Package ‘depend.truncation’

August 11, 2017

Type Package
Title Statistical Inference for Parametric and Semiparametric Models Based on Dependently Truncated Data
Version 2.8
Date 2017-08-11
Author Takeshi Emura
Maintainer Takeshi Emura <takeshiemura@gmail.com>
Description Suppose that one can observe bivariate random variables \((X, Y)\) only when \(X \leq Y\) holds. Data \((X_j, Y_j)\), subject to \(X_j \leq Y_j\), for all \(j = 1, \ldots, n\), are called truncated data. For truncated data, several different approaches are implemented for statistical inference on \((X, Y)\), when \(X\) and \(Y\) are dependent. Also included is truncated data on the number of deaths at each year (1963-1980) for Japanese male centenarians.
License GPL-2
Depends mvtnorm
NeedsCompilation no
Repository CRAN
Date/Publication 2017-08-11 10:30:02 UTC

R topics documented:

- depend.truncation-package
- centenarian
- CHAIEB.Clayton
- CHAIEB.Frank
- dependAFT.reg
- EMURA.Clayton
- EMURA.Frank
- NPMLE.Frank
- NPMLE.Indep
- NPMLE.Normal
- NPMLE.Plackett
Statistical Inference for Parametric and Semiparametric Models Based on Dependent Truncated Data

Description
Suppose that one can observe bivariate random variables (X, Y) only when X ≤ Y holds. Data (X_j, Y_j), subject to X_j ≤ Y_j for all j = 1,...,n, are called truncated data (or truncation data). This package performs parametric and semiparametric inference methods for models on (X, Y) where X and Y are not independent. Semi-parametric approaches are based on Chaieb et al. (2006), Emura et al. (2011), Emura and Wang (2012) and Emura and Murotani (2015). Parametric approaches are based on Emura and Konno (2012). A regression approach is based on Emura and Wang (2016). This package also includes truncated data on the number of deaths at each year (1963-1980) for Japanese male centenarians (Emura and Murotani 2015).

Details

Package: depend.truncation
Type: Package
Version: 2.7
Date: 2017-08-11
License: GPL-2

Author(s)
Takeshi Emura <takeshiemura@gmail.com>

References

---

Japanese Centenarians Data

Description

The number of deaths at each year (1963-1980) for Japanese male centenarians (Table 2 of Emura and Murotani (2015)). See also the original reference Sibuya & Hanayama (2004).

Usage

data("centenarian")

Format

A data frame with 21 observations on the following 19 variables.

x : the age at death
x1963 : the number of deaths between 1963 and 1964
x1964 : the number of deaths between 1964 and 1965
x1965 : the number of deaths between 1965 and 1966
x1966 : the number of deaths between 1966 and 1967
x1967 : the number of deaths between 1967 and 1968
x1968 : the number of deaths between 1968 and 1969
x1969 : the number of deaths between 1969 and 1970
x1970 : the number of deaths between 1970 and 1971
x1971 : the number of deaths between 1971 and 1972
x1972 : the number of deaths between 1972 and 1973
x1973 : the number of deaths between 1973 and 1974
x1974 : the number of deaths between 1974 and 1975
x1975 : the number of deaths between 1975 and 1976
x1976 : the number of deaths between 1976 and 1977
x1977 : the number of deaths between 1977 and 1978
x1978 : the number of deaths between 1978 and 1979
x1979 : the number of deaths between 1979 and 1980
x1980 : the number of deaths between 1980 and 1981
References


Examples

```r
## Below is the centenarians data analysis of Emura & Murotani (2015) ##
data(centenarian)
Death=centenarian[,1]
Year=1963:1980
data.mat=centenarian[,-1]

X=rep(0,6)
year=c(1,2,3,4,5)
data.mat=data.mat[1:6,]
data.Nmat=data.mat
x=x[1]
t=t[1]
for(i in 1:length(Death)){
  for(j in 1:length(Year)){
    x=c( x,rep(Death[i],data.mat[i,j]) )
    t=c( t,rep(Year[j]-i+1,data.mat[i,j]) )
  }
}
x.trunc=x
z.trunc=max(Year)+0.5-T+100
m=length(x.trunc)
d=rep(1,m)
set.seed(1)
x.trunc=x.trunc+runif(length(x.trunc),min=-0.01,max=0)
z.trunc=z.trunc+runif(length(z.trunc),min=0,max=0.01)

## Copula-based estimator ####
# CHAIEB.Frank(x.trunc, z.trunc, d, a = 1/10)
```

Description

A copula-based estimation based on dependent truncation data under the Clayton copula model (Lakhal-Chaieb, Rivest & Abdous 2006; Emura & Murotani 2015). The forms of the marginal distributions for X and Y are completely unspecified, but the parametric form of copula is specified as the Clayton copula.

Usage

`CHAIEB.Clayton(x.trunc, z.trunc, d, a = 1/10, plotX = TRUE, plotY = TRUE)`
Arguments

- **x.trunc**: vector of variables satisfying \( x.trunc \leq z.trunc \)
- **z.trunc**: vector of variables satisfying \( x.trunc \leq z.trunc \)
- **d**: censoring indicator (0=censoring, 1=failure) for \( z.trunc \)
- **a**: tuning parameter adjusting for small the risk sets (pp.360-361: Emura, Wang & Hung 2011 Sinica)
- **plotX**: if TRUE, plot the distribution function of \( X \)
- **plotY**: if TRUE, plot the survival function of \( Y \)

Details

The function produces the moment-based estimate for the marginal distributions and the estimate of the association parameter under the Clayton copula model. The method can handle right-censoring for \( Y \) in which \( Z = \min(Y, C) \) and \( I(Y \leq C) \) are observed with censoring variable \( C \).

