Package: icsp2 (via r-universe)

May 29, 2026

Title Stratified interval-censored semi-parametric survival models

Version 0.0.0.9000

Description Semi-parametric survival models for interval censored data with stratification, based on the `icenReg` package, including model fitting and testing.

BugReports https://github.com/igrave/icsp2/issues

URL https://igrave.github.io/icsp2/

License LGPL (>= 2.0, < 3)

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Depends R (>= 4.2.0)

LinkingTo Rcpp, RcppEigen

Imports Rcpp, survival

Suggests testthat (>= 3.0.0), icenReg, vdiffr, knitr, rmarkdown

Config/testthat/edition 3

LazyData TRUE

language en

VignetteBuilder knitr

SystemRequirements GNU make

Config/pak/sysreqs make

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

Date/Publication 2026-05-06 20:30:51 UTC

RemoteUrl https://github.com/igrave/icsp2

RemoteRef HEAD

RemoteSha f9e1f689a0a9fc2b5a5f99cf8d3c324506774b22

Contents

`cpp_logrank` *Internal fast log-rank test implementation*

Description

This is an internal implementation of right-censored logrank test for use within ic_logrank where many tests need to be performed efficiently for variance calculation.

Usage

```
cpp_logrank(time, event, group)
```

Arguments

`time`	Numeric vector of survival/censoring times.
`event`	Logical vector of event indicators (TRUE = event, FALSE = censored).
`group`	Integer vector of group labels, coded as consecutive integers (i.e. `1, 2, ...`). This is not checked.

Value

An object of class `"cpp_logrank"`, a list with components:

observed Numeric vector of observed events per group.

expected Numeric vector of expected events per group.

variance Numeric variance matrix

n_by_group Numeric vector of sample sizes per group.

Examples

```
set.seed(1992)
time  <- rexp(100)
event <- sample(c(TRUE, FALSE), 100, replace = TRUE)
group <- sample(1:2, 100, replace = TRUE)
cpp_logrank(time, event, group)
```

---

ic_logrank *Sun's Log-Rank Test for Interval-Censored Data*

Description

Sun's Log-Rank Test for Interval-Censored Data

Usage

```
ic_logrank(
  formula,
  data,
  subset,
  na.action,
  B = c(0, 1),
  n_samples = 1000,
  type = c("sas", "hly"),
  ...
)
```

Arguments

formula	A formula with `Surv(l, u, type = 'interval2')` response and a single grouping variable on the right-hand side. May also contain `strata()` terms for stratified analysis.
data	A data frame containing the variables in the formula.
subset	Optional expression indicating which subset of rows to use.
na.action	Function to handle missing values.
B	A vector of length 2 giving bounds for observation times. Default is c(0, 1).
n_samples	The number of "imputation" samples for the variance calculation. Default is 1000.
type	The method for calculating the test statistic and variance. Options are `"sas"` (default) or `"hly"`.
...	Additional arguments (currently unused).

Details

Performs Sun's (1996) non-parametric log-rank test for comparing survival distributions across groups with interval-censored data. This test compares group-specific NPMLE survival curves without assuming proportional hazards.

The test statistic is $Q = U(0)'I(0)^{-1}U(0)$, which follows a chi-squared distribution with k-1 degrees of freedom under the null hypothesis, where k is the number of groups.

The default test type is `"sas"`, which uses Sun's test statistic combined with variance estimated based on the Huang, Lee and Yu (2008) procedure sampling exact observation times from the Turnbull intervals.

Alternatively, the `"hly"` type calculates the test statistic and variance using the multiple imputation approach of Huang, Lee and Yu (2008) directly.

Stratified tests are constructed by calculating the U and V matrices for each stratum separately and then summing the stratum-specific matrices to give a global test statistic $\sum \bar{U}'(\sum \hat{V})^{-1} \sum \bar{U}$. This is the procedure described in the SAS documentation for PROC ICLIFETEST.

