

# Package ‘EL2Surv’

June 28, 2017

**Title** Empirical Likelihood (EL) for Comparing Two Survival Functions

**Version** 1.0

**Description**

Functions for computing critical values and implementing the one-sided/two-sided EL tests.

**Depends** R (>= 2.13.0)

**Imports** survival, stats

**License** GPL (>= 2)

**LazyData** true

**Author** Hsin-wen Chang [aut, cre] <hwchang@stat.sinica.edu.tw>

**Maintainer** Pei-Yuan Tsai <pyt@stat.sinica.edu.tw>

**Archs** i386, x64

**RoxygenNote** 6.0.1

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2017-06-28 06:21:46 UTC

## R topics documented:

hepatitis . . . . .	2
intELtest . . . . .	3
ptwiseELtest . . . . .	6
supELtest . . . . .	7
<b>Index</b>	<b>9</b>

---

hepatitis

*Survival from Severe Alcoholic Hepatitis*

---

### Description

The data frame `hepatitis` is obtained by digitizing the published Kaplan-Meier curves in Nguyen-Khac et al (2011). The method of digitizing is described in Guyot et al. (2012). See [intELtest](#) and [supELtest](#) for the application.

### Usage

`hepatitis`

### Format

The `hepatitis` is a data frame with 174 observations of 3 variables, and has the following columns:

- `time` the survival time
- `censor` the censoring indicator
- `group` the grouping variable

### Source

Nguyen-Khac et al., "Glucocorticoids plus N-Acetylcysteine in Severe Alcoholic Hepatitis," *The New England Journal of Medicine*, Vol. 365, No. 19, pp. 1781-1789 (2011). <http://www.nejm.org/doi/full/10.1056/NEJMoa1101214#t=article>

### References

P. Guyot, A. E. Ades, M. J. N. M. Ouwens, and N. J. Welton, "Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves," *BMC Medical Research Methodology*, 12(1):9. <http://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-12-9>

### See Also

[intELtest](#), [supELtest](#)

intELtest

*The integrated likelihood ratio test***Description**

intELtest gives a class of the weighted likelihood ratio statistics:

$$\sum_{t \in U} w(t) \{-2 \log R(t)\},$$

where  $w(t)$  is an objective weight function, and  $R(t)$  is an empirical likelihood (EL) ratio that compares two survival functions at each time point  $t$  in the set of observed uncensored lifetimes,  $U$ .

**Usage**

```
intELtest(data, g1 = 1, t1 = 0, t2 = Inf, sided = 2, nboot = 1000,
  wt = "p.event", alpha = 0.05, compo = FALSE, seed = 1011,
  nlimit = 200)
```

**Arguments**

data	a data frame/matrix with 3 columns. The first column is the survival time. The second is the censoring indicator. The last is the grouping variable. An example as the input to data provided is <a href="#">hepatitis</a> .
g1	the group with the longer survival that should take a value from the third column of data
t1	pre-specified $t_1$ based on domain knowledge with the default value of 0
t2	pre-specified $t_2$ based on domain knowledge with the default value of $\infty$
sided	2 if two-sided test, and 1 if one-sided test. It assumes the default value of 2.
nboot	number of bootstrap replications in calculating critical values
wt	a string for the integral statistic with a specific weight function. There are four types of integral statistics provided: "p.event", "dt", "db", and "dF". See 'Details' for more about the integral statistics.
alpha	pre-specified significance level of the test
compo	A logical value specifying whether to change the local statistics at time points outside the overlapping region of the two samples. The default is FALSE. TRUE if uses a functional of the difference between the Kaplan–Meier estimators (see "Details" for more).
seed	the parameter to <a href="#">set.seed</a> for the random number generator in R. The <code>set.seed</code> is used implicitly in intELtest.
nlimit	the splitting unit. To deal with large data problems, the bootstrap algorithm is to split the number of bootstrap replicates into <code>nsplit</code> parts. The number <code>nsplit</code> is the smallest integer not less than $\ U\  / \text{nlimit}$ .

## Details

intELtest calculates the weighted likelihood ratio statistics:

$$\sum_{i=1}^h w_i \cdot \{-2 \log R(t_i)\},$$

where  $w_1, \dots, w_h$  are the values of the weight function evaluated at the distinct ordered uncensored times  $t_1, \dots, t_h$  in  $U$ . There are four types of weight functions considered.

