# Package: ProbYX (via r-universe)

September 3, 2024

**Type** Package

**Title** Inference for the Stress-Strength Model R = P(Y<X)

**Version** 1.1-0.1

**Date** 2015-12-17

**Author** Giuliana Cortese

**Maintainer** Giuliana Cortese <gcortese@stat.unipd.it>

**Depends** R (>= 3.0-0), rootSolve

**Description** Confidence intervals and point estimation for R under various parametric model assumptions; likelihood inference based on classical first-order approximations and higher-order asymptotic procedures.

**License** GPL-2

**LazyLoad** yes

**Imports** stats, graphics

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2022-06-21 08:57:52 UTC

## Contents

---

`ProbYX-package` *Inference on the stress-strength model R = P(Y<X)*

---

### Description

Compute confidence intervals and point estimates for R, under parametric model assumptions for Y and X. Y and X are two independent continuous random variables from two different populations.

### Details

| | |
|---|---|
| Package: | ProbYX |
| Type: | Package |
| Version: | 1.1 |
| Date: | 2012-03-20 |
| License: | GPL-2 |
| LazyLoad: | yes |

The package can be used for computing accurate confidence intervals and point estimates for the stress-strength (reliability) model R = P(Y<X); maximum likelihood estimates, Wald statistic, signed log-likelihood ratio statistic and its modified version ca be computed.
The main function is `Prob`, which evaluates confidence intervals and point estimates under different approaches and parametric assumptions.

### Author(s)

Giuliana Cortese

Maintainer: Giuliana Cortese <gcortese@stat.unipd.it>

### References

Cortese G., Ventura L. (2013). Accurate higher-order likelihood inference on P(Y<X). Computational Statistics, 28:1035-1059.

Kotz S, Lumelskii Y, Pensky M. (2003). The Stress-Strength Model and its Generalizations. Theory and Applications. World Scientific, Singapore.

### Examples

```
# data from the first population
Y <- rnorm(15, mean=5, sd=1)
# data from the second population
X <- rnorm(10, mean=7, sd=1.5)
   level <- 0.01        # \eqn{\alpha} level
   # estimate and confidence interval under the assumption of two
   # normal variables with different variances.
Prob(Y, X, "norm_DV", "RPstar", level)
# method has to be set equal to "RPstar".
```

loglik *Log-likelihood of the bivariate distribution of (Y,X)*

## Description

Computation of the log-likelihood function of the bivariate distribution (Y,X). The log-likelihood is reparametrized with the parameter of interest $\psi$, corresponding to the quantity R, and the nuisance parameter $\lambda$.

## Usage

```
loglik(ydat, xdat, lambda, psi, distr = "exp")
```

## Arguments

| | |
|---|---|
| `ydat` | data vector of the sample measurements from Y. |
| `xdat` | data vector of the sample measurements from X. |
| `lambda` | nuisance parameter vector, $\lambda$. Values can be determined from the reparameterisation of the original parameters of the bivariate distribution chosen in `distr`. |
| `psi` | scalar parameter of interest, $\psi$, for the probability R. Value can be determined from the reparameterisation of the original parameters of the bivariate distribution chosen in `distr`. |
| `distr` | character string specifying the type of distribution assumed for $X_1$ and $X_2$. Possible choices for `distr` are "exp" (default) for the one-parameter exponential, "norm_EV" and "norm_DV" for the Gaussian distribution with, respectively, equal or unequal variances assumed for the two random variables. |

## Details

For further information on the random variables Y and X, see help on `Prob`.
Reparameterisation in order to determine $\psi$ and $\lambda$ depends on the assumed distribution. Here the following relashonships have been used:

**Exponential models:** $\psi = \frac{\alpha}{(\alpha+\beta)}$ and $\lambda = \alpha + \beta$, with $Y \sim e^{\alpha}$ and $X \sim e^{\beta}$;

**Gaussian models with equal variances:** $\psi = \Phi\left(\frac{\mu_2-\mu_1}{\sqrt{2\sigma^2}}\right)$ and $\lambda = (\lambda_1, \lambda_2) = (\frac{\mu_1}{\sqrt{2\sigma^2}}, \sqrt{2\sigma^2})$, with $Y \sim N(\mu_1, \sigma^2)$ and $X \sim N(\mu_2, \sigma^2)$;

**Gaussian models with unequal variances:** $\psi = \Phi\left(\frac{\mu_2-\mu_1}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)$ and $\lambda = (\lambda_1, \lambda_2, \lambda_3) = (\mu_1, \sigma_1^2, \sigma_2^2)$, with $Y \sim N(\mu_1, \sigma_1^2)$ and $X \sim N(\mu_2, \sigma_2^2)$.