Value

- **alpha**: association parameter
- **tau**: Kendall’s tau between \( X \) and \( Y \)
- **c**: inclusion probability, defined by \( c = \Pr(X \leq Z) \)
- **Fx**: marginal distribution function of \( X \) at (ordered) observed points of \( X \)
- **Sy**: marginal survival function of \( Y \) at (ordered) observed points of \( Y \)

Author(s)

Takeshi EMURA

References


Examples

```
CCCCCCCCCCC aids data of kalbfleisch & lawless (1989, jasa) CCCCCCCCCCCC
inf2=c(36, 12, 50, 45, 33, 55, 37, 46, 62, 57, 34, 56, 57, 64, 57, 42, 54, 28, 36, 66, 19, 45, 53, 67, 68, 54, 40, 60, 54, 42, 74, 71, 43, 61, 68, 52, 75, 46, 62, 67, 48, 68, 58, 55, 70, 33, 56, 47, 64, 15)
inf4=c(46, 68, 76, 64, 50, 59, 80, 46, 78, 26, 62, 19, 22, 26, 76, 27, 62, 75, 76, 57, 58, 59, 8, 41, 70, 57, 58, 63, 37, 75, 58, 39, 38, 73, 72, 41, 56, 50, 79, 83, 76, 29, 17, 69, 86, 29, 65, 75, 74, 42)
inf5=c(65, 61, 84, 41, 83, 58, 45, 83, 80, 84, 59, 85, 59, 65, 47, 64, 73, 81, 79, 36, 67, 87, 85, 29, 72, 77, 72, 67, 53, 54, 17, 61, 65, 48, 57, 29, 36, 30, 45, 40, 75, 43, 76, 66, 86, 57, 75, 71, 51)
```
Description

A copula-based estimation based on dependent truncation data under the Frank copula model (Lakhal-Chaieb, Rivest & Abdous 2006; Emura & Murotani 2015). The forms of the marginal distributions for X and Y are completely unspecified, but the parametric form of copula is specified as the Frank copula.
**Usage**

`CHAIEB.Frank(x.trunc, z.trunc, d, a = 1/10, plotX = TRUE, plotY = TRUE)`

**Arguments**

- `x.trunc`: vector of variables satisfying `x.trunc<=z.trunc`
- `z.trunc`: vector of variables satisfying `x.trunc<=z.trunc`
- `d`: censoring indicator(0=censoring,1=failure) for `z.trunc`
- `a`: tuning parameter adjusting for small the risk sets (pp.360-361: Emura, Wang & Hung 2011 Sinica)
- `plotX`: if TRUE, plot the distribution function of X
- `plotY`: if TRUE, plot the survival function of Y

**Details**

The function produces the moment-based estimate for the marginal distributions and the estimate of the association parameter under the Frank copula model. The method can handle right-censoring for Y in which Z=min(Y, C) and I(Y<=C) are observed with censoring variable C.

**Value**

- `alpha`: association parameter
- `tau`: Kendall’s tau between X and Y
- `c`: inclusion probability, defined by `c=Pr(X<=Z)`
- `Fx`: marginal distribution function of X at at (ordered) observed points of X
- `Sy`: marginal survival function of Y at (ordered) observed points of Y

**Author(s)**

Takeshi EMURA

**References**


**Examples**

```
# AIDS data of Kalbfleisch & Lawless (1989, JASA)
inf2=c(36, 12, 58, 45, 33, 55, 37, 46, 62, 57, 34, 56, 57, 64, 57, 42, 54, 28, 36, 66, 19, 45, 53, 67, 68, 54, 48, 60, 54, 42, 74, 71, 43, 61, 68, 52, 75, 46, 62, 67, 48, 68, 58, 55, 70, 33, 56, 47, 64, 15)
```
Regression estimation for AFT regression models based on left-truncated and right-censored data, which is proposed by Emura and Wang (2015). The dependency of truncation on lifetime is modeled through the AFT regression form.

**Usage**

```R
dependAFT.reg(t.trunc, y.trunc, d, x1.trunc, initial = c(0, 0), LY = FALSE, beta1_low = -0.2, beta1_up = 0.2, alpha = 1, epsilon = 1/50)
```

---

### Description

Regression estimation for AFT regression models based on left-truncated and right-censored data, which is proposed by Emura and Wang (2015). The dependency of truncation on lifetime is modeled through the AFT regression form.

### Usage

```R
dependAFT.reg(t.trunc, y.trunc, d, x1.trunc, initial = c(0, 0), LY = FALSE, beta1_low = -0.2, beta1_up = 0.2, alpha = 1, epsilon = 1/50)
```
Arguments

- `t.trunc`: vector of left-truncation variables satisfying \( t.trunc \leq y.trunc \)
- `y.trunc`: vector of lifetime variables satisfying \( t.trunc \leq y.trunc \)
- `d`: vector of censoring indicators
- `x1.trunc`: vector of 1-dimensional covariates
- `initial`: a pair of initial values for (\( \beta, \gamma \))
- `LY`: Lai and Ying’s estimator for initial values
- `beta1_low`: lower bound for \( \beta \)
- `beta1_up`: upper bound for \( \beta \)
- `alpha`: some tuning parameter for optimization, \( \alpha = 1 \) is default
- `epsilon`: some tuning parameter for kernel methods

Details

Only the univariate regression (only one covariate) is allowed.

