

Package: emplikCS (via r-universe)

June 3, 2026

Version 0.3

Title Empirical Likelihood with Current Status Data for Mean,
Probability, Hazard

Maintainer Mai Zhou <maizhou@gmail.com>

Depends R (>= 4.0.0), quadprog, monotone

Imports stats

Description Compute the empirical likelihood ratio, -2LogLikRatio
(Wilks) statistics, based on current status data for the
hypotheses about the parameters of mean or probability or
weighted cumulative hazard.

License GPL (>= 2)

NeedsCompilation no

Author Mai Zhou [aut, cre]

Repository <https://maizhou.r-universe.dev>

Date/Publication 2026-03-05 04:57:57 UTC

RemoteUrl <https://github.com/cran/emplikCS>

RemoteRef HEAD

RemoteSha 64129fdf4f3f7773fc9ceeb3e59399be455bc66e

Contents

CSbj	2
CSdataclean	3
el.CS.2prob	4
el.CS.Hz	6
el.CS.mean	8
el.CS.prob	9
isotNEW2	11
L2S	12
S2L	13

Index	15
--------------	-----------

CSbj *Current Status Data Buckley-James Estimator For Linear Regression Models*

Description

In an AFT regression model, when the responses are current status censored (observe y_i either $> t_i$ or $\leq t_i$), we may still estimate the regression coefficients by the Buckley-James (extension from right censored case). We assume the inspection time t_i have a larger support to cover the support of error ϵ_i , which is assumed iid.

Usage

```
CSbj(x, delta, Itime, maxiter = 99, error = 0.0001)
```

Arguments

<code>x</code>	Design matrix. N rows p columns.
<code>delta</code>	Either 0 or 1. I[$y_i \leq t_i$]. Length N. $y_i = \beta x_i + \epsilon_i$
<code>Itime</code>	The inspection times. Length N.
<code>maxiter</code>	Default to 99. Control the iteration.
<code>error</code>	Default to 0.0001. Control the iteration.

Details

This function is an implementation of the Buckley-James estimator for the regression parameter β in the AFT regression model when the responses, y_i , are current status censored. Similar to the Binary Choice model in econometrics where all the inspection times are fixed at zero. I wrote an S-plus function for the binary choice model (name `bibj`). It is easily adapted to the current status situation, and this is the function. The AFT model we considered here has an intercept term. But we try to estimate the regression parameter β , without the intercept term first. The estimator of the intercept can be obtained (if needed) as the mean of the iid errors/residuals after we got the estimator of the slope parameter β .

This function depends on the functions `monotone` from package `monotone` and `lsfit` from the basic stats package.

At this point, we do not have a good estimate for the variance for the Buckley-James estimators. Bootstrap is one method one can try.

Value

It returns a list containing

<code>est</code>	The Buckley-James estimator of the regression coefficients.
<code>iterN</code>	Number of iterations done.
<code>distFx</code>	Locations of the jumps of final estimator of the error distribution.

distFy Probabilities of the final estimator of the error distribution at jump locations. Mean of this error distribution is the intercept term estimator of the regression model.

Author(s)

Mai Zhou <maizhou@gmail.com>.

References

Zhou, M. (2026). *Empirical Likelihood Method in Survival Analysis 2nd Edition* Chapman & Hall/CRC

Wang, W., and Zhou, M. (1995). *Iterative least squares estimator of binary choice models: a semi-parametric approach*. Tech. Report, University of Kentucky. <https://www.ms.uky.edu/~mai/research/eco5wz.pdf>

Buckley, J. J., and James, I. R. (1979). *Linear regression with censored data*. *Biometrika* **66**, 429–436.

Examples

```
y <- c(10, 209, 273, 279, 324, 391, 566, 785)
x <- c(21, 38, 39, 51, 77, 185, 240, 289, 524)
```

CSdataclean	<i>Order the Given Current Status Data and Output with Some Additional Information</i>
-------------	----------------------------------------------------------------------------------------

Description

Given N current status data, order according to inspection times, and find the positions (in the ordered data) of the first 1 and last 0 of the delta.

Usage

```
CSdataclean(itime, delta)
```

Arguments

itime	The inspection times. Length N.
delta	Either 0 or 1. I[yi <= itimei]. Length N.