- (wt = "dt")

By means of an extension of the integral statistic derived by Pepe and Fleming (1989),

$$w_i = \begin{cases} t_{i+1} - t_i & \text{if } i \neq h \\ 0 & \text{if } i = h \end{cases}$$

- (wt = "p.event")

According to the integral statistic derived by Uno et al. (2013):

$$w_i = \frac{1}{n_1 + n_2},$$

where  $n_1$  and  $n_2$  are the sample sizes of each group. The role of  $w_i$  assigns equal weight  $1/n$  to each observation when no tie is involved ( $n = n_1 + n_2$ ); otherwise, it assigns heavy weight to the observations with multiplicity.

- (wt = "dF")

Based on the integral statistic built by Barimi and McKeague (2013), the weight function is the derivative of the empirical distribution function  $\hat{F}(t)$ .

$$w_i = \left. \frac{d\hat{F}(t)}{dt} \right|_{t=t_i}$$

This is an empirical version of taking expectation.

- (wt = "db")

The weight function is of the form:

$$w_i = \left. \frac{d\hat{b}(t)}{dt} \right|_{t=t_i}$$

with  $\hat{b}(t) = \hat{\sigma}^2(t)/(1 + \hat{\sigma}^2(t))$ . The  $\hat{b}(t)$  is chosen so that the limiting distribution

$$\int_{x_1}^{x_2} \frac{B_+^2(x)}{x(1-x)} dx$$

is the same as the asymptotic null distribution in Barimi and McKeague (2013) with the  $[x_1, x_2]$  restriction.

For time points outside the overlapping region of the two samples, we also consider an alternative local statistic by taking a functional of the difference between the KM estimators inspired by Pepe and Fleming (1989). Here we take a standardized square of the difference for two-sided testing and the positive part of that for one-sided testing, inspired from the self-studentization property of EL.

**Value**

intELtest returns a list with three elements:

- critval the critical value
- teststat the resulting integrated test statistic
- pvalue the p-value based on the integrated statistic

**References**

- H. W. Chang, "Empirical likelihood tests for stochastic ordering based on censored and biased data," *Columbia University Academic Commons* (2014). <http://academiccommons.columbia.edu/catalog/ac%3A177230>
- M. S. Pepe and T. R. Fleming, "Weighted Kaplan-Meier Statistics: A Class of Distance Tests for Censored Survival Data," *Biometrics*, Vol. 45, No. 2, pp. 497-507 (1989). [https://www.jstor.org/stable/2531492?seq=1#page\\_scan\\_tab\\_contents](https://www.jstor.org/stable/2531492?seq=1#page_scan_tab_contents)
- H. Uno, L. Tian, B. Claggett, and L. J. Wei, "A versatile test for equality of two survival functions based on weighted differences of Kaplan-Meier curves," *Statistics in Medicine*, Vol. 34, No. 28, pp. 3680-3695 (2015). <http://onlinelibrary.wiley.com/doi/10.1002/sim.6591/abstract>
- H. E. Barimi and I. W. McKeague, "Empirical likelihood-based tests for stochastic ordering," *Bernoulli*, Vol. 19, No. 1, pp. 295-307 (2013). <https://projecteuclid.org/euclid.bj/1358531751>

**See Also**

[hepatitis](#), [supELtest](#), [ptwiseELtest](#)

**Examples**

```
library(EL2Surv)
intELtest(hepatitis, 1, sided = 2, wt = "p.event")

## OUTPUT:
## $critval
## [1] 0.8993514
##
## $teststat
## [1] 1.406029
##
## $pvalue
## [1] 0.012
```

ptwiseELtest

*The pointwise likelihood ratio test***Description**

ptwiseELtest gives pointwise EL statistic values at uncensored time span. The pointwise statistic considers only the decision on each single time point; thus, it is different from the [integral type](#) and [sup type](#) statistics.

**Usage**

```
ptwiseELtest(data, g1 = 1, t1 = 0, t2 = Inf, sided = 2, nboot = 1000,
  alpha = 0.05, compo = FALSE, seed = 1011, nlimit = 200)
```

**Arguments**

data	a data frame/matrix with 3 columns. The first column is the survival time. The second is the censoring indicator. The last is the grouping variable. An example as the input to data provided is <a href="#">hepatitis</a> .
g1	the group with the longer survival that should take a value from the third column of data
t1	pre-specified $t_1$ based on domain knowledge with the default value of 0
t2	pre-specified $t_2$ based on domain knowledge with the default value of $\infty$
sided	2 if two-sided test, and 1 if one-sided test. It assumes the default value of 2.
nboot	number of bootstrap replications in calculating critical values
alpha	pre-specified significance level of the test
compo	A logical value specifying whether to change the local statistics at time points outside the overlapping region of the two samples. The default is FALSE. TRUE if uses a functional of the difference between the Kaplan–Meier estimators (see "Details" for more).
seed	the parameter to <a href="#">set.seed</a> for the random number generator in R. The <code>set.seed</code> is used implicitly in <code>supELtest</code> .
nlimit	the splitting unit. To deal with large data problems, the bootstrap algorithm is to split the number of bootstrap replicates into <code>nsplit</code> parts. The number <code>nsplit</code> is the smallest integer not less than $\lceil U \rceil / \text{nlimit}$ .