The Standard Normal cumulative distribution function is indicated with $\Phi$.

## Value

Value of the log-likelihood function computed in $\psi =$`psi` and $\lambda =$`lambda`.

### Author(s)

Giuliana Cortese

### References

Cortese G., Ventura L. (2013). Accurate higher-order likelihood inference on $P(Y < X)$. Computational Statistics, 28:1035-1059.

### See Also

MLEs

### Examples

```
# data from the first population
Y <- rnorm(15, mean=5, sd=1)
   # data from the second population
X <- rnorm(10, mean=7, sd=1)
   mu1 <- 5
   mu2 <- 7
   sigma <- 1
   # parameter of interest, the R probability
   interest <- pnorm((mu2-mu1)/(sigma*sqrt(2)))
   # nuisance parameters
   nuisance <- c(mu1/(sigma*sqrt(2)), sigma*sqrt(2))
   # log-likelihood value
   loglik(Y, X, nuisance, interest, "norm_EV")
```

---

## MLEs — *Maximum likelihood estimates of the stress-strength model R = P(Y<X).*

---

### Description

Compute maximum likelihood estimates of R, considered as the parameter of interest. Maximum likelihood estimates of the nuisance parameter are also supplied.

### Usage

```
MLEs(ydat, xdat, distr)
```

### Arguments

| | |
|---|---|
| `ydat` | data vector of the sample measurements from Y. |
| `xdat` | data vector of the sample measurements from X. |
| `distr` | character string specifying the type of distribution assumed for Y and X. Possible choices for `distr` are "exp" (default) for the one-parameter exponential, "norm_EV" and "norm_DV" for the Gaussian distribution with, respectively, equal or unequal variances assumed for the two random variables. |

### Details

The two independent random variables Y and X with given distribution `distr` are measurements of a certain characteristics on two different populations. For the relationship of the parameter of interest (R) and nuisance parameters with the original parameters of `distr`, look at the details in `loglik`.

### Value

Vector of estimetes of the nuisance parameters and the R quantity (parameter of interest), respectively.

### Author(s)

Giuliana Cortese

### References

Kotz S, Lumelskii Y, Pensky M. (2003). The Stress-Strength Model and its Generalizations. Theory and Applications. World Scientific, Singapore.

### See Also

`loglik`, `Prob`

### Examples

```
# data from the first population
Y <- rnorm(15, mean=5, sd=1)
# data from the second population
   X <- rnorm(10, mean=7, sd=1.5)
   # vector of MLEs for the nuisance parameters and the quantity R
   MLEs(Y, X, "norm_DV")
```

---

## Prob     *Estimation of the stress-strength model R = P(Y<X)*

---

### Description

Compute confidence intervals and point estimates for the probability R, under parametric model assumptions for Y and X. Y and X are two independent continuous random variable from two different populations.

### Usage

```
Prob(ydat, xdat, distr = "exp", method = "RPstar", level = 0.05)
```

## Arguments

| | |
|---|---|
| ydat | data vector of the sample measurements from Y. |
| xdat | data vector of the sample measurements from X. |
| distr | character string specifying the type of distribution assumed for Y and X. Possible choices for `distr` are "exp" (default) for the one-parameter exponential, "norm_EV" and "norm_DV" for the Gaussian distribution with, respectively, equal or unequal variances assumed for the two random variables. |
| method | character string specifying the methodological approach used for inference (confidence intervals and point estimates) on the AUC. The argument `method` can be set equal to "Wald", "RP" or RPstar" (default), according as inference is based on the Wald statistic, the signed log-likelihood ratio statistic (directed likelihhod, $r_p$) or the modified signed log-likelihood ratio statistic (modified directed likelihood, $r_p^*$), respectively. |
| level | it is the $\alpha$ that supplies the nominal level $(1-\alpha)$ chosen for the confidence interval. |

## Value

| | |
|---|---|
| PROB | Point estimate of $R = P(Y < X)$. This value corresponds to the maximum likelihoos estimate if method "Wald" or "RP" is chosen; otherwise, when method "RPstar" is selected, estimate is obtained from the estimating equaltion $r_p^* = 0$. |
| C.Interval | Confidence interval of R at confidence level $(1-\alpha)$. |

## Author(s)

Giuliana Cortese

## References

Cortese G., Ventura L. (2013). Accurate higher-order likelihood inference on $R = P(Y < X)$. Computational Statistics, 28:1035-1059.