Value

An object of class `ic_logrank` containing:

`logrank`	The log-rank statistics "observed - expected" for all groups and strata
`logrank_overall`	The log-rank statistics "observed - expected" for all groups
`statistic`	The overall chi-squared test statistic based on imputation
`df`	Degrees of freedom
`p.value`	P-value from chi-squared distribution
`var`	Variance-covariance matrix
`groups`	Group levels being compared
`n`	Sample sizes per group
`method`	Description of test method
`data.name`	Name of the data
`call`	The matched call

References

Sun, J. (1996). A non-parametric test for interval-censored failure time data with application to AIDS studies. *Statistics in Medicine*, 15(13), 1387-1395.

Huang, J., Lee, C., and Yu, Q. (2008). A Generalized Log-Rank Test for Interval-Censored Failure Time Data via Multiple Imputation. *Statistics in Medicine 27:3217–3226*. http://dx.doi.org/10.1002/sim.3211

SAS Institute Inc. (2026). *SAS/STAT® 26.03 User's Guide: The ICLIFETEST Procedure*. https://documentation.sas.com/doc/en/statug/latest/statug_iclifetest_details01.htm Accessed 14 April 2026.

Examples

```
# Simple two-group comparison
data(miceData)
ic_logrank(Surv(l, u, type = "interval2") ~ grp, data = miceData)
```

`ic_sp_control` *Control ic_sp2 fitting options*

Description

Control ic_sp2 fitting options

Usage

```
ic_sp_control(
  useGA = TRUE,
  maxIter = 10000,
  baseUpdates = 5,
  regStart = NULL,
  derivMethod = c(12, 1),
  updateCovars = TRUE
)
```

Arguments

`useGA`	Should constrained gradient ascent step be used?
`maxIter`	Maximum iterations
`baseUpdates`	number of baseline updates (ICM + GA) per iteration
`regStart`	Initial values for regression parameters
`derivMethod`	Method for derivative calculations.
`updateCovars`	Should covariates be updated during fitting? @description Creates the control options for `ic_sp2()`. Defaults not intended to be changed for use in standard analyses.

Details

The constrained gradient step, controlled by `useGA = TRUE`, is a step that was added to improve the convergence in a special case. The option to turn it off is only in place to help demonstrate its utility.

`regStart` also for seeding of initial value of regression parameters. Intended for use in "warm start" for bootstrap samples and providing fixed regression parameters.

ic_sp2 *Fit a semi-parametric model for interval-censored data*

Description

Fit a semi-parametric model for interval-censored data

Usage

```
ic_sp2(
  formula,
  data,
  weights,
  subset,
  na.action,
  B = c(0, 1),
  control = ic_sp_control(...),
  model = c("ph", "po"),
  profile_ci = 0,
  ...
)

ic_sp_ph(
  formula,
  data,
  weights,
  subset,
  na.action,
  B = c(0, 1),
  control = ic_sp_control(...),
  model = c("ph", "po"),
  profile_ci = 0,
  ...
)

ic_sp_po(
  formula,
  data,
  weights,
  subset,
  na.action,
  B = c(0, 1),
  control = ic_sp_control(...),
  model = c("ph", "po"),
  profile_ci = 0,
  ...
)
```

Arguments

formula	A model formula with Surv(l, u, type = 'interval2') response and covariates on the right-hand side. May also contain strata() terms.
data	A data frame containing the variables in the formula, including strata terms.
weights	Optional vector of weights for each observation, or the name of a variable in data containing the weights.
subset	Optional expression indicating a subset of the rows of data to be used in the fit.
na.action	Optional function to handle missing data. Default is na.omit.
B	A vector of length 2 giving the lower and upper bounds for the observation times. Default is c(0, 1).
control	A list of control settings, with defaults created by ic_sp_control().
model	Type of model to fit. Choices are "ph" for proportional hazards and "po" for proportional odds. Default is "ph". This is normally determined by the function aliases ic_sp_ph and ic_sp_po.
profile_ci	Confidence level for profile likelihood confidence intervals. Default is 0.95. Set to NULL to skip profile likelihood confidence interval calculations.
...	Additional arguments passed to control.

Value

A list containing the fitted model information, including coefficients, variance-covariance matrix, and other details.

IR_diabetes *Interval censored time from diabetes onset to diabetic nephronpathy*

Description

Data set contains interval censored survival time for time from onset of diabetes to to diabetic nephronpathy. Identical to the diabetes dataset found in the package glrt.