Value

- `beta`: inference results for \( \beta \)
- `gamma`: inference results for \( \gamma \)
- `beta_LY`: the estimator of Lai & Ying (1991)
- `S2_Minimum`: minimum of the objective function
- `detail`: detailed results for minimizing the estimating objective function "optim"

Author(s)

Takeshi Emura

References


Examples

```r
y.trunc=c(-0.52, 0.22, -1.42, 0.05, 0.32, -1.02, -0.47, 0.10, -0.38, -0.18, 0.97, 0.04, -0.10, 0.50, 0.57, -0.80, -0.24, 0.07, -0.04, 0.88, -0.52, -0.28, -0.55, 0.53, 0.99, -0.52, -0.59, -0.48, -0.07, 0.20, -0.34, 1.00, -0.52)
t.trunc=c(-2.05, -0.25, -2.43, -0.32, -0.27, -1.06, -0.95, -0.82, -0.66, -0.28, -1.14, -0.32, -1.19, -2.18, -0.45, -1.71, -0.84, -1.93, -1.04, -2.58, -1.97, -2.15, -0.59, -0.74, -1.26, -2.57, ...)```
EMURA.Clayton

Semi-parametric Inference for Copula Models with Dependently Truncated Data

Description

A copula-based estimation based on dependent truncation data under the Clayton copula model (Emura, Wang & Hung 2011; Emura & Murotani 2015). The forms of the marginal distributions for X and Y are completely unspecified, but the parametric form of copula is specified as the Clayton copula.

Usage

EMURA.Clayton(x.trunc, z.trunc, d, a = 1/10, plotX = TRUE, plotY = TRUE)

Arguments

x.trunc vector of variables satisfying x.trunc<=z.trunc
z.trunc vector of variables satisfying x.trunc<=z.trunc
d censoring indicator(0=censoring,1=failure) for z.trunc
a tuning parameter adjusting for small the risk sets (pp.360-361: Emura, Wang & Hung 2011 Sinica)
plotX if TRUE, plot the distribution function of X
plotY if TRUE, plot the survival function of Y
Details

The function produces the moment-based estimate for the marginal distributions and the estimate of the association parameter under the Clayton copula model. The method can handle right-censoring for \( Y \) in which \( Z = \min(Y, C) \) and \( I(Y \leq C) \) are observed with censoring variable \( C \).

Value

- \( \alpha \): association parameter
- \( \tau \): Kendall’s tau between \( X \) and \( Y \)
- \( \alpha \): association parameter
- \( c \): inclusion probability, defined by \( c = \Pr(X \leq Z) \)
- \( F_X \): marginal distribution function of \( X \) at (ordered) observed points of \( X \)
- \( S_Y \): marginal survival function of \( Y \) at (ordered) observed points of \( Y \)

Author(s)

Takeshi Emura

References


Examples

```
# AIDS data of Kalbfleisch & Lawless (1989, JASA)
inf1 = c(23, 38, 38, 36, 27, 45, 23, 48, 25, 42, 33, 45, 33, 39, 34, 48, 50, 35, 26, 43, 3, 53, 48, 53,
28, 34, 42, 66, 19, 21, 37, 33, 31, 48, 32, 43, 17, 64, 58, 67, 49, 67, 35, 12, 19, 60, 53, 56, 65, 53)
inf2 = c(36, 12, 50, 45, 33, 55, 37, 46, 62, 57, 34, 56, 57, 64, 57, 42, 54, 28, 36, 66, 19, 45, 53, 67,
68, 54, 40, 60, 54, 42, 74, 71, 43, 61, 68, 52, 75, 46, 62, 67, 48, 68, 58, 55, 70, 33, 56, 47, 64, 15)
inf3 = c(41, 35, 67, 35, 71, 67, 36, 69, 45, 76, 66, 22, 74, 49, 42, 15, 29, 62, 65, 75, 65, 69, 62, 56,
52, 82, 46, 75, 27, 56, 70, 49, 66, 73, 76, 43, 50, 41, 49, 68, 39, 67, 61, 82, 69, 65, 56, 59, 57, 57)
inf4 = c(46, 68, 76, 64, 50, 59, 80, 46, 78, 26, 62, 19, 22, 26, 76, 27, 62, 75, 76, 57, 58, 59, 6, 41,
70, 57, 58, 63, 37, 75, 58, 39, 38, 73, 72, 41, 56, 50, 79, 83, 76, 29, 17, 69, 86, 29, 65, 75, 74, 42)
inf5 = c(65, 61, 84, 41, 83, 58, 45, 83, 80, 84, 59, 85, 59, 65, 47, 64, 73, 81, 79, 36, 67, 87, 85, 29,
72, 77, 72, 67, 53, 47, 61, 65, 48, 57, 29, 36, 45, 40, 75, 43, 76, 66, 86, 57, 75, 71, 51)
inf6 = c(60, 41, 53, 48, 36, 72, 60, 36, 48, 55, 58, 60, 89, 37, 82, 41, 68, 71, 63, 49, 37, 60, 78, 52,
60, 85, 68, 37, 39, 22, 12, 63, 80, 45, 47, 85, 29, 60, 84, 76, 61, 69, 77, 63)
x1 = c(27, 14, 15, 18, 28, 10, 34, 17, 34, 17, 29, 17, 29, 23, 29, 15, 13, 29, 38, 21, 61, 12, 25, 12, 38,
32, 24, 0, 48, 46, 30, 34, 37, 21, 37, 26, 53, 6, 13, 4, 22, 4, 37, 60, 53, 13, 20, 17, 8, 20)
x2 = c(38, 62, 24, 29, 41, 19, 37, 28, 13, 18, 41, 19, 18, 11, 18, 33, 21, 48, 40, 10, 58, 32, 24, 11, 10,
25, 19, 19, 25, 37, 5, 8, 37, 20, 13, 29, 6, 35, 19, 14, 33, 13, 23, 27, 12, 49, 26, 35, 18, 68)
x3 = c(42, 48, 16, 48, 12, 16, 47, 14, 38, 8, 18, 62, 10, 35, 42, 69, 55, 22, 20, 10, 20, 16, 23, 30, 34, 4,
40, 11, 59, 30, 16, 37, 28, 13, 10, 43, 36, 48, 38, 19, 48, 20, 26, 5, 18, 23, 32, 29, 31, 31)
x4 = c(42, 28, 12, 24, 38, 29, 8, 43, 11, 63, 27, 70, 67, 63, 13, 62, 27, 15, 14, 33, 32, 32, 33, 50, 21,
34, 33, 29, 55, 17, 34, 53, 54, 19, 20, 52, 37, 43, 14, 10, 17, 64, 76, 25, 8, 65, 29, 19, 20, 52)
x5 = c(29, 33, 10, 53, 11, 36, 49, 12, 15, 11, 36, 10, 36, 31, 49, 32, 23, 15, 17, 60, 29, 9, 11, 67, 24,
...
Semi-parametric Inference for Copula Models with Dependently Truncated Data

