by formulating and solving a set of simultaneous equations. These equations relate the estimated sample quantiles to their theoretical counterparts of the respective distribution (GLD or SLD). Finally, the mean for each scenario is calculated by integrating functions of the estimated quantile function.

Value

mean: numeric value representing the estimated mean of the sample.

References

Alysha De Livera, Luke Prendergast, and Udara Kumaranathunga. A novel density-based approach for estimating unknown means, distribution visualisations, and meta-analyses of quantiles. *Submitted for Review*, 2024, pre-print available here: https://arxiv.org/abs/2411.10971

Dehui Luo, Xiang Wan, Jiming Liu, and Tiejun Tong. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Statistical methods in medical research*, 27(6):1785–1805,2018.

Xiang Wan, Wenqian Wang, Jiming Liu, and Tiejun Tong. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC medical research methodology*, 14:1–13, 2014.

Sean McGrath, XiaoFei Zhao, Russell Steele, Brett D Thombs, Andrea Benedetti, and DEPRESsion Screening Data (DEPRESSD) Collaboration. Estimating the sample mean and standard deviation from commonly reported quantiles in meta-analysis. *Statistical methods in medical research*, 29(9):2520–2537, 2020b.

Marshall Freimer, Georgia Kollia, Govind S Mudholkar, and C Thomas Lin. A study of the generalized tukey lambda family. *Communications in Statistics-Theory and Methods*, 17(10):3547–3567, 1988.

Warren Gilchrist. *Statistical modelling with quantile functions*. Chapman and Hall/CRC, 2000.

P. J. van Staden and R. A. R. King. The quantile-based skew logistic distribution. *Statistics & Probability Letters*, 96:109–116, 2015.

Examples

```
#Generate 5-point summary data
set.seed(123)
n <- 1000
x <- stats::rlnorm(n, 4, 0.3)
quants <- c(min(x), stats::quantile(x, probs = c(0.25, 0.5, 0.75)), max(x))
obs_mean <- mean(x)

#Estimate sample mean using s3 (5 number summary)
est_mean_s3 <- est.mean(min = quants[1], q1 = quants[2], med = quants[3], q3 = quants[4],
                        max = quants[5], n=n, method = "gld/sld")
est_mean_s3

#Estimate sample mean using s1 (min, median, max)
est_mean_s1 <- est.mean(min = quants[1], med = quants[3], max = quants[5],
                        n=n, method = "gld/sld")
est_mean_s1

#Estimate sample mean using s2 (q1, median, q3)
est_mean_s2 <- est.mean(q1 = quants[2], med = quants[3], q3 = quants[4],
                        n=n, method = "gld/sld")
est_mean_s2
```

---

est.mean.2g   *Estimating Sample Means of Two Groups using Quantiles*

Description

This function estimates the sample means from a two group study presenting quantile summary measures with the sample size (n). The quantile summaries of each group can fall into one of the following categories:

- S_1: { minimum, median, maximum }
- S_2: { first quartile, median, third quartile }
- S_3: { minimum, first quartile, median, third quartile, maximum }

The `est.mean.2g` function uses a novel quantile-based distribution methods for estimating sample mean for two groups such as 'Treatment' and 'Control' (De Livera et al., 2024). The method is based on the following quantile-based distributions for estimating sample means:

- Generalized Lambda Distribution (GLD) for estimating sample means using 5-number summaries (S_3).
- Skew Logistic Distribution (SLD) for estimating sample means using 3-number summaries (S_1 and S_2).

Usage

```
est.mean.2g(
    min.g1 = NULL,
    q1.g1 = NULL,
    med.g1 = NULL,
    q3.g1 = NULL,
    max.g1 = NULL,
    min.g2 = NULL,
    q1.g2 = NULL,
    med.g2 = NULL,
    q3.g2 = NULL,
    max.g2 = NULL,
    n.g1,
    n.g2,
    opt = TRUE
)
```

Arguments

`min.g1`	numeric value representing the sample minimum of group 1.
`q1.g1`	numeric value representing the first quartile of group 1.
`med.g1`	numeric value representing the median of group 1.
`q3.g1`	numeric value representing the third quartile of group 1.
`max.g1`	numeric value representing the sample maximum of group 1.
`min.g2`	numeric value representing the sample minimum of group 2.
`q1.g2`	numeric value representing the first quartile of group 2.
`med.g2`	numeric value representing the median of group 2.
`q3.g2`	numeric value representing the third quartile of group 2.
`max.g2`	numeric value representing the sample maximum of group 2.
`n.g1`	numeric value specifying the sample size of group 1.
`n.g2`	numeric value specifying the sample size of group 2.
`opt`	logical value indicating whether to apply the optimization step in estimating the parameters of GLD or SLD. Default is `TRUE`.

