Package 'SurvRegCensCov'

October 8, 2015

Type Package

Title Weibull Regression for a Right-Censored Endpoint with Interval-Censored Covariate

Version 1.4

Date 2015-10-02

Author Stanislas Hubeaux <stanislas.hubeaux@alumni.epfl.ch> and Kaspar Rufibach <kaspar.rufibach@gmail.com>

Maintainer Stanislas Hubeaux <stanislas.hubeaux@alumni.epfl.ch>

Depends R (>= 2.10), survival, stats, graphics

Imports numDeriv

Description The main function of this package allows estimation of a Weibull Regression for a right-censored endpoint, one interval-censored covariate, and an arbitrary number of non-censored covariates. Additional functions allow to switch between different parametrizations of Weibull regression used by different R functions, inference for the mean difference of two arbitrarily censored Normal samples, and estimation of canonical parameters from censored samples for several distributional assumptions.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2015-10-08 20:17:22

R topics documented:

`SurvRegCensCov-package`

Weibull Regression for a Right-Censored Endpoint with Interval-Censored Covariates

---

Description

The main function `SurvRegCens` of this package allows estimation of a Weibull Regression for a right-censored endpoint, one interval-censored covariate, and an arbitrary number of non-censored covariates. Additional functions allow to switch between different parametrizations of Weibull regression used by different `R` functions (`ConvertWeibull`, `WeibullReg`, `WeibullDiag`), inference for the mean difference of two arbitrarily censored Normal samples (`NormalMeanDiffCens`), and estimation of canonical parameters from censored samples for several distributional assumptions (`ParamSampleCens`).

Details

Package:	SurvRegCensCov
Type:	Package
Version:	1.4
Date:	2015-10-02
License:	GPL (>=2)

Author(s)

Stanislas Hubeaux (maintainer), `<stanislas.hubeaux@alumni.epfl.ch>`

Kaspar Rufibach, `<kaspar.rufibach@gmail.com>`
`http://www.kasparrufibach.ch`

We thank Sarah Haile for contributing the functions `ConvertWeibull`, `WeibullReg`, `WeibullDiag` to the package.

References

Hubeaux, S. (2013). Estimation from left- and/or interval-censored samples. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Hubeaux, S. (2013). Parametric Surival Regression Model with left- and/or interval-censored covariate. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Hubeaux, S. and Rufibach, K. (2014). **SurvRegCensCov**: Weibull Regression for a Right-Censored Endpoint with a Censored Covariate. Preprint, http://arxiv.org/abs/1402.0432.

Lynn, H. S. (2001). Maximum likelihood inference for left-censored HIV RNA data. *Stat. Med.*, **20**, 33–45.

Sattar, A., Sinha, S. K. and Morris, N. J. (2012). A Parametric Survival Model When a Covariate is Subject to Left-Censoring. *Biometrics & Biostatistics*, **S3**(2).

Examples

```
# The main functions in this package are illustrated in their respective help files.
```

CDF *Cumulative distribution function*

Description

Evaluates the cumulative distribution function using the integral of its density function.

Usage

```
CDF(c, density)
```

Arguments

c	Value at which the CDF is to be evaluated.
density	Density function.

Note

Function not intended to be invoked by the user.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

censorContVar *Censor a vector of continuous numbers*

Description

Given a vector of realizations of a continuous random variable, interval-, left-, or right-censor these numbers at given boundaries. Useful when setting up simulations involving censored observations.

Usage

```
censorContVar(x, LLOD = NA, ULOD = NA)
```

Arguments

x	Vector of random numbers.
LLOD	Lower limit where x should be left-censored at. If no left-censoring, set to NA.
ULOD	Upper limit where x should be left-censored at. If no left-censoring, set to NA.

Value

A data.frame as specified by code = interval2 in Surv.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

Examples

```
## random vector
x <- rnorm(200)

## interval-censor this vector at -1 and 0.5
censorContVar(x, -1, 0.5)
```

coef.src *Extract coefficients of Weibull regression with an interval-censored covariate*

Description

coef method for class "src".

Usage

```
## S3 method for class 'src'
coef(object, ...)
```

Arguments

object	An object of class "src", usually a result of a call to SurvRegCens.
...	Further arguments.

Value

The function coef.src returns the estimated parameters of the Weibull regression when calling SurvRegCens.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

References

Hubeaux, S. (2013). Parametric Surival Regression Model with left- and/or interval-censored covariate. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Hubeaux, S. and Rufibach, K. (2014). **SurvRegCensCov**: Weibull Regression for a Right-Censored Endpoint with a Censored Covariate. Preprint, http://arxiv.org/abs/1402.0432.

Sattar, A., Sinha, S. K. and Morris, N. J. (2012). A Parametric Survival Model When a Covariate is Subject to Left-Censoring. *Biometrics & Biostatistics*, **S3**(2).