Description

A copula-based estimation based on dependent truncation data under the Frank copula model (Emura, Wang & Hung 2011; Emura & Murotani 2015). The forms of the marginal distributions for X and Y are completely unspecified, but the parametric form of copula is specified as the Frank copula.

Usage

EMURA.Frank(x.trunc, z.trunc, d, a = 1/10, plotX = TRUE, plotY = TRUE)

Arguments

- x.trunc: vector of variables satisfying x.trunc<=z.trunc
- z.trunc: vector of variables satisfying x.trunc<=z.trunc
- d: censoring indicator(0=censoring,1=failure) for z.trunc
- a: tuning parameter adjusting for small the risk sets (pp.360-361: Emura, Wang & Hung 2011 Sinica)
- plotX: if TRUE, plot the distribution function of X
- plotY: if TRUE, plot the survival function of Y
Details

The function produces the moment-based estimate for the marginal distributions and the estimate of the association parameter under the Frank copula model. The method can handle right-censoring for Y in which \( Z = \min(Y, C) \) and \( I(Y \leq C) \) are observed with censoring variable C.

Value

- **alpha**: association parameter
- **tau**: Kendall’s tau between X and Y
- **c**: inclusion probability, defined by \( c = \Pr(X \leq Z) \)
- **Fx**: marginal distribution function of X at (ordered) observed points of X
- **Sy**: marginal survival function of Y at (ordered) observed points of Y

Author(s)

Takeshi EMURA

References


Examples

---------
### AIDS data of Kalbfleisch & Lawless (1989, JASA)
---------

inf2=c(36, 12, 50, 45, 33, 55, 37, 46, 62, 57, 34, 56, 57, 64, 57, 42, 54, 28, 36, 66, 19, 45, 53, 67, 68, 54, 49, 60, 54, 42, 74, 71, 43, 61, 68, 52, 75, 46, 62, 67, 48, 68, 58, 55, 70, 33, 56, 47, 64, 15)
inf4=c(46, 68, 76, 64, 50, 59, 80, 46, 78, 26, 62, 19, 22, 26, 76, 27, 62, 75, 76, 57, 58, 59, 81, 70, 57, 58, 63, 37, 75, 58, 39, 38, 73, 72, 41, 56, 58, 79, 83, 76, 29, 17, 69, 86, 29, 65, 75, 74, 42)
inf5=c(65, 61, 84, 41, 83, 58, 45, 83, 80, 64, 59, 85, 59, 65, 47, 64, 73, 81, 79, 36, 67, 87, 85, 29, 72, 77, 27, 67, 53, 54, 17, 61, 65, 48, 57, 29, 36, 45, 48, 75, 43, 76, 66, 86, 57, 75, 71, 51)
inf6=c(60, 41, 53, 48, 36, 72, 60, 36, 48, 55, 58, 60, 89, 37, 82, 41, 68, 71, 63, 49, 37, 60, 78, 52, 60, 85, 68, 37, 39, 22, 12, 63, 80, 45, 47, 85, 29, 60, 84, 70, 61, 69, 77, 63)
x1=c(27, 14, 15, 18, 28, 10, 34, 17, 34, 17, 29, 17, 29, 23, 29, 15, 13, 29, 38, 21, 61, 12, 25, 12, 38, 32, 24, 0, 48, 46, 30, 34, 21, 37, 26, 53, 6, 13, 4, 22, 4, 37, 60, 53, 13, 20, 17, 8, 20)
x2=c(38, 62, 24, 29, 41, 19, 37, 28, 13, 18, 41, 19, 18, 11, 18, 33, 21, 48, 40, 10, 58, 32, 24, 11, 10, 25, 39, 19, 25, 37, 5, 8, 37, 20, 13, 29, 6, 35, 19, 14, 33, 13, 23, 27, 12, 49, 26, 35, 18, 68)
x3=c(42, 48, 16, 48, 12, 16, 47, 14, 38, 8, 18, 62, 10, 35, 42, 69, 55, 22, 28, 10, 20, 16, 23, 30, 34, 4, 40, 11, 59, 30, 16, 37, 20, 13, 10, 43, 36, 48, 19, 48, 20, 26, 5, 18, 23, 32, 29, 31, 31)
x4=c(42, 29, 48, 12, 24, 38, 29, 8, 43, 11, 63, 27, 70, 67, 63, 13, 62, 27, 15, 14, 33, 32, 32, 83, 50, 21, 34, 33, 29, 55, 17, 34, 53, 54, 19, 20, 52, 37, 43, 14, 10, 17, 64, 76, 25, 8, 65, 29, 19, 20, 52)
x5=c(29, 33, 10, 53, 11, 36, 49, 12, 15, 11, 36, 10, 36, 31, 49, 32, 23, 15, 17, 68, 29, 9, 11, 67, 24, 19, 24, 29, 43, 43, 80, 36, 32, 49, 40, 68, 61, 68, 53, 58, 23, 55, 22, 32, 12, 41, 23, 27, 47)
Nonparametric maximum likelihood estimation for dependent truncation data under the Frank copula models (Emura & Wang, 2012). The forms of the marginal distributions for X and Y are completely unspecified, but the parametric form of copula is specified as the Frank copula.