Details

The `est.mean.2g` function implement the methods proposed by De Livera et al. (2024) for the two group case by incorporating shared information across the two groups to improve the accuracy of the estimates.

The generalised lambda distribution (GLD) is a four parameter family of distributions defined by its quantile function under the FKML parameterisation (Freimer et al., 1988). De Livera et al. propose that the GLD quantlie function can be used to approximate a sample's distribution using 5-point summaries (S_3). The four parameters of GLD quantile function include: a location parameter (λ_1), an inverse scale parameter ($\lambda_2 > 0$), and two shape parameters (λ_3 and λ_4). The `est.mean.sld.2g` function considers the case where the underlying distribution in each group has the same shape (i.e., common λ_3 and λ_4), and differ only in location and scale. Weights are used in the optimisation step in estimating λ_3 and λ_4 to put more emphasis on the group with the larger sample size.

The quantile-based skew logistic distribution (SLD), introduced by Gilchrist (2000) and further modified by van Staden and King (2015) is used to approximate the sample's distribution using 3-point summaries (S_1 and S_2). The SLD quantile function is defined using three parameters: a location parameter (λ), a scale parameter (η), and a skewing parameter (δ). In `est.mean.2g`, an assumption of a common skewing parameter (δ) is used for the two groups, so a pooled estimate of δ is computed using weights based on the sample sizes.

Under each scenario, the parameters of the respective distributions are estimated by formulating and solving a series of simultaneous equations which relate the estimated quantiles with the population counterparts. The estimated mean is then obtained via integration of functions of the estimated quantile function.

Value

A list containing the estimated sample means for the two groups:

- `mean.g1`: numeric value representing the estimated mean of group 1.
- `mean.g2`: numeric value representing the estimated mean of group 2.

References

Alysha De Livera, Luke Prendergast, and Udara Kumaranathunga. A novel density-based approach for estimating unknown means, distribution visualisations, and meta-analyses of quantiles. *Submitted for Review*, 2024, pre-print available here: https://arxiv.org/abs/2411.10971

Marshall Freimer, Georgia Kollia, Govind S Mudholkar, and C Thomas Lin. A study of the generalized tukey lambda family. *Communications in Statistics-Theory and Methods*, 17(10):3547–3567, 1988.

Warren Gilchrist. *Statistical modelling with quantile functions*. Chapman and Hall/CRC, 2000.

P. J. van Staden and R. A. R. King. The quantile-based skew logistic distribution. *Statistics & Probability Letters*, 96:109–116, 2015.

See Also

`est.mean` for estimating means from one-group quantile data.

Examples

```
#Generate 5-point summary data for two groups
set.seed(123)
n_t <- 1000
n_c <- 1500
x_t <- stats::rlnorm(n_t, 4, 0.3)
x_c <- 1.1*(stats::rlnorm(n_c, 4, 0.3))
q_t <- c(min(x_t), stats::quantile(x_t, probs = c(0.25, 0.5, 0.75)), max(x_t))
q_c <- c(min(x_c), stats::quantile(x_c, probs = c(0.25, 0.5, 0.75)), max(x_c))
obs_mean_t <- mean(x_t)
obs_mean_c <- mean(x_c)

#Estimate sample mean using s3 (5 number summary)
est_means_s3 <- est.mean.2g(q_t[1],q_t[2],q_t[3],q_t[4],q_t[5],
                            q_c[1],q_c[2],q_c[3],q_c[4],q_c[5],
                            n.g1 = n_t,
                            n.g2 = n_c)
est_means_s3

#Estimate sample mean using s1 (min, med, max)
est_means_s1 <- est.mean.2g(min.g1=q_t[1], med.g1=q_t[3], max.g1=q_t[5],
                            min.g2=q_c[1], med.g2=q_c[3], max.g2=q_c[5],
                            n.g1 = n_t,
                            n.g2 = n_c)
est_means_s1

#Estimate sample mean using s2 (q1, med, q3)
est_means_s2 <- est.mean.2g(q1.g1=q_t[2], med.g1=q_t[3], q3.g1=q_t[4],
                            q1.g2=q_c[2], med.g2=q_c[3], q3.g2=q_c[4],
                            n.g1 = n_t,
                            n.g2 = n_c)
est_means_s2
```