Details

When calculate NPMLE of the CDF $F(t)$ from current status data, it is obvious that $F(t) = 0$ before the location (in `itime`) of first $\delta=1$ occurrence in the delta list, and similarly, $F(t) = 1$ after the last $\delta=0$ occurrence. So in the calculation of NPMLE we need only to consider estimating $F(\cdot)$ when time t is from `itime[first]` to `itime[last]`.

We take the definition of NPMLE $\hat{F}_n(t)$ as right continuous.

Usually, the current status data are stored in either long format or short format. The short format is often used when there are many tied inspection times. This function, `CSdataclean`, takes input of current status data in the long format: `itime=(t1, t2, ... tN)` and `delta=(0, 1, ... 1)`. The only values in the delta are 0 or 1. If needed you can convert the short format to long format using function `L2S`.

Value

It returns a list containing

<code>itime</code>	The ordered inspection times.
<code>delta</code>	The delta, ordered according to <code>itime</code> .
<code>Istart</code>	<code>delta[Istart]</code> is the first 1 in the ordered delta output.
<code>Iend</code>	<code>delta[Iend]</code> is the last 0 in the ordered delta output.

Author(s)

Mai Zhou <maizhou@gmail.com>.

References

Zhou, M. (2026). *Empirical Likelihood Method in Survival Analysis 2nd Edition* Chapman & Hall/CRC

Examples

```
y <- c(10, 209, 273, 279, 324, 391, 566, 785)
x <- c(21, 38, 39, 51, 77, 185, 240, 289, 524)
```

el.CS.2prob

*Current Status Data Empirical Likelihood Test for Two Parameters:
F(t_1) and F(t_2) Jointly*

Description

Given n current status data, we may estimate the CDF $F(t)$ by NPMLE (e.g. by `isotNEW2()` function in this package). This function, `el.CS.2prob`, uses empirical likelihood to test the hypothesis that $F(t)$ at two given locations (`t01` and `t02`) equal to two given values (`Ft01` and `Ft02`): i.e. $H_0: F(t01) = Ft01$ and $F(t02) = Ft02$ jointly.

Empirical likelihood ratio test returns the Wilks statistics, $-2LLR$. By our working conjecture, the $-2 \log$ likelihood ratio times $(5/3)$ under H_0 is approximately chi square $DF=2$ distributed. See reference below.

Usage

```
el.CS.2prob(ti, di, t01, Ft01, t02, Ft02)
```

Arguments

ti	The inspection times, a vector of length n.
di	Either 0 or 1. I[$y_i \leq t_i$]. length n.
t01	The given time where F() value is tested.
Ft01	The hypothesized value of F(t01). Must be within (0, 1).
t02	The given time where F() value is tested.
Ft02	The hypothesized value of F(t02). Must be within (0, 1).

Details

This function tests the null hypothesis that $F(t01) = Ft01$ and $F(t02) = Ft02$ versus at least one not equal. We assume the data given is current status censored data.

We require t01 (and also t02) be equal to one of the inspection times. If not, you have to do something using the right continuity of the NPMLE (change t01 to the closest ti on the left, etc.).

The NPMLE $F(t)$ is convergent at cubic root n speed and the -2LLR times (5/3) has chi square DF=2 null distribution, according to the working conjecture.

It goes without saying that we assume the NPMLE has finite asymptotic variance (when normalized by cubic root n).

Value

It returns a list containing

"-2LLR"	The Wilks statistics of the EL test, after multiply by (5/3) has approximate chi SQ DF=2 distribution under null hypothesis.
LogLik0	The log lik value achieved by the un-constrained NPMLE.
LogLik1	The log lik value achieved by the constrained NPMLE.

Author(s)

Mai Zhou <maizhou@gmail.com>.

References

- Banerjee, M., and Wellner, J. (2001). Likelihood ratio tests for monotone functions. *Annals of Statistics* 29, 1699-1731.
- Banerjee, M., and Wellner, J. (2005). Score statistics for current status data: Comparisons with likelihood ratio and Wald statistics. *The International Journal of Biostatistics* 1, 1-29.
- Zhou, M. (2026). *Empirical Likelihood Method in Survival Analysis 2nd Edition* Chapman & Hall/CRC.
- Sun, J. (2006). *The Statistical Analysis of Interval-Censored Failure Time Data* Springer, New York.