Usage

NPMLE.Frank(x.trunc, y.trunc,
  x.fix = median(x.trunc), y.fix = median(y.trunc), plotX = TRUE)

Arguments

x.trunc vector of variables satisfying x.trunc<=y.trunc
y.trunc vector of variables satisfying x.trunc<=y.trunc
x.fix vector of fixed points at which marginal distribution function of X is calculated
y.fix vector of fixed points at which marginal survival function of Y are calculated
plotX if TRUE, the cumulative distribution function for X is plotted
Details

The function produces the nonparametric maximum likelihood estimate (NPMLE) for the marginal distributions and the estimate of the association parameter under the Frank copula model. The method follows Emura & Wang (2012). The maximization of the likelihood (L) is conducted by minimizing -logL by "nlm". Standard errors are calculated from the diagonal elements in the observed Fisher information matrix, which is obtained from the output of the "nlm". The method can handle ties but cannot handle right-censoring.

Value

- **alpha**: estimate of association parameter
- **alpha_se**: standard error of the estimate of association parameter
- **hx**: estimate of the marginal cumulative reverse-hazard function for X
- **hx_se**: standard error of the estimate of the marginal cumulative reverse-hazard function for X
- **Ay**: estimate of the marginal cumulative hazard function for Y
- **Ay_se**: standard error of the estimate of the marginal cumulative hazard function for Y
- **Fx**: estimate of the marginal distribution function for X
- **Fx_se**: standard error of the estimate of the marginal distribution function for X
- **Sy**: estimate of the marginal survival function for Y
- **Sy_se**: standard error of the estimate of the marginal survival function for Y
- **conv**: an integer indicating why the optimization process terminated in "nlm": If conv=1, the maximization of the likelihood function is properly done. Please refer R function "nlm" for more details.
- **iteration**: the number of iterations until convergence
- **Grad**: L_2 norm for the gradient vector at the solution (close to zero if the solution is proper)
- **MinEigen**: Minimum eigenvalue of the Hessian matrix at the solution (positive if the solution is proper)

Author(s)

Takeshi EMURA

References


Examples

```r
x.trunc=c(0.53, 0.43, 1.63, 0.98, 0.62)
y.trunc=c(1.20, 4.93, 2.09, 2.57, 1.52)
NPMLE.Frank(x.trunc, y.trunc, x.fix =1, y.fix = 2, plotX = TRUE)
```
Description

Nonparametric maximum likelihood estimation for dependent truncation data under the independence copula models (Emura & Wang, 2012). The forms of the marginal distributions for X and Y are completely unspecified, but the parametric form of copula is specified as the independence copula.

Usage

```r
NPMLE.Indep(x.trunc, y.trunc,
  x.fix = median(x.trunc), y.fix = median(y.trunc), plotX = TRUE)
```

Arguments

- `x.trunc`: vector of variables satisfying `x.trunc<=y.trunc`
- `y.trunc`: vector of variables satisfying `x.trunc<=y.trunc`
- `x.fix`: vector of fixed points at which marginal distribution function of X is calculated
- `y.fix`: vector of fixed points at which marginal survival function of Y are calculated
- `plotX`: if TRUE, the cumulative distribution function for X is plotted

Details

The function produces the nonparametric maximum likelihood estimate (NPMLE) for the marginal distributions and the estimate of the association parameter under the independence copula model. The method follows Emura & Wang (2012). The maximization of the likelihood (L) is conducted by minimizing -logL by "nlm". Standard errors are calculated from the diagonal elements in the observed Fisher information matrix, which is obtained from the output of the "nlm". The method can handle ties but cannot handle right-censoring.

Value

- `Hx`: estimate of the marginal cumulative reverse-hazard function for X
- `Hx_se`: standard error of the estimate of the marginal cumulative reverse-hazard function for X
- `Ay`: estimate of the marginal cumulative hazard function for Y
- `Ay_se`: standard error of the estimate of the marginal cumulative hazard function for Y
- `Fx`: estimate of the marginal distribution function for X
- `Fx_se`: standard error of the estimate of the marginal distribution function for X
- `Sy`: estimate of the marginal survival function for Y
- `Sy_se`: standard error of the estimate of the marginal survival function for Y
conv

An integer indicating why the optimization process terminated in "nlm": If conv=1, the maximization of the likelihood function is properly done. Please refer R function "nlm" for more details.