---

est.sd   *Estimating Sample Standard Deviation using Quantiles*

Description

This function estimates the sample standard deviation from a study presenting quantile summary measures with the sample size (n). The quantile summaries can fall into one of the following categories:

- S_1: { minimum, median, maximum }
- S_2: { first quartile, median, third quartile }
- S_3: { minimum, first quartile, median, third quartile, maximum }

The est.sd function implements newly proposed flexible quantile-based distribution methods for estimating sample standard deviation by De Livera et al. (2024) as well as other existing methods for estimating sample standard deviations by Shi et al. (2020) and McGrath et al. (2020).

Usage

```
est.sd(
    min = NULL,
    q1 = NULL,
    med = NULL,
    q3 = NULL,
    max = NULL,
    n = NULL,
    method = "shi/wan",
    opt = TRUE
    )
```

Arguments

min	numeric value representing the sample minimum.
q1	numeric value representing the first quartile of the sample.
med	numeric value representing the median of the sample.
q3	numeric value representing the third quartile of the sample.
max	numeric value representing the sample maximum.
n	numeric value specifying the sample size.
method	character string specifying the approach used to estimate the sample standard deviations. The options are the following:

> 'shi/wan' The default option. Method of Shi et al. (2020).
>
> 'gld/sld' The method proposed by De Livera et al. (2024). Estimation using the Generalized Lambda Distribution (GLD) for 5-number summaries (S_3), and the Skew Logistic Distribution (SLD) for 3-number summaries (S_1 and S_2).
>
> 'wan' The method proposed by Wan et al. (2014).
>
> 'bc' Box-Cox method proposed by McGrath et al. (2020).
>
> 'qe' Quantile Matching Estimation method proposed by McGrath et al. (2020).

opt	logical value indicating whether to apply the optimization step of 'gld/sld' method, in estimating their parameters using theoretical quantiles. The default value is TRUE.

Details

For details explaining the new method 'gld/sld', check est.mean.

Value

sd: numeric value representing the estimated standard deviation of the sample.

References

Alysha De Livera, Luke Prendergast, and Udara Kumaranathunga. A novel density-based approach for estimating unknown means, distribution visualisations, and meta-analyses of quantiles. *Submitted for Review*, 2024, pre-print available here: https://arxiv.org/abs/2411.10971

Jiandong Shi, Dehui Luo, Hong Weng, Xian-Tao Zeng, Lu Lin, Haitao Chu, and Tiejun Tong. Optimally estimating the sample standard deviation from the five-number summary. *Research synthesis methods*, 11(5):641–654, 2020.

Xiang Wan, Wenqian Wang, Jiming Liu, and Tiejun Tong. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC medical research methodology*, 14:1–13, 2014.

Sean McGrath, XiaoFei Zhao, Russell Steele, Brett D Thombs, Andrea Benedetti, and DEPRESSion Screening Data (DEPRESSD) Collaboration. Estimating the sample mean and standard deviation from commonly reported quantiles in meta-analysis. *Statistical methods in medical research*, 29(9):2520–2537, 2020b.