Examples

```

N <- 300
set.seed(12345)
itime <- sort(c(rexp(N-2), 0.3, 0.6) )      ##### inspection times
Stime <- rexp(N)                          ##### survival times
delta <- as.numeric(Stime <= itime)        ##### current status censoring

```

el.CS.Hz

Current Status Data Empirical Likelihood Test for a Parameter Defined by Weighted Cumulative Hazard

Description

Given n current status data, we may estimate the CDF $F(t)$ by NPMLE (e.g. by isotNEW2 in this package). By the 1 to 1 correspondence, $\Lambda(t) = -\log[1 - F(t)]$, we can also calculate the NPMLE of the cumulative hazard function $\Lambda(t)$. We are interested in a parameter defined by weighted average of cumulative hazard function (or integrated cumulative hazard) defined below. Empirical likelihood ratio test is performed and return a Wilks statistics. The Wilks statistics, $-2 \log$ likelihood ratio, under H_0 is approximately chi square $DF=1$ distributed.

Usage

```
el.CS.Hz(ti, di, Pfun, thetaMU, error=1e-11, maxit=25)
```

Arguments

ti	The inspection times, a vector of length n .
di	Either 0 or 1. $I[y_i \leq t_i]$. length n .
Pfun	The function used in calculate the weighted cumulative hazard. See an example below.
thetaMU	The null hypothesis value of the weighted cumulative hazard to be tested.
error	Default to $1e-11$. Control the SQP iteration.
maxit	Default to 25. Control the SQP iteration.

Details

This function tests the null hypothesis that the parameter of weighted (or integrated) cumulative hazard function equal to the given value (thetaMU). We assume the data given is current status censored data.

The (unknown) true value of the parameter of interest is defined by

$$\theta(\Lambda_0) = \int_0^M \Lambda_0(s) d\Psi(s)$$

where $\Lambda_0(s)$ is the true cumulative hazard function. And $\Psi(s)$ is the weight function given(Pfun, used to generate weights). The function $\Psi(s)$ is assumed to be continuous and at least piecewise differentiable.

The NPMLE of cumulative hazard, $\hat{\Lambda}_n(t)$, is obtained via the NPMLE of CDF $\hat{F}_n(t)$:

$$\hat{\Lambda}_n(t) = -\log[1 - \hat{F}_n(t)].$$

It goes without saying that we assume the parameter is well defined (not equal to infinity) and its NPMLE has finite asymptotic variance.

Value

It returns a list containing

"-2LLR"	The Wilks statistics of the EL test, has approximate chi SQ DF=1 distribution under null hypothesis.
location	The locations of jumps of the NPMLE of cumulative hazard function (also of CDF).
Haz	The constrained NPMLE of cumulative hazard function.
iter	Number of iterations done.
error	The error at final iteration, = sum(abs(difference))
Loglik1	The log lik value achieved by the constrained NPMLE.
Check	The weighted cumulative hazard by the constrained NPMLE. This should equal to thetaMU (of input). A check of convergence.
thetaMLE	The NPMLE of the (non-constrained) weighted cumulative hazard. The input value thetaMU should not deviate from this value by too much.

Author(s)

Mai Zhou <maizhou@gmail.com>.

References

Zhou, M. (2026). *Empirical Likelihood Method in Survival Analysis 2nd Edition* Chapman & Hall/CRC

Examples

```
##### An example of the Pfun (or Psi(s) or weight function) and calculation #####
mydgfun2 <- function(t){0.3*t*(10-t)*as.numeric(t<=10)}

N <- 3000
set.seed(12345)
itime <- sort(rexp(N, rate=0.1))      ##### inspection times
Stime <- rexp(N, rate=0.1)           ##### survival times
delta <- as.numeric(Stime <= itime)  ##### current status censoring

el.CS.Hz(ti=itime, di=delta, Pfun=mydgfun2, thetaMU= -5) ## -5 is the true value of parameter.

##### You should get
## $^-2LLR^
## [1] 1.04782          ##### and more.
```

el.CS.mean	<i>Current Status Data Empirical Likelihood Test for the Parameter of Mean: mu(F)</i>
------------	---------------------------------------------------------------------------------------

Description

Given n current status data, we may estimate the CDF $F(t)$ by NPMLE (e.g. by the `isotNEW2()` function in this package). Based on the NPMLE $\hat{F}_n(t)$ we can estimate the mean $\mu(F)$.