## See Also

`wald`, `rp`, `rpstar`

## Examples

```
# data from the first population
Y <- rnorm(15, mean=5, sd=1)
# data from the second population
X <- rnorm(10, mean=7, sd=1.5)
   level <- 0.01                  ## \eqn{\alpha} level
   # estimate and confidence interval under the assumption of two
   # normal variables with different variances.
Prob(Y, X, "norm_DV", "RPstar", level)
# method has to be set equal to "RPstar".
```

ROC.plot *Estimated ROC curves*

## Description

Plot of ROC curves estimated under parametric model assumptions on the continuous diagnostic marker.

## Usage

```
ROC.plot(ydat, xdat, distr = "exp", method = "RPstar", mc = 1)
```

## Arguments

| | |
|---|---|
| ydat | data vector of the diagnostic marker measurements on the sample of non-diseased individuals (from Y). |
| xdat | data vector of the diagnostic marker measurements on the sample of diseased individuals (from X). |
| distr | character string specifying the type of distribution assumed for Y and X. Possible choices for `distr` are "exp" (default) for the one-parameter exponential, "norm_EV" and "norm_DV" for the Gaussian distribution with, respectively, equal or unequal variances assumed for the two random variables. |
| method | character string specifying the methodological approach used for estimating the probability R, which is here interpreted as the area under the ROC curve (AUC). The argument `method` can be set equal to "Wald", "RP" or RPstar" (default), according as inference is based on the Wald statistic, the signed log-likelihood ratio statistic (directed likelihhod, $r_p$) or the modified signed log-likelihood ratio statistic (modified directed likelihood, $r_p^*$), respectively. For estimating the ROC curve parametrically, methods "Wald" and "RP" are equivalent and supply maximum likelihood estimation (MLE), whereas, by using method "RPstar", estimate of the ROC curve is based on the modified signed log-likelihood ratio statistic ($r_p^*$). See `rpstar` for details on this statistic. |
| mc | a numeric value indicating single or multiple plots in the same figure. In case `mc` is equal to 1 (default), only the method specified in `method` is applied and the corresponding estimated ROC curve is plotted. If `mc` is different from 1, both MLE and $r_p^*$-based methods are applied, and two differently estimated ROC curves are plotted. |

## Details

If `mc` is different from 1, `method` does not need to be specified.

## Value

Plot of ROC curves

## Note

The two independent random variables Y and X with given distribution `distr` are measurements of the diagnostic marker on the diseased and non-diseased subjects, respectively.
In "Wald" method, or equivalently "RP" method, MLEs for parameters of the Y and X distributions are computed and then used to estimate specificity and sensitivity. These measures are evaluated as $P(Y < t)$ and $P(X > t)$, respectively.
In "RPstar" method, parameters of the Y and X distributions are estimated from the $r_p^*$-based estimate of the AUC.

## Author(s)

Giuliana Cortese

## References

Cortese G., Ventura L. (2013). Accurate higher-order likelihood inference on $P(Y < X)$. Computational Statistics, 28:1035-1059.

## See Also

Prob

## Examples

```
# data from the non-diseased population
Y <- rnorm(15, mean=5, sd=1)
# data from the diseased population
X <- rnorm(10, mean=7, sd=1.5)
 ROC.plot(Y, X, "norm_DV", method = "RP", mc = 2)
```

---

# rp *Signed log-likelihood ratio statistic*

---

## Description

Compute the signed log-likelihood ratio statistic ($r_p$) for a given value of the stress strength R = P(Y<X), that is the parameter of interest, under given parametric model assumptions.

## Usage

```
rp(ydat, xdat, psi, distr = "exp")
```

### Arguments

| | |
|---|---|
| ydat | data vector of the sample measurements from Y. |
| xdat | data vector of the sample measurements from X. |
| psi | scalar for the parameter of interest. It is the value of R, treated as a parameter under the parametric model construction. |
| distr | character string specifying the type of distribution assumed for Y and X. Possible choices for `distr` are "exp" (default) for the one-parameter exponential, "norm_EV" and "norm_DV" for the Gaussian distribution with, respectively, equal or unequal variances assumed for the two random variables. |

### Details

The two independent random variables Y and X with given distribution `distr` are measurements of the diagnostic marker on the diseased and non-diseased subjects, respectively. For the relationship of the parameter of interest (R) and nuisance parameters with the original parameters of `distr`, look at the details in `loglik`.

### Value

Value of the signed log-likelihood ratio statistic $r_p$.

### Note

The $r_p$ values can be also used for testing statistical hypotheses on the probability R.

### Author(s)

Giuliana Cortese

### References

Cortese G., Ventura L. (2013). Accurate higher-order likelihood inference on P(Y<X). Computational Statistics, 28:1035-1059.

Severini TA. (2000). Likelihood Methods in Statistics. Oxford University Press, New York.

Brazzale AR., Davison AC., Reid N. (2007). Applied Asymptotics. Case-Studies in Small Sample Statistics. Cambridge University Press, Cambridge.

