Package: ROCsurf (via r-universe)

August 22, 2024

Type Package

Title ROC Surface Analysis Under the Three-Class Problems

Version 0.1.1

Maintainer Ertan Akgenç <statistician.ertan@gmail.com>

Description Receiver Operating Characteristic (ROC) analysis is performed assuming samples are from the proposed distributions. In addition, the volume under the ROC surface and true positive fractions values are evaluated by ROC surface analysis.

License GPL-3

URL https://github.com/ErtanSU/ROCsurf, https://ertansu.github.io/ROCsurf/

BugReports https://github.com/ErtanSU/ROCsurf/issues

Imports plotly, pracma, stats

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

Config/testthat/edition 3

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Language en-US

Repository https://ertansu.r-universe.dev

RemoteUrl https://github.com/ertansu/rocsurf

RemoteRef HEAD

RemoteSha 79581850beb64b06131a5d2bd74d917008993a01

Contents

ROCsurf *ROC Surface Analysis Under the Three-class Problems*

Description

ROC surface analysis is performed under the three-class classification problems. The volume under the ROC surface and true positive fractions values are evaulated by ROC surface analysis.

Usage

```
dG(x, alpha, beta)

dW(x, alpha, beta)

dL(z, alpha, beta)

pG(x, alpha, beta)

pW(y, alpha, beta)

pL(y, alpha, beta)

qG(p, alpha, beta)

qW(p, alpha, beta)

qL(p, alpha, beta)

rG(n, alpha, beta)

rW(n, alpha, beta)

rL(n, alpha, beta)

r.tc_vus(
  x,
  y,
  z,
  init_param = c(alpha1 = 1, beta1 = 1, alpha2 = 1, beta2 = 1, alpha3 = 1, beta3 = 1),
  true_param = c(alpha1 = 1, beta1 = 1, alpha2 = 1, beta2 = 1, alpha3 = 1, beta3 = 1),
  model = c("GWL", "GGW", "WGW", "WWW", "GGG", "LLL"),
  method = c("MLE", "AD", "CvM", "LSE", "WLSE", "TRUE")
)

r.tc_index(
  x,
  y,
```

```
  z,
 init_param = c(alpha1 = 1, beta1 = 1, alpha2 = 1, beta2 = 1, alpha3 = 1, beta3 = 1),
 true_param = c(alpha1 = 1, beta1 = 1, alpha2 = 1, beta2 = 1, alpha3 = 1, beta3 = 1),
  init_index = c(x, y),
  model = c("GWL", "GGW", "WGW", "WWW", "GGG", "LLL"),
  method = c("MLE", "AD", "CvM", "LSE", "WLSE", "TRUE")
)

r.tc_graph(
  x,
  y,
  z,
 init_param = c(alpha1 = 1, beta1 = 1, alpha2 = 1, beta2 = 1, alpha3 = 1, beta3 = 1),
 true_param = c(alpha1 = 1, beta1 = 1, alpha2 = 1, beta2 = 1, alpha3 = 1, beta3 = 1),
  empirical = TRUE,
  model = c("GWL", "GGW", "WGW", "WWW", "GGG", "LLL"),
  method = c("MLE", "AD", "CvM", "LSE", "WLSE", "TRUE")
)
```

Arguments

x, y, z	vector of quantiles.
alpha	shape parameter.
beta	scale parameter.
p	vector of probabilities.
n	number of observations. If length(n) > 1, the length is taken to be the number required.
init_param	initial paremeter values for the estimation method.
true_param	true parameter values.
model	selected model. The default value for the model is "GWL".
method	estimation method. The default value for the method is "MLE".
init_index	initial index value for the optimization calculation.
empirical	empirical must be TRUE or FALSE.
alpha1	shape parameter of distribution of first sample.
beta1	scale parameter of distribution of first sample.
alpha2	shape parameter of distribution of second sample.
beta2	scale parameter of distribution of second sample.
alpha3	location parameter of distribution of third sample.
beta3	scale parameter of distribution of third sample.

Details

The Gamma, Weibull, and Logistic distributions are widely used in statistical modeling and analysis. Below are the descriptions of their probability density functions (PDF), cumulative distribution functions (CDF), and quantile functions, including their parameter constraints.

Gamma Distribution

The Gamma distribution is a continuous probability distribution characterized by its shape (α) and scale (β) parameters. It is commonly used to model waiting times or lifetimes of objects.

- **Probability Density Function (PDF)**:

$$f(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), x > 0$$

  ,where $\alpha > 0$ and $\beta > 0$.

- **Cumulative Distribution Function (CDF)**:

$$F(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)} \gamma\left(\alpha, \frac{x}{\beta}\right)$$

  ,where $\gamma(\alpha, x/\beta)$ is the lower incomplete Gamma function.

- **Quantile Function**:
  The quantile function is the inverse of the CDF, denoted as $Q(p;\alpha,\beta) = F^{-1}(p;\alpha,\beta)$,where $0 < p < 1$.

Weibull Distribution

The Weibull distribution is used in reliability analysis and failure time analysis, characterized by its shape (α) and scale (β) parameters.

- **Probability Density Function (PDF)**:

$$f(x;\alpha,\beta) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right), x \geq 0$$

  ,where $\alpha > 0$, and $\beta > 0$.

- **Cumulative Distribution Function (CDF)**:

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

  ,where $\alpha > 0$, and $\beta > 0$.

- **Quantile Function**:

$$Q(p;\alpha,\beta) = \beta\left(-\log(1-p)\right)^{1/\alpha}$$

  ,where $0 < p < 1$, $\alpha > 0$, and $\beta > 0$.

Logistic Distribution

The Logistic distribution is used for growth models and logistic regression, characterized by its location (α) and scale (β) parameters.