This function, `el.CS.mean`, uses empirical likelihood to test the hypothesis that $\mu(F)$ equal to a given value(μ): i.e. $H_0: \mu(F) = \mu$.

Empirical likelihood ratio test returns the Wilks statistics, $-2LLR$. The -2 log likelihood ratio under H_0 is asymptotically chi square $DF=1$ distributed. See reference below.

Usage

```
el.CS.mean(mu, Itime, delta, Pfun)
```

Arguments

<code>mu</code>	The hypothesized mean value.
<code>Itime</code>	The inspection times, a vector of length n .
<code>delta</code>	Either 0 or 1. $I[y_i \leq t_i]$. length n .
<code>Pfun</code>	A given function, $\Psi(s)$, used to define the (weighted) mean.

Details

This function tests the null hypothesis that $\mu(F) = \mu$ versus not equal. We assume the data given is current status censored data.

The definition of the mean, $\mu(F)$ is

$$\mu(F) = \int_0^M [1 - F(t)] d\Psi(t)$$

and its estimator based on (δ_i, t_i) or \hat{F}_n is (assume $\min(t_i) = 0$ or $t_{(1)} = 0$)

$$\mu(\hat{F}_n) = \sum_{i=1}^n [1 - \hat{F}_n(t_{(i)})] \Delta\Psi(t_{(i)}),$$

where $\Psi(t)$ is a given function and $\Delta\Psi(t_{(i)}) = \Psi(t_{(i+1)}) - \Psi(t_{(i)})$. If $\Psi(t) = t$ in the above, then this is the ordinary mean (assuming $F(t)$ has support $(0, M)$).

The NPMLE $\hat{F}_n(t)$ is convergent at cubic root n speed, but the mean estimator is convergent at ordinary root n speed. The $-2LLR$ has chi square $DF=1$ null distribution asymptotically.

It goes without saying that we assume the NPMLE $\mu(\hat{F})$ has finite asymptotic variance (when normalized by root n).

Value

It returns a list containing

"-2LLR" The Wilks statistics of the EL test, has approximate chi SQ DF=1 distribution under null hypothesis.

Author(s)

Mai Zhou <maizhou@gmail.com>.

References

Zhou, M. (2026). *Empirical Likelihood Method in Survival Analysis 2nd Edition* Chapman & Hall/CRC

Huang, J. and Wellner, J. (1995). *Asymptotic normality of the NPMLE of linear functionals for interval censored data, case 1* Statistica Neerlandica **49**, 2, 153–163.

Sun, J. (2006). *The Statistical Analysis of Interval-Censored Failure Time Data* Springer, New York.

Examples

```
N <- 300
set.seed(12345)
itime <- sort(c(rexp(N-1), 0.5) )        ##### inspection times
Stime <- rexp(N)                        ##### survival times
delta <- as.numeric(Stime <= itime)    ##### current status censoring
```

e1.CS.prob	<i>Current Status Data Empirical Likelihood Test for the Probability $F(t_0)$.</i>
------------	-----------------------------------------------------------------------------------------------

Description

Given n current status data, we may estimate the CDF $F(t)$ by NPMLE (e.g. by `isotNEW2()` function in this package). This function, `e1.CS.prob`, uses empirical likelihood to test the hypothesis that $F(t)$ at a given location (t_0) equal to a given value (Ft_0): i.e. $H_0: F(t_0) = Ft_0$.

Empirical likelihood ratio test returns the Wilks statistics, `-2LLR`. According to our working conjecture, the -2 log likelihood ratio times $(5/3)$ under H_0 is approximately chi square DF=1 distributed. See references below.

Usage

```
e1.CS.prob(ti, di, t0=0.5, Ft0=0.5)
```

Arguments

t_i	The inspection times, a vector of length n .
d_i	Either 0 or 1. $I[y_i \leq t_i]$. length n .
t_0	The given time where $F()$ value is tested.
F_{t_0}	The hypothesized value of $F(t_0)$. Must be within $(0, 1)$.

Details

This function tests the null hypothesis that $F(t_0) = F_{t_0}$ versus not equal. We assume the data given is current status censored data.

We require t_0 equals to one of the inspection times. If not, you have to do something by the right continuity of the NPMLE (change t_0 to the closest t_i on the left).