Examples

```
#Generate 5-point summary data
set.seed(123)
n <- 1000
x <- stats::rlnorm(n, 5, 0.5)
quants <- c(min(x), stats::quantile(x, probs = c(0.25, 0.5, 0.75)), max(x))
obs_sd <- sd(x)

#Estimate sample SD using s3 (5 number summary)
est_sd_s3 <- est.sd(min = quants[1], q1 = quants[2], med = quants[3], q3 = quants[4],
                    max = quants[5], n=n, method = "gld/sld")
est_sd_s3

#Estimate sample SD using s1 (min, median, max)
est_sd_s1 <- est.sd(min = quants[1], med = quants[3], max = quants[5],
                    n=n, method = "gld/sld")
est_sd_s1

#Estimate sample SD using s2 (q1, median, q3)
est_sd_s2 <- est.sd(q1 = quants[2], med = quants[3], q3 = quants[4],
                    n=n, method = "gld/sld")
est_sd_s2
```

---

est.sd.2g   *Estimating Sample Standard Deviations of Two Groups using Quantiles*

Description

This function estimates the sample standard deviations (SD) from a two group study presenting quantile summary measures with the sample size (n). The quantile summaries of each group can fall into one of the following categories:

- S_1: { minimum, median, maximum }
- S_2: { first quartile, median, third quartile }
- S_3: { minimum, first quartile, median, third quartile, maximum }

The est.sd.2g function uses a novel quantile-based distribution methods for estimating sample SD for two groups such as 'Treatment' and 'Control' (De Livera et al., 2024). The method is based on the following quantile-based distributions:

- Generalized Lambda Distribution (GLD) for estimating sample SDs using 5-number summaries (S_3).
- Skew Logistic Distribution (SLD) for estimating sample SDs using 3-number summaries (S_1 and S_2).

Usage

```
est.sd.2g(
    min.g1 = NULL,
    q1.g1 = NULL,
    med.g1 = NULL,
    q3.g1 = NULL,
    max.g1 = NULL,
    min.g2 = NULL,
    q1.g2 = NULL,
    med.g2 = NULL,
    q3.g2 = NULL,
    max.g2 = NULL,
    n.g1,
    n.g2,
    opt = TRUE
)
```

Arguments

min.g1	numeric value representing the sample minimum of group 1.
q1.g1	numeric value representing the first quartile of group 1.
med.g1	numeric value representing the median of group 1.
q3.g1	numeric value representing the third quartile of group 1.
max.g1	numeric value representing the sample maximum of group 1.
min.g2	numeric value representing the sample minimum of group 2.
q1.g2	numeric value representing the first quartile of group 2.
med.g2	numeric value representing the median of group 2.

q3.g2	numeric value representing the third quartile of group 2.
max.g2	numeric value representing the sample maximum of group 2.
n.g1	numeric value specifying the sample size of group 1.
n.g2	numeric value specifying the sample size of group 2.
opt	logical value indicating whether to apply the optimization step in estimating the parameters of GLD or SLD. Default is TRUE.

Details

For details explaining the method of estimating using GLD or SLD, check `est.mean.2g`.

Value

A list containing the estimated sample SDs for the two groups:

- `sd.g1`: numeric value representing the estimated standard deviation of group 1.
- `sd.g2`: numeric value representing the estimated standard deviation of group 2.

References

Alysha De Livera, Luke Prendergast, and Udara Kumaranathunga. A novel density-based approach for estimating unknown means, distribution visualisations, and meta-analyses of quantiles. *Submitted for Review*, 2024, pre-print available here: https://arxiv.org/abs/2411.10971

See Also

`est.sd` for estimating standard deviation from one-group quantile data.

Examples

```
#Generate 5-point summary data for two groups
set.seed(123)
n_t <- 1000
n_c <- 1500
x_t <- stats::rlnorm(n_t, 5, 0.5)
x_c <- 1.1*(stats::rlnorm(n_c, 5, 0.5))
q_t <- c(min(x_t), stats::quantile(x_t, probs = c(0.25, 0.5, 0.75)), max(x_t))
q_c <- c(min(x_c), stats::quantile(x_c, probs = c(0.25, 0.5, 0.75)), max(x_c))
obs_sd_t <- sd(x_t)
obs_sd_c <- sd(x_c)

#Estimate sample SD using s3 (5 number summary)
est_sds_s3 <- est.sd.2g(q_t[1],q_t[2],q_t[3],q_t[4],q_t[5],
                        q_c[1],q_c[2],q_c[3],q_c[4],q_c[5],
                        n.g1 = n_t,
                        n.g2 = n_c)
est_sds_s3

#Estimate sample SD using s1 (min, med, max)
est_sds_s1 <- est.sd.2g(min.g1=q_t[1], med.g1=q_t[3], max.g1=q_t[5],
```

```
                              min.g2=q_c[1], med.g2=q_c[3], max.g2=q_c[5],
                              n.g1 = n_t,
                              n.g2 = n_c)
      est_sds_s1

      #Estimate sample SD using s2 (q1, med, q3)
      est_sds_s2 <- est.sd.2g(q1.g1=q_t[2], med.g1=q_t[3], q3.g1=q_t[4],
                              q1.g2=q_c[2], med.g2=q_c[3], q3.g2=q_c[4],
                              n.g1 = n_t,
                              n.g2 = n_c)
      est_sds_s2
```