Grad

L_2 norm for the gradient vector at the solution (close to zero if the solution is proper)

MinEigen

Minimum eigenvalue of the Hessian matrix at the solution (positive if the solution is proper)

Author(s)

Takeshi EMURA

References


Examples

```r
### Generate independent truncated data with unit exponential marginals ###
set.seed(1)
m=35 ### sample size ###
x.trunc=y.trunc=numeric(m)
l=1
while(l<=m){
  x=rexp(1,rate=1)
  y=rexp(1,rate=1)
  if((x<y)){
    x.trunc[l]=x; y.trunc[l]=y
    l=l+1
  }
}

npmle.indep(x.trunc, y.trunc, x.fix =1, y.fix = 1, plotX = TRUE)

### compare the above estimate with the true value ###
c(Fx_true=1-exp(-1),Sy_true=exp(-1))
```

Description

Nonparametric maximum likelihood estimation for dependent truncation data under the normal(Gaussian) copula models (Emura & Wang, 2012). The forms of the marginal distributions for X and Y are completely unspecified, but the parametric form of copula is specified as the normal(Gaussian) copula.
Usage

NPMLE.Normal(x.trunc, y.trunc,
          x.fix = median(x.trunc), y.fix = median(y.trunc), plotX = TRUE)

Arguments

- x.trunc: vector of variables satisfying x.trunc<=y.trunc
- y.trunc: vector of variables satisfying x.trunc<=y.trunc
- x.fix: vector of fixed points at which marginal distribution function of X is calculated
- y.fix: vector of fixed points at which marginal survival function of Y are calculated
- plotX: if TRUE, the cumulative distribution function for X is plotted

Details

The function produces the nonparametric maximum likelihood estimate (NPMLE) for the marginal distributions and the estimate of the association parameter under the normal(Gaussian) copula model. The method follows Emura & Wang (2012). The maximization of the likelihood (L) is conducted by minimizing -logL by "nlm". Standard errors are calculated from the diagonal elements in the observed Fisher information matrix, which is obtained from the output of the "nlm". The method can handle ties but cannot handle right-censoring.

Value

- alpha: estimate of association parameter
- alpha_se: standard error of the estimate of association parameter
- Hx: estimate of the marginal cumulative reverse-hazard function for X
- Hx_se: standard error of the estimate of the marginal cumulative reverse-hazard function for X
- Ay: estimate of the marginal cumulative hazard function for Y
- Ay_se: standard error of the estimate of the marginal cumulative hazard function for Y
- Fx: estimate of the marginal distribution function for X
- Fx_se: standard error of the estimate of the marginal distribution function for X
- Sy: estimate of the marginal survival function for Y
- Sy_se: standard error of the estimate of the marginal survival function for Y
- conv: an integer indicating why the optimization process terminated in "nlm": If conv=1, the maximization of the likelihood function is properly done. Please refer R function "nlm" for more details.
- iteration: the number of iterations until convergence
- Grad: L_2 norm for the gradient vector at the solution (close to zero if the solution is proper)
- MinEigen: Minimum eigenvalue of the Hessian matrix at the solution (positive if the solution is proper)
Author(s)
Takeshi EMURA

References

Examples
```r
x.trunc=c(0.53, 0.43, 1.63, 0.98, 0.62)
y.trunc=c(1.20, 4.93, 2.09, 2.57, 1.52)
# NPMLE.Normal(x.trunc, y.trunc, x.fix =1, y.fix = 2, plotX = TRUE)
```

NPMLE.Plackett

Semiparametric Inference for Copula Models with Dependent Truncation Data

Description
Nonparametric maximum likelihood estimation for dependent truncation data under the Plackett copula models (Emura & Wang, 2012). The forms of the marginal distributions for X and Y are completely unspecified, but the parametric form of copula is specified as the Plackett copula.

Usage
```r
NPMLE.Plackett(x.trunc, y.trunc,
x.fix = median(x.trunc), y.fix = median(y.trunc), plotX = TRUE)
```

Arguments
- `x.trunc`: vector of variables satisfying x.trunc<=y.trunc
- `y.trunc`: vector of variables satisfying x.trunc<=y.trunc
- `x.fix`: vector of fixed points at which marginal distribution function of X is calculated
- `y.fix`: vector of fixed points at which marginal survival function of Y are calculated
- `plotX`: if TRUE, the cumulative distribution function for X is plotted