Usage

```
IR_diabetes
```

Format

An object of class data.frame with 731 rows and 3 columns.

Fields

left left side of observation interval

right right side of observation interval

gender gender of subject

References

Borch-Johnsens, K, Andersen, P and Decker, T (1985). "The effect of proteinuria on relative mortality in Type I (insulin-dependent) diabetes mellitus." Diabetologia, 28, 590-596.

Examples

```
head(IR_diabetes)
diabetes_fit <- ic_sp_po(
    Surv(left, right, type = 'interval2') ~ gender,
    data = IR_diabetes
    )
```

`lines.ic_sp2` *Lines method for ic_sp2 objects*

Description

Lines method for ic_sp2 objects

Usage

```
## S3 method for class 'ic_sp2'
lines(x, y, type = "s", strata, ...)
```

Arguments

`x`	an object of class 'ic_sp2'
`y`	ignored
`type`	the type of plot to produce, default is "l" for lines
`strata`	Optional vector of indices to subset the strata to be plotted.
`...`	additional arguments passed to the lines function

Details

Due to the interval censoring, the survival curves are not unique, so two lines are drawn for each curve: one for the lower bound and one for the upper bound of the interval. A pair of lines is drawn for each stratum, if strata are present in the model. Therefore to unambiguously specify graphics parameters such as colour or line type, the user should specify them as a vector of length equal to the number of strata (for strata specific pars) or twice the number of strata (for line specific pars, e.g. strata 1 lower, strata 1 upper, strata 2 lower, strata 2 upper).

miceData *Lung Tumor Interval Censored Data from Hoel and Walburg 1972*

Description

RFM mice were sacrificed and examined for lung tumors. This resulted in current status interval censored data: if the tumor was present, this implied left censoring and if no tumor was present this implied right censoring. Mice were placed in two different groups: conventional environment or germ free environment.

Usage

```
miceData
```

Format

An object of class `data.frame` with 144 rows and 3 columns.

Fields

`l` left side of observation interval

`u` right side of observation interval

`grp` Group for mouse. Either `"ce"` (conventional environment) or `"ge"` (germ-free environment)

References

Hoel D. and Walburg, H.,(1972), Statistical analysis of survival experiments, \emph{The Annals of Statis
18, 1259-1294

Examples

```
head(miceData)
mice_ph_fit <- ic_sp_ph(Surv(l, u, type = 'interval2') ~ grp, data = miceData)

summary(mice_ph_fit)
```

print.ic_sp2 *Print method for ic_sp2 objects*

Description

Print method for ic_sp2 objects

Usage

```
## S3 method for class 'ic_sp2'
print(x, ...)
```

Arguments

x	Fitted model object from ic_sp
...	Additional arguments.

---

profile_fit *Refit the model with fixed regression parameters*

Description

Refit the model with fixed regression parameters

Usage

```
profile_fit(object, beta = object$coefficients)
```

Arguments

object	Fitted model object from ic_sp
beta	Vector of regression parameters to fix in the refit. Default is the original fitted regression parameters.

Value

Raw fitted model object with fixed regression parameters.

---

simIC_weib *Simulates interval censored data from regression model with a Weibull baseline*

Description

Simulates interval censored data from a regression model with a Weibull baseline distribution. Used for demonstration

Usage

```
simIC_weib(
  n = 100,
  b1 = 0.5,
  b2 = -0.5,
  model = "ph",
  shape = 2,
  scale = 2,
  inspections = 2,
  inspectLength = 2.5,
  rndDigits = NULL,
  prob_cen = 1
)
```

Arguments

n	Number of samples simulated
b1	Value of first regression coefficient
b2	Value of second regression coefficient
model	Type of regression model. Options are 'po' (prop. odds) and 'ph' (Cox PH)
shape	shape parameter of baseline distribution
scale	scale parameter of baseline distribution
inspections	number of inspections times of censoring process
inspectLength	max length of inspection interval
rndDigits	number of digits to which the inspection time is rounded to, creating a discrete inspection time. If rndDigits = NULL, the inspection time is not rounded, resulting in a continuous inspection time
prob_cen	probability event being censored. If event is uncensored, l == u

Details

Exact event times are simulated according to regression model: covariate x1 is distributed rnorm(n) and covariate x2 is distributed 1 - 2 * rbinom(n, 1, 0.5). Event times are then censored with a case II interval censoring mechanism with inspections different inspection times. Time between inspections is distributed as runif(min = 0, max = inspectLength). Note that the user should be careful in simulation studies not to simulate data where nearly all the data is right censored (or more over, all the data with x2 = 1 or -1) or this can result in degenerate solutions!