---

plotdist	*Visualising Densities using Quantiles*

Description

The function estimates and visualizes the density curves of one-group or two-group studies presenting quantile summary measures with the sample size (n). The quantile summaries can fall into one of the following categories:

- S_1: { minimum, median, maximum }
- S_2: { first quartile, median, third quartile }
- S_3: { minimum, first quartile, median, third quartile, maximum }

The `plotdist` function uses the following quantile-based distribution methods for visualising densities using qantiles (De Livera et al., 2024).

- Generalized Lambda Distribution (GLD) when 5-number summaries present (S_3).
- Skew Logistic Distribution (SLD) when 3-number summaries present (S_1 and S_2).

Usage

```
plotdist(
    data,
    xmin = NULL,
    xmax = NULL,
    ymax = NULL,
    length.out = 1000,
    title = "",
    xlab = "x",
    ylab = "Density",
    line.size = 0.5,
    title.size = 12,
    lab.size = 10,
    color.g1 = "pink",
    color.g2 = "skyblue",
```

```
    color.g1.pooled = "red",
    color.g2.pooled = "blue",
    label.g1 = NULL,
    label.g2 = NULL,
    display.index = FALSE,
    display.legend = FALSE,
    pooled.dist = FALSE,
    pooled.only = FALSE,
    opt = TRUE
)
```

Arguments

data
: data frame containing the quantile summary data. For one-group studies, the input dataset may contain the following columns depending on the quantile scenario:

  'stduy.index' stduy index or name

  'min.g1' minimum value

  'q1.g1' first quartile

  'med.g1' median

  'q3.g1' third quartile

  'max.g1' maximum value

  'n.g1' sample size

  For two-group studies, the data frame may also contain the following columns for the second group: min.g2, q1.g2, med.g2, q3.g2, max.g2 and n.g2. Note that, for three-point summaries (S_1 and S_2), only the relevant columns should be included.

xmin
: numeric value for the lower limit of the x-axis for density calculation. It is recommended to set this to a value smaller than the smallest value across the quantile summaries to ensure the density curve is fully captured. If xmin is not provided, the minimum value of the 'min.' columns will be used for scenario S_1 or S_3. Note that for scenario S_2, no default calculation is performed for xmin.

xmax
: numeric value for the upper limit of the x-axis for density calculation. It is recommended to set this to a value larger than the largest value across the quantile summaries to ensure the density curve is fully captured. If xmax is not provided, the maximum value of the 'max.' columns will be used for scenario S_1 or S_3. Similarly, for scenario S_2, no default calculation is performed for xmax.

ymax
: numeric value for the upper limit of the y-axis. If NULL, the highest density value will be used.

length.out
: integer specifying the number of points along the x-axis for density calculation. Default is 1000.

title
: character string for the plot title. Default is an empty string.

xlab
: character string for the x-axis label. Default is "x".

ylab
: character string for the y-axis label. Default is "Density".

line.size
: numeric. Thickness of the density curve lines. Default is 0.5.