Details
The function produces the nonparametric maximum likelihood estimate (NPMLE) for the marginal distributions and the estimate of the association parameter under the Plackett copula model. The method follows Emura & Wang (2012). The maximization of the likelihood (L) is conducted by minimizing -logL by "nlm". Standard errors are calculated from the diagonal elements in the observed Fisher information matrix, which is obtained from the output of the "nlm". The method can handle ties but cannot handle right-censoring.
**Value**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>alpha</td>
<td>estimate of association parameter</td>
</tr>
<tr>
<td>alpha_se</td>
<td>standard error of the estimate of association parameter</td>
</tr>
<tr>
<td>Hx</td>
<td>estimate of the marginal cumulative reverse-hazard function for X</td>
</tr>
<tr>
<td>Hx_se</td>
<td>standard error of the estimate of the marginal cumulative reverse-hazard function for X</td>
</tr>
<tr>
<td>Ay</td>
<td>estimate of the marginal cumulative hazard function for Y</td>
</tr>
<tr>
<td>Ay_se</td>
<td>standard error of the estimate of the marginal cumulative hazard function for Y</td>
</tr>
<tr>
<td>Fx</td>
<td>estimate of the marginal distribution function for X</td>
</tr>
<tr>
<td>Fx_se</td>
<td>standard error of the estimate of the marginal distribution function for X</td>
</tr>
<tr>
<td>Sy</td>
<td>estimate of the marginal survival function for Y</td>
</tr>
<tr>
<td>Sy_se</td>
<td>standard error of the estimate of the marginal survival function for Y</td>
</tr>
<tr>
<td>conv</td>
<td>an integer indicating why the optimization process terminated in &quot;nlm&quot;: If conv=1, the maximization of the likelihood function is properly done. Please refer R function &quot;nlm&quot; for more details.</td>
</tr>
<tr>
<td>iteration</td>
<td>the number of iterations until convergence</td>
</tr>
<tr>
<td>Grad</td>
<td>L_2 norm for the gradient vector at the solution (close to zero if the solution is proper)</td>
</tr>
<tr>
<td>MinEigen</td>
<td>Minimum eigenvalue of the Hessian matrix at the solution (positive if the solution is proper)</td>
</tr>
</tbody>
</table>

**Author(s)**

Takeshi EMURA

**References**


**Examples**

```r
### Generate truncated data from Plackett copula with unit exponential marginals ###
set.seed(1)
m=25 ### sample size ###
alpha_true=1/5.11 ### Kendall's tau on (X,Y) = 0.5 ###
x.trunc=y.trunc=numeric(m)
l=1
while(l<m){
  u=runif(1,min=0,max=1)
  v=runif(1,min=0,max=1)
  K=function(w){
    A=sqrt((1+(alpha_true-1)*(u+w))^2-4*alpha_true*(alpha_true-1)*u*w )
    2*v+( 1+(alpha_true-1)*u-(alpha_true+1)*w )/A-1
  }
  low=0.00001*u
```
\begin{verbatim}
up=1-0.00001*v
for(j in 1:100){
    mid=(low+up)/2;
    if(HM>0)(up=mid)
    else(low=mid)
}

w=mid
x=-log(1-u);y=-log(w)
if(\{x<y\}){
    x.trunc[1]=x;y.trunc[1]=y
    l=l+1
}

PMLE.Plackett(x.trunc, y.trunc, x.fix =1, y.fix = 1, plotX = TRUE)
\end{verbatim}

### compare the above estimate with the true value ###
c(\alpha_{true}=\alpha_{true}, F_x_{true}=1-exp(-1), S_y_{true}=exp(-1))

---

**PMLE.Clayton.Exponential**

*Parametric Inference for Exponential Models with Dependent Truncation Data*

---

**Description**

Maximum likelihood estimation (MLE) for dependent truncation data under the Clayton copula with Exponential margins for a bivariate lifetimes (L, X). The truncated data (L_j, X_j), subject to L_j<=X_j for all j=1, ..., n, are used to obtain the MLE for the population parameters of (L, X).

**Usage**

PMLE.Clayton.Exponential(l.trunc, x.trunc, GOF = TRUE,
                         Err=3, alpha_max=20, alpha_min=10^-4)

**Arguments**

- `l.trunc` vector of truncation variables satisfying l.trunc<=x.trunc
- `x.trunc` vector of variables satisfying l.trunc<=x.trunc
- `GOF` if TRUE, a goodness-of-fit test statistics is computed
- `Err` tuning parameter in the NR algorithm
- `alpha_max` upper bound for the copula parameter
- `alpha_min` lower bound for the copula parameter

**Details**

Original paper is submitted for review
Value

- $n$: sample size
- $\alpha$: dependence parameter
- $\lambda_L$: scale parameter of $L$
- $\lambda_X$: scale parameter of $X$
- $\mu_X$: Mean lifetime of $X$, defined as $E[X]$
- $\log L$: Maximized log-likelihood
- $c$: inclusion probability, defined by $c = \Pr(L \leq X)$
- $C$: Cramer-von Mises goodness-of-fit test statistics
- $K$: Kolmogorov-Smirnov goodness-of-fit test statistics

Author(s)

Takeshi Emura, Chi-Hung Pan

References

Emura T, Pan CH (2017-). Parametric likelihood inference and goodness-of-fit for dependently left-truncated data, a copula-based approach (in revision, Statistical Papers).

Examples

```r
```

```r
x.trunc = c(38.701, 49.173, 42.409, 73.823, 46.738, 44.071, 61.904, 39.327, 49.828, 46.314, 56.150, 50.549, 54.930, 54.839, 49.170, 44.795, 72.238, 107.783, 81.609, 45.228, 124.637, 64.018, 82.957, 143.550, 43.382, 69.644, 74.750, 32.881, 51.483, 31.767, 77.633, 63.745, 82.965, 24.818, 68.762, 68.762, 89.100, 64.979, 65.127, 59.289, 53.926, 79.370, 47.385, 61.395, 72.826, 53.980, 37.220, 44.224, 50.826, 65.460, 86.726, 43.819, 100.605, 67.615, 89.542, 68.266, 103.580, 82.570, 87.960, 42.385, 68.914, 95.666, 78.135, 83.643, 18.617, 92.629, 42.415, 34.346, 186.569, 20.758, 52.003, 77.179, 68.934, 78.661, 165.543, 79.547, 55.009, 46.774, 124.526, 92.504, 109.986, 101.161, 59.422, 27.772, 33.598, 69.038, 75.222, 58.373, 185.610, 56.158, 55.913, 83.770, 123.468, 68.994, 101.869, 87.627, 38.798, 74.734)
```

```r
u.min=10
```
**PMLE.Clayton.Weibull**