The NPMLE $F(t)$ is convergent at cubic root speed and the $-2LLR$ times $(5/3)$ has chi square $DF=1$ null distribution, according to our working conjecture.

It goes without saying that we assume the NPMLE has finite asymptotic variance (when normalized by cubic root n).

Value

It returns a list containing

"-2LLR"	The Wilks statistics of the EL test, when multiply by $(5/3)$ has approximate chi SQ $DF=1$ distribution under null hypothesis.
LogLik0	The log lik value achieved by the un-constrained NPMLE.
LogLik1	The log lik value achieved by the constrained NPMLE.

Author(s)

Mai Zhou <maizhou@gmail.com>.

References

- Banerjee, M., and Wellner, J. (2001). Likelihood ratio tests for monotone functions. *Annals of Statistics* 29, 1699-1731.
- Banerjee, M., and Wellner, J. (2005). Score statistics for current status data: Comparisons with likelihood ratio and Wald statistics. *The International Journal of Biostatistics* 1, 1-29.
- Zhou, M. (2026). *Empirical Likelihood Method in Survival Analysis 2nd Edition* Chapman & Hall/CRC
- Sun, J. (2006). *The Statistical Analysis of Interval-Censored Failure Time Data* Springer, New York.

Examples

```

N <- 300
set.seed(12345)
itime <- sort(c(rexp(N-1), 0.5) )      ##### inspection times
Stime <- rexp(N)                      ##### survival times
delta <- as.numeric(Stime <= itime)    ##### current status censoring

el.CS.prob( ti=itime, di=delta, t0=0.5, Ft0=pexp(0.5) )

##### You should get
## $^-2LLR^
## [1] 1.867655      ##### and more.

```

isotNEW2

Given Current Status Data, Calculate the NPMLE of CDF $F(t)$ by Calling the monotone Function from monotone Package

Description

Using improved PAVA algorithm to calculate the NPMLE of CDF $F(t)$. Input inspection times (x) can have ties. We require two extra inputs a and b to make sure the output is a proper CDF: $F(t_{(1)} - a) = 0$, and $F(t_{(n)} + b) = 1$ always, and the rest are increase from 0 to 1.

Usage

```
isotNEW2(x, y, a, b, LONG=TRUE)
```

Arguments

x	Inspection time of the current status data.
y	Equivalent to delta in the current status data. Either 0 or 1 as in $I[\text{survTi} \leq x_i]$. length N .
a	To make sure the output $F(\cdot)$ is a proper CDF: $F(x[1]-a) = 0$ always.
b	To make sure the output $F(\cdot)$ is a proper CDF: $F(x[n]+b) = 1$ always.
LONG	Should the output in the LONG format or not? Default is TRUE.

Details

Usually, the current status data are stored in either long format or short format. The short format is often used when there are many tied inspection times. This function, `isotNEW2`, takes in the current status data in the long form: $x=(t_1, t_2, \dots, t_N)$ and $y=(0, 1, \dots, 1)$. The only values of the y are 0 or 1.

For more details please refer to monotone package.

May be we should put $a=1, b=1$ as default?

Since the NPMLE of $F(t)$ has very few number of jumps (for sample size N , the number of positive jumps are about cubic root N), a short format output can same space. For current status data of sample size 1000, usually the NPMLE $F(t)$ has about 10 jumps. So the short format has length about 10 and the long format has length 1000.

Value

It returns a list in either the long format or short format containing

x The ordered inspection times, including $x[1]-a$, and $x[n]+b$.
 y The NPML of $F(x)$ at the x time, $F(x[1]-a)=0$ always; $F(x[n]+b)=1$ always.

Author(s)

Mai Zhou <maizhou@gmail.com>.