title.size	numeric. Font size for the plot title. Default is 12.
lab.size	numeric. Font size for axis labels. Default is 10.
color.g1	character string specifying the color for individual density curves of group 1 for each study (row). Default is "pink".
color.g2	character string specifying the color for individual density curves of group 2 for each study (row). Default is "skyblue".
color.g1.pooled	
	character string specifying the color for pooled density curve of group 1. Default is "red".
color.g2.pooled	
	character string specifying the color for pooled density curve of group 2. Default is "blue".
label.g1	character string indicating label or name for group 1 (eg., 'Treatment')
label.g2	character string indicating label or name for group 2 (eg., 'Control').
	If 'label.g1' and 'label.g2' are not provided, the function will assign labels as 'Group 1' and 'Group 2'.
display.index	logical. If TRUE, the 'study.index' of each quantile set (row) will be displayed alongside the corresponding density curve. The default is FALSE, meaning no labels will be shown. The label text size is controlled by the lab.size parameter.
display.legend	logical. If TRUE, legends ('label.g1' and/or 'label.g2') will be displayed on the right side of the plot. The default is FALSE. The legend text size is controlled by the lab.size parameter.
pooled.dist	logical. If TRUE, pooled density curves for group 1 and/or group 2 will be plotted along with the individual density curves. The default is FALSE.
pooled.only	logical. If TRUE, only the pooled density curves of group 1 and/or group 2 will be plotted, excluding the individual density curves. The default is FALSE.
opt	logical value indicating whether to apply the optimization step when estimating GLD or SLD parameters. The default value is TRUE.

Details

The generalised lambda distribution (GLD) is a four parameter family of distributions defined by its quantile function under the FKML parameterisation (Freimer et al., 1988). De Livera et al. propose that the GLD quantile function can be used to approximate a sample's distribution using 5-point summaries. The four parameters of GLD quantile function include: a location parameter (λ_1), an inverse scale parameter ($\lambda_2 > 0$), and two shape parameters (λ_3 and λ_4).

The quantile-based skew logistic distribution (SLD), introduced by Gilchrist (2000) and further modified by van Staden and King (2015) is used to approximate the sample's distribution using 3-point summaries. The SLD quantile function is defined using three parameters: a location parameter (λ), a scale parameter (η), and a skewing parameter (δ).

These parameters of GLD and SLD are estimated by formulating and solving a series of simultaneous equations which relate the estimated quantiles with the population counterparts of respective distribution (GLD or SLD). The plotdist uses these estimated parameters, to compute the density data using [dgl](#) function from the **gld** package and [dsl](#) function from the **sld** package.

If one needs to generate pooled density plots, they can use the `pooled.dist` or `pooled.only` arguments as described in the *Arguments* section. The pooled density curves represent a weighted average of individual study densities, with weights determined by sample sizes. The method is similar to obtaining pooled estimates of effects in a standard meta-analysis and it serves as a way to visualize combined estimated distributional information across studies.

Value

An interactive plotly object visualizing the estimated density curve(s) for one or two groups.

References

Alysha De Livera, Luke Prendergast, and Udara Kumaranathunga. A novel density-based approach for estimating unknown means, distribution visualisations, and meta-analyses of quantiles. *Submitted for Review*, 2024, pre-print available here: https://arxiv.org/abs/2411.10971

Marshall Freimer, Georgia Kollia, Govind S Mudholkar, and C Thomas Lin. A study of the generalized tukey lambda family. *Communications in Statistics-Theory and Methods*, 17(10):3547–3567, 1988.

Warren Gilchrist. *Statistical modelling with quantile functions*. Chapman and Hall/CRC, 2000.

P. J. van Staden and R. A. R. King. The quantile-based skew logistic distribution. *Statistics & Probability Letters*, 96:109–116, 2015.

Examples

```
#Example dataset of 3-point summaries (min, med, max) for 2 groups
data_3num_2g <- data.frame(
  study.index = c("Study 1", "Study 2", "Study 3"),
  min.g1 = c(15, 15, 13),
  med.g1 = c(66, 68, 63),
  max.g1 = c(108, 101, 100),
  n.g1 = c(226, 230, 200),
  min.g2 = c(18, 19, 15),
  med.g2 = c(73, 82, 81),
  max.g2 = c(110, 115, 100),
  n.g2 = c(226, 230, 200)
 )
print(data_3num_2g)

#Density plots of two groups along with the pooled plots
plot_2g <- plotdist(
  data_3num_2g,
  xmin = 10,
  xmax = 125,
  title = "Example Density Plots of Two Groups",
  xlab = "x data",
  color.g1 = "skyblue",
  color.g2 = "pink",
  color.g1.pooled = "blue",
  color.g2.pooled = "red",
  label.g1 = "Treatment",
```

```
    label.g2 = "Control",
    display.legend = TRUE,
    pooled.dist = TRUE
)
print(plot_2g)
```

Index