Description

Maximum likelihood estimation (MLE) for dependent truncation data under the Clayton copula with Weibull margins for a bivariate lifetimes ($L, X$). The truncated data ($L_j, X_j$), subject to $L_j \leq X_j$ for all $j=1, ..., n$, are used to obtain the MLE for the population parameters of $(L, X)$.

Usage

```r
PMLE.Clayton.Weibull(l.trunc, x.trunc, GOF = TRUE,
                         Err = 2, alpha_max = 20, alpha_min = 1e-6)
```

Arguments

- `l.trunc`: vector of truncation variables satisfying $l.trunc \leq x.trunc$
- `x.trunc`: vector of variables satisfying $l.trunc \leq x.trunc$
- `GOF`: if TRUE, a goodness-of-fit test statistics is computed
- `Err`: tuning parameter in the NR algorithm
- `alpha_max`: upper bound for the copula parameter
- `alpha_min`: lower bound for the copula parameter

Details

Relevant paper is submitted for review

Value

- `n`: sample size
- `alpha`: dependence parameter
- `lambda_L`: scale parameter of $L$
- `lambda_X`: scale parameter of $X$
- `nu_L`: shape parameter of $L$
- `nu_X`: shape parameter of $X$
- `mean_X`: Mean lifetime of $X$, defined as $E[X]$
- `logL`: Maximized log-likelihood
- `c`: inclusion probability, defined by $c = \Pr(L \leq X)$
- `C`: Cramer-von Mises goodness-of-fit test statistics
- `K`: Kolmogorov-Smirnov goodness-of-fit test statistics
PMLE.Normal

Description

Maximum likelihood estimation (MLE) for dependent truncation data under the bivariate normal distribution. A bivariate normal distribution is assumed for bivariate random variables (L, X). The truncated data (L_j, X_j), subject to L_j<=X_j for all j=1, ..., n, are used to obtain the MLE for the population parameters of (L, X).
Usage

PMLE.Normal(l.trunc, x.trunc, testimator = FALSE, GOF=TRUE)

Arguments

- `l.trunc` vector of truncation variables satisfying \( l.trunc \leq x.trunc \)
- `x.trunc` vector of variables satisfying \( l.trunc \leq x.trunc \)
- `testimator` if TRUE, testimator is computed instead of MLE
- `GOF` if TRUE, goodness-of-fit test is performed

Details

PMLE.Normal performs the maximum likelihood estimation for dependently left-truncated data under the bivariate normal distribution. "PMLE.Normal" implements the methodologies developed in Emura T. & Konno Y. (2012, Statistical Papers 53, 133-149) and can produce the maximum likelihood estimates and their standard errors. Furthermore, "PMLE.Normal" tests the independence assumption between truncation variable and variable of interest via likelihood ratio test. The MLE is obtained by minimizing \(-\log L\) using "nlm", where \( L \) is the log-likelihood.

Value

- `mu_L` mean of \( L \) and its standard error
- `mu_X` mean of \( X \) and its standard error
- `var_L` variance of \( L \) and its standard error
- `var_X` variance of \( X \) and its standard error
- `cov_LX` covariance between \( L \) and \( X \) and its standard error
- `c` inclusion probability, defined by \( c = \Pr(L \leq X) \), and its standard error
- `test` Likelihood ratio statistic and p-value
- `C` Cramer-von Mises goodness-of-fit test statistics
- `K` Kolmogorov-Smirnov goodness-of-fit test statistics

Author(s)

Takeshi EMURA

References


Examples

```
l.trunc=c(1,2,3,4,5,6,7,8,8)
x.trunc=c(2,4,4,5,5,7,9,10)
PMLE.Normal(l.trunc, x.trunc, testimator=FALSE)
```
Index

*Topic **Copula**
  CHAIEB. Clayton, 4
  CHAIEB. Frank, 6
  EMURA. Clayton, 10
  EMURA. Frank, 12
  NPMLE. Frank, 14
  NPMLE. Indep, 16
  NPMLE. Normal, 17
  NPMLE. Plackett, 19

*Topic **Parametric method**
  PMLE. Clayton. Exponential, 21
  PMLE. Clayton. Weibull, 23
  PMLE. Normal, 24

*Topic **Regression**
  depend.AFT.reg, 8

*Topic **Semiparametric method**
  CHAIEB. Clayton, 4
  CHAIEB. Frank, 6
  depend.AFT.reg, 8
  EMURA. Clayton, 10
  EMURA. Frank, 12
  NPMLE. Frank, 14
  NPMLE. Indep, 16
  NPMLE. Normal, 17
  NPMLE. Plackett, 19

*Topic **datasets**
  centenarian, 3
  centenarian, 3
  CHAIEB. Clayton, 4
  CHAIEB. Frank, 6
  depend.truncation
    (depend.truncation-package), 2
  depend.truncation-package, 2
  depend.AFT.reg, 8
  EMURA. Clayton, 10
  EMURA. Frank, 12
  NPMLE. Frank, 14