- **Probability Density Function (PDF)**:

$$f(x;\alpha,\beta) = \frac{\exp\left(-\frac{x-\alpha}{\beta}\right)}{\beta\left(1+\exp\left(-\frac{x-\alpha}{\beta}\right)\right)^2}, -\infty < x < \infty$$

  ,where $\alpha \in \mathbb{R}$, and $\beta > 0$.

- **Cumulative Distribution Function (CDF)**:

$$F(x;\alpha,\beta) = \frac{1}{1+\exp\left(-\frac{x-\alpha}{\beta}\right)}$$

  ,where $\alpha \in \mathbb{R}$, and $\beta > 0$.

- **Quantile Function**:

$$Q(p;\alpha,\beta) = \alpha + \beta \log\left(\frac{p}{1-p}\right)$$

  ,where $0 < p < 1$, $\alpha \in \mathbb{R}$, and $\beta > 0$.

Additionally, the estimation methods Anderson-Darling "AD", Cramér-von Mises "CvM", least squares "LS" and weighted least squares "WLS" as well as the "TRUE" option for the true value, are available. Please note that the default value for the method parameter is maximum likelihood "ML" estimation. Also, models such as "GWL," "WWW," and "WGW" are defined for evaluating ROC surface analysis under three-class classification problems.

The cut-off point values corresponding to the generalized Youden's J index (J), The Perfection method (PM), The Maximum Volume (MV), and the newly proposed indices (NI, M) are provided.

Value

`dG` gives the probability density function of Gamma Distribution.

`dW` gives the probability density function of Weibull Distribution.

`dL` gives the probability density function of Logistic Distribution.

`pG` gives the cumulative density function of Gamma Distribution.

`pW` gives the cumulative density function of Weibull Distribution.

`pL` gives the cumulative density function of Logistic Distribution.

`qG` gives the quantile function of Gamma Distribution.

`qW` gives the quantile function of Weibull Distribution.

`qL` gives the quantile function of Logistic Distribution.

`rG` gives random numbers from Gamma Distribution.

`rW` gives random numbers from Weibull Distribution.

`rL` gives random numbers from Logistic Distribution.

`r.tc_vus` gives the Volume Under the Surface (VUS) when the data conforms to the proposed three distributions.

`r.tc_index` gives index values when the data conforms to the proposed three distributions.

`r.tc_graph` gives the ROC curve when the data conforms to the proposed three distributions.

References

Akgenç, E., and Kuş, C., 2023, *Statistical Inference for ROC Surface Analysis Under the Three-Class Problems*, 7th International Congress of Researchers, Statisticians and Young Statisticians (IRSYSC-2023).

B. R. Mosier and L. E. Bantis., 2021, *Estimation and construction of confidence intervals for biomarker cutoff-points under the shortest euclidean distance from the roc surface to the perfection corner*, Statistics in medicine, 40(20):4522–4539. doi:10.1002/sim.9077

G. Jameson., 2016, *The incomplete gamma functions.*, The Mathematical Gazette, 100(548):298–306. doi:10.1017/mag.2016.67

T. Dong., 2014, *Selected Topics on Statistical Methods for Three and Multiple Class Diagnostic Studies*, State University of New York at Buffalo.

J. Luo and C. Xiong., 2013, *Youden index and associated cut-points for three ordinal diagnostic groups*, Communications in Statistics-Simulation and Computation, 42(6):1213–1234. doi:10.1080/03610918.2012.661906

F. Edition, A. Papoulis, and S. U. Pillai., 2002, *Pobability, random variables, and stochastic processes*, McGraw-Hill Europe: New York, NY, USA.

A. J. Hallinan Jr., 1993, *A review of the weibull distribution*, Journal of Quality Technology, 25(2):85–93. doi:10.1080/00224065.1993.11979431

N. Balakrishnan., 1991, *Handbook of the logistic distribution*, CRC Press.

Examples

```
dG(c(1,2,3,4,5,200,1000),alpha=6,beta=.8)
dW(c(1,2,3,4,5,200,10000),alpha=1,beta=2)
dL(c(1,2,3,4,5,200),alpha=1,beta=.1)
pG(c(.5,1,2,3,4,25),alpha=6,beta=.8)
pW(c(.5,1,2,3,4,100),alpha=1,beta=2)
pL(c(.5,1,2),alpha=1,beta=.1)
qG(c(.9971,0.5,0.3),alpha=6,beta=.8)
qW(c(.9971,0.5,0.3),alpha=1,beta=2)
qL(c(.9971,0.5,0.3),alpha=1,beta=.1)
rG(10,alpha=6,beta=.8)
rW(10,alpha=1,beta=2)
rL(10,alpha=1,beta=.1)
x<- rW(100,  2,  1)
y <- rG(100,  2, 2)
z <- rW(100,  6,  9)
r.tc_vus(x=x,y=y,z=z,
        init_param=c(alpha1=2,beta1=1,alpha2=2,beta2=2,
                     alpha3=6,beta3=9),
        model=c("WGW"), method=c("MLE"))
x<- rW(100,  2,  1)
y <- rG(100,  2, 2)
z <- rW(100,  6,  9)
r.tc_index(x=x,y=y,z=z,
          init_param=c(alpha1=2,beta1=1,alpha2=2,
                       beta2=2,alpha3=6,beta3=9),
          init_index=c(median(x),median(y)),
```

```
            model=c("WGW"),
            method=c("MLE"))
x<- rW(100,  2,  1)
y <- rG(100,  2, 2)
z <- rW(100,  6,  9)
r.tc_graph(x=x,y=y,z=z,
            init_param=c(alpha1=2,beta1=1,alpha2=2,
                         beta2=2,alpha3=6,beta3=9),
            empirical=FALSE,model=c("WGW"),
            method=c("MLE"))
```

Index