Author(s)

Clifford Anderson-Bergman

Examples

```
set.seed(1)
sim_data <- simIC_weib(n = 500, b1 = .3, b2 = -.3, model = 'ph',
                       shape = 2, scale = 2, inspections = 6,
                       inspectLength = 1)
#simulates data from a cox-ph with beta Weibull distribution.

ic_sp_ph(Surv(l, u, type = 'interval2') ~ x1 + x2, data = sim_data)
ic_sp_po(Surv(l, u, type = 'interval2') ~ x1 + x2, data = sim_data)

#'ph' fit looks better than 'po'; the difference between the transformed survival
#function looks more constant
```

vcov.ic_sp2 *Profile Likelihood Covariance for Semi-Parametric Models*

Description

Profile Likelihood Covariance for Semi-Parametric Models

Usage

```
## S3 method for class 'ic_sp2'
vcov(object, type = "oim_curvature", fixed = 5, typical = 1, large = 2, ...)
```

Arguments

object	Fitted model object from ic_sp
type	One of "oim_curvature" (default), "oim_fixed", or "opg_fixed". See details for explanation.
fixed	A fixed factor to multiply by n^(-1/2) to determin the perturbation size for fixed types.
typical	A typical value for the regression parameters, used to determine the scale of h_n. Default is 1. This is required for the "oim_curvature" type.
large	A large value for the regression parameters, used to determine the scale of h_n. Default is 2. This is required for the "oim_curvature" type.
...	Unused.

Details

The covariance matrix is calculated using the profile likelihood approach. (Murphey and Vand Der Vaart 2000). This method involves perturbing the regression parameters, updating the baseline hazard estimates using the profile_fit function with the perturbed parameters, and calculating the change in log-likelihood for each perturbation, which is then used to compute the covariance matrix. We borrowing the naming convention from the Stata stintcox manual https://www.stata.com/manuals/ststintcox.pdf.

Type "oim_curvature" (Boruvka and Cook 2015) uses the curvature of the log-likelihood function to determine the perturbation size for the profile likelihood calculations, as described in Boruvka and Cook (2014). The typical and large parameters are used to determine the scale of the perturbations.

Type "oim_fixed" (Zeng et al 2016) uses a fixed perturbation size based on the fixed parameter and the sample size n.

Type "opg_fixed" (Zeng et al 2017) uses a fixed perturbation size based on the fixed parameter and the sample size n, but uses the outer product of gradients instead of the observed information matrix for the covariance calculation.

For larged values of fixed the model fitting may fail to converge.

Value

Variance-covariance matrix of the regression parameters.

References

Murphy, S. A., & Van Der Vaart, A. W. (2000). On Profile Likelihood. Journal of the American Statistical Association, 95(450), 449–465. https://doi.org/10.1080/01621459.2000.10474219

Boruvka, A., and Cook, R. J. (2015), A Cox-Aalen Model for Interval-censored Data. Scand J Statist, 42, 414–426. doi: 10.1111/sjos.12113.

Donglin Zeng, Lu Mao, D. Y. Lin, Maximum likelihood estimation for semiparametric transformation models with interval-censored data, Biometrika, Volume 103, Issue 2, June 2016, Pages 253–271, https://doi.org/10.1093/biomet/asw013

Donglin Zeng, Fei Gao, D. Y. Lin, Maximum likelihood estimation for semiparametric regression models with multivariate interval-censored data, Biometrika, Volume 104, Issue 3, September 2017, Pages 505–525, https://doi.org/10.1093/biomet/asx029

Index