**Value**

ptwiseELtest returns a list with four elements:

- `time_pts` the values of statistics at each uncensored time point
- `decision` logical values. See `stat_ptwise`.
- `stat_ptwise` the decision of the test in which the null hypothesis is rejected at a specific day if the decision exhibits 1 and not rejected if otherwise
- `critval_ptwise` the critical values of the statistic at each uncensored time point

## References

H. W. Chang, "Empirical likelihood tests for stochastic ordering based on censored and biased data," *Columbia University Academic Commons* (2014). <http://academiccommons.columbia.edu/catalog/ac%3A177230>

## See Also

[hepatitis](#), [intELtest](#), [supELtest](#)

## Examples

```
library(EL2Surv)
ptwiseELtest(hepatitis)
## It produces the estimates on 47 distinct uncensored days
## out of 57 possibly repeated uncensored days.

ptwiseELtest(hepatitis, t1 = 30, t2 = 60)
## It produces the estimates on 12 distinct uncensored days
## on the restricted time interval [30, 60].
```

---

supELtest

---

*The maximally selected likelihood ratio test*


---

## Description

supELtest provides a maximal deviation type statistics that is better adapted at detecting local differences:

$$\sup_{t \in U} \{-2 \log R(t)\},$$

where  $w(t)$  is an objective weight function, and  $R(t)$  is an empirical likelihood (EL) ratio that compares two survival functions at each time point  $t$  in the set of observed uncensored lifetimes,  $U$ .

## Usage

```
supELtest(data, g1 = 1, t1 = 0, t2 = Inf, sided = 2, nboot = 1000,
  alpha = 0.05, compo = FALSE, seed = 1011, nlimit = 200)
```

## Arguments

data	a data frame/matrix with 3 columns. The first column is the survival time. The second is the censoring indicator. The last is the grouping variable. An example as the input to data provided is <a href="#">hepatitis</a> .
g1	the group with the longer survival that should take a value from the third column of data
t1	pre-specified $t_1$ based on domain knowledge with the default value of 0
t2	pre-specified $t_2$ based on domain knowledge with the default value of $\infty$

sided	2 if two-sided test, and 1 if one-sided test. It assumes the default value of 2.
nboot	number of bootstrap replications in calculating critical values
alpha	pre-specified significance level of the test
compo	A logical value specifying whether to change the local statistics at time points outside the overlapping region of the two samples. The default is FALSE. TRUE if uses a functional of the difference between the Kaplan–Meier estimators (see "Details" for more).
seed	the parameter to <code>set.seed</code> for the random number generator in R. The <code>set.seed</code> is used implicitly in <code>supELtest</code> .
nlimit	the splitting unit. To deal with large data problems, the bootstrap algorithm is to split the number of bootstrap replicates into <code>nsplit</code> parts. The number <code>nsplit</code> is the smallest integer not less than $\ U\ /nlimit$ .

### Value

`supELtest` returns a list with three elements:

- `critval` the critical value
- `teststat` the resulting integrated test statistic
- `pvalue` the p-value based on the integrated statistic

### References

H. W. Chang, "Empirical likelihood tests for stochastic ordering based on censored and biased data," *Columbia University Academic Commons* (2014). <http://academiccommons.columbia.edu/catalog/ac%3A177230>

### See Also

`hepatitis`, `intELtest`, `ptwiseELtest`

### Examples

```
library(EL2Surv)
supELtest(hepatitis, 1, sided = 2)

## OUTPUT:
## $critval
## [1] 7.247709
##
## $teststat
## [1] 10.3581
##
## $pvalue
## [1] 0.013
```



# Index

\*Topic **datasets**

hepatitis, [2](#)

hepatitis, [2](#), [3](#), [5–8](#)

intELtest, [2](#), [3](#), [7](#), [8](#)

ptwiseELtest, [5](#), [6](#), [8](#)

set.seed, [3](#), [6](#), [8](#)

supELtest, [2](#), [5](#), [7](#), [7](#)