### See Also

`wald`, `rpstar`, `MLEs`, `Prob`

### Examples

```
# data from the first population
Y <- rnorm(15, mean=5, sd=1)
# data from the second population
X <- rnorm(10, mean=7, sd=1.5)
```

```
# value of \eqn{r_p} for \code{psi=0.9}
rp(Y, X, 0.9,"norm_DV")
```

---

rpstar *Modified signed log-likelihood ratio statistic*

---

### Description

Compute the modified signed log-likelihood ratio statistic ($r_p^*$) for a given value of the stress strength R = P(Y<X), that is the parameter of interest, under given parametric model assumptions.

### Usage

```
rpstar(ydat, xdat, psi, distr = "exp")
```

### Arguments

| | |
|---|---|
| ydat | data vector of the sample measurements from Y. |
| xdat | data vector of the sample measurements from X. |
| psi | scalar for the parameter of interest. It is the value of R, treated as a parameter under the parametric model construction. |
| distr | character string specifying the type of distribution assumed for Y and X. Possible choices for `distr` are "exp" (default) for the one-parameter exponential, "norm_EV" and "norm_DV" for the Gaussian distribution with, respectively, equal or unequal variances assumed for the two random variables. |

### Details

The two independent random variables Y and X with given distribution `distr` are measurements from two different populations. For the relationship of the parameter of interest (R) and nuisance parameters with the original parameters of `distr`, look at the details in `loglik`.

### Value

| | |
|---|---|
| rp | Value of the signed log-likelihood ratio statistic $r_p$. |
| rp_star | Value of the modified signed log-likelihood ratio statistic $r_p^*$. |

### Note

The statistic $r_p^*$ is a modified version of $r_p$ which provides more statistically accurate estimates. The $r_p^*$ values can be also used for testing statistical hypotheses on the probability R.

### Author(s)

Giuliana Cortese

### References

Cortese G., Ventura L. (2013). Accurate higher-order likelihood inference on P(Y<X). Computational Statistics, 28:1035-1059.

Severini TA. (2000). Likelihood Methods in Statistics. Oxford University Press, New York.

Brazzale AR., Davison AC., Reid N. (2007). Applied Asymptotics. Case-Studies in Small Sample Statistics. Cambridge University Press, Cambridge.

### See Also

wald, rp, MLEs, Prob

### Examples

```
# data from the first population
Y <- rnorm(15, mean=5, sd=1)
# data from the second population
X <- rnorm(10, mean=7, sd=1.5)
# value of \eqn{r_p^*} for \code{psi=0.9}
rpstar(Y, X, 0.9,"norm_DV")
# method has be set equal to "RPstar".
```

---

`wald` *Wald statistic*

---

### Description

Compute the Wald statistic for a given value of the stress-strength R = P(Y<X), that is the parameter of interest, under given parametric model assumptions.

### Usage

```
wald(ydat, xdat, psi, distr = "exp")
```

### Arguments

| | |
|---|---|
| `ydat` | data vector of the sample measurements from Y. |
| `xdat` | data vector of the sample measurements from X. |
| `psi` | scalar for the parameter of interest. It is the value of the quantity R, treated as a parameter under the parametric model construction. |
| `distr` | character string specifying the type of distribution assumed for Y and X. Possible choices for `distr` are "exp" (default) for the one-parameter exponential, "norm_EV" and "norm_DV" for the Gaussian distribution with, respectively, equal or unequal variances assumed for the two random variables. |

## Details

The two independent random variables Y and X with given distribution `distr` are measurements from two different populations. For the relationship of the parameter of interest (R) and nuisance parameters with the original parameters of `distr`, look at the details in `loglik`.

## Value

| | |
|---|---|
| `Wald` | Value of the Wald statistic for a given `psi` |
| `Jphat` | Observed profile Fisher information |

## Note

Values of the Wald statistic can be also used for testing statistical hypotheses on the probability R.

## Author(s)

Giuliana Cortese

## References

Cortese G., Ventura L. (2013). Accurate higher-order likelihood inference on P(Y<X). Computational Statistics, 28:1035-1059.

Brazzale AR., Davison AC., Reid N. (2007). Applied Asymptotics. Case-Studies in Small Sample Statistics. Cambridge University Press, Cambridge.

## See Also

`rp`, `rpstar`, `MLEs`, `Prob`

## Examples

```
# data from the first population
Y <- rnorm(15, mean=5, sd=1)
# data from the second population
X <- rnorm(10, mean=7, sd=1.5)
# value of Wald for \code{psi=0.9}
wald(Y, X, 0.9,"norm_DV")
```

# Index