References

Busing, F. (2022). *Monotone regression: A simple and fast $O(n)$ PAVA implementation*. Journal of Statistical Software, Code Snippets 102, 1, 1–25. doi: [10.18637/jss.v102.c01](https://doi.org/10.18637/jss.v102.c01)

Zhou, M. (2026). *Empirical Likelihood Method in Survival Analysis 2nd Edition* Chapman & Hall/CRC

Examples

```
itime <- c(1, 2, 3, 4, 5, 6, 7, 8)
delta <- c(0, 1, 0, 1, 1, 1, 0, 1)

isotNEW2(x=itime, y=delta, a=0.5, b=0.5)
## $x
## [1] 0.5 1.0 2.0 3.0 4.0 5.0 6.0 7.0 8.0 8.5
##
## $y
## [1] 0.00 0.00 0.50 0.50 0.75 0.75 0.75 0.75 1.00 1.00
#### the correct answer is F(t) = 0, .5, .5, .75, .75, .75, .75, 1
#### at the ordered itime, augmented by a and b.
isotNEW2(x=itime, y=delta, a=0.5, b=0.5, LONG=FALSE)
## $x
## [1] 2 4 8
##
## $y
## [1] 0.50 0.75 1.00 #### for time t < 2, F(t) = 0 by right cont.
```

L2S

Convert the Current Status Data from Long Format to Short Format

Description

Given current status data (two vectors of length n): one vector of inspection times, the other vector = $I[Y_i \leq t_i]$. We assume the input is ordered according to $itime$. Convert the same data into short format. Useful when there are a lot of tied inspection times. (saves space).

Usage

```
L2S(itime, delta)
```

Arguments

`itime` The inspection times. Length `n`.
`delta` Either 0 or 1. $I[y_i \leq \text{itime}_i]$. Length `n`.

Details

Same data set of 850 cases Hepatitis A tests from Bulgaria recorded in long format `hepABulg` in the R package `csci` compare to the short format `hepatitisA` in the R package `curstatCI`.

This function (and the sister function `S2L`) can be used to convert between the short and long format.

Value

It returns a list containing

`Sitime` The ordered inspection times. Short format. Tied itimes removed.
`fi` Number of delta's=1, with identical itime `ti`.
`ni` Number of itimes tied at `ti`.

Author(s)

Mai Zhou <maizhou@gmail.com>.

References

Zhou, M. (2026). *Empirical Likelihood Method in Survival Analysis 2nd Edition* Chapman & Hall/CRC

Examples

```
y <- c(10, 209, 273, 279, 324, 391, 566, 785)
x <- c(21, 38, 39, 51, 77, 185, 240, 289, 524)
```

S2L

Convert the Current Status Data from Short Format to Long Format

Description

Given current status data in short format (three vectors), convert it to the long format (one vector of inspection times, another vector= $I[Y_i \leq t_i]$). We assume the input is ordered according to inspection times `itime`.

Usage

```
S2L(itime, fi, ni, OneFirst=TRUE)
```

Arguments

<code>itime</code>	The unique inspection times. Tie removed. Length k .
<code>fi</code>	Frequency of 1's in the delta. Definition is $I[y_i \leq itime_i]$. Length k .
<code>ni</code>	Frequency of inspection times t_i tied at this value. Length k .
<code>OneFirst</code>	In the output of long format data, within the same inspection times, $\delta=1$'s comes before 0's. Default to TRUE.

Details

Same data set of 850 cases Hepatitis A tests from Bulgaria recorded in long format `hepABulg` in the R package `csci` compare to the short format `hepatitisA` in the R package `curstatCI`.

This function (and the sister function `S2L`) can be used to convert between the short and long format.

Value

It returns a list containing

<code>litime</code>	The ordered inspection times. Long format.
<code>delta</code>	This is $I[Y_i \leq t_i]$.

Author(s)

Mai Zhou <maizhou@gmail.com>.

References

Zhou, M. (2026). *Empirical Likelihood Method in Survival Analysis 2nd Edition* Chapman & Hall/CRC

Examples

```
y <- c(10, 209, 273, 279, 324, 391, 566, 785)
x <- c(21, 38, 39, 51, 77, 185, 240, 289, 524)
```

Index

* nonparametric

CSbj, [2](#)

CSdataclean, [3](#)

e1.CS.2prob, [4](#)

e1.CS.Hz, [6](#)

e1.CS.mean, [8](#)

e1.CS.prob, [9](#)

isotNEW2, [11](#)

L2S, [12](#)

S2L, [13](#)

CSbj, [2](#)

CSdataclean, [3](#)

e1.CS.2prob, [4](#)

e1.CS.Hz, [6](#)

e1.CS.mean, [8](#)

e1.CS.prob, [9](#)

isotNEW2, [11](#)

L2S, [12](#)

S2L, [13](#)