Examples

```
## See help file of function "SurvRegCens".
```

`ConvertWeibull` *Transformation of survreg output for the Weibull distribution*

Description

Transforms output from `survreg` using the Weibull distribution to a more natural parameterization. See details and the vignette for more information.

Usage

```
ConvertWeibull(model, conf.level = 0.95)
```

Arguments

`model`	A `survreg` model, with `dist = "weibull"`.
`conf.level`	Confidence level used to produce two-sided $1-\alpha/2$ confidence intervals for the hazard and event time ratios.

Details

The `survreg` function fits a Weibull accelerated failure time model of the form

$$\log t = \mu + \alpha^T Z + \sigma W,$$

where Z is a matrix of covariates, and W has the extreme value distribution, μ is the intercept, α is a vector of parameters for each of the covariates, and σ is the scale. The usual parameterization of the model, however, is defined by hazard function

$$h(t|Z) = \gamma\lambda t^{\gamma-1}\exp(\beta^T Z).$$

The transformation is as follows: $\gamma = 1/\sigma$, $\lambda = \exp(-\mu/\sigma)$, and $\beta = -\alpha/\sigma$, and estimates of the standard errors can be found using the delta method.

The Weibull distribution has the advantage of having two separate interpretations. The first, via proportional hazards, leads to a hazard ratio, defined by $\exp\beta$. The second, of accelerated failure times, leads to an event time ratio (also known as an acceleration factor), defined by $\exp(-\beta/\gamma)$.

Further details regarding the transformations of the parameters and their standard errors can be found in Klein and Moeschberger (2003, Chapter 12). An explanation of event time ratios for the accelerated failure time interpretation of the model can be found in Carroll (2003). A general overview can be found in the `vignette("weibull")` of this package.

Value

`vars`	A matrix containing the values of the transformed parameters and their standard errors
`HR`	A matrix containing the hazard ratios for the covariates, and $1-$`conf.level`$/2$ confidence intervals.
`ETR`	A matrix containing the event time ratios for the covariates, and $1-$`conf.level`$/2$ confidence intervals.

Author(s)

Sarah R. Haile, <sarah.haile@gmail.com>

References

Carroll, K. (2003). On the use and utility of the Weibull model in the analysis of survival data. *Controlled Clinical Trials*, **24**, 682–701.

Klein, J. and Moeschberger, M. (2003). *Survival analysis: techniques for censored and truncated data*. 2nd edition, Springer.

See Also

This function is used by `WeibullReg`.

Examples

```
data(larynx)
ConvertWeibull(survreg(Surv(time, death) ~ stage + age, larynx), conf.level = 0.95)
```

---

larynx *Survival Times of Larynx Cancer Patients*

Description

A study of 90 males with laryngeal cancer was performed, comparing survival times. Each patient's age, year of diagnosis, and disease stage was noted, see Kardaun (1983) and Klein and Moeschberger (2003).

Usage

```
data(larynx)
```

Format

A data frame with 90 observations on the following 5 variables.

`stage` Disease stage (1-4) from TNM cancer staging classification.

`time` Time from first treatment until death, or end of study.

`age` Age at diagnosis.

`year` Year of diagnosis.

`death` Indicator of death [1, if patient died at time t; 0, otherwise].

Source

http://www.mcw.edu/FileLibrary/Groups/Biostatistics/Publicfiles/DataFromSection/DataFromSectionTXT/Data_from_section_1.8.txt

References

Kardaun, O. (1983). Statistical survival analysis of male larynx-cancer patients-a case study. *Statistica Neerlandica*, **37**, 103–125.

Klein, J. and Moeschberger, M. (2003). *Survival analysis: techniques for censored and truncated data*. 2nd edition, Springer.

Examples

```
library(survival)
data(larynx)
Surv(larynx$time, larynx$death)
```

`logLik.src` *Extract value of log-likelihood at maximum for Weibull regression with an interval-censored covariate*

Description

`logLik` method for class `"src"`.

Usage

```
## S3 method for class 'src'
logLik(object, ...)
```

Arguments

`object`	An object of class `"src"`, usually a result of a call to `SurvRegCens`.
`...`	Further arguments.

Value

The function `logLik.src` returns the value of the log-likelihood at the maximum likelihood estimate, as well as the corresponding degrees of freedom.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

References

Hubeaux, S. (2013). Parametric Surival Regression Model with left- and/or interval-censored covariate. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Hubeaux, S. and Rufibach, K. (2014). **SurvRegCensCov**: Weibull Regression for a Right-Censored Endpoint with a Censored Covariate. Preprint, http://arxiv.org/abs/1402.0432.

Sattar, A., Sinha, S. K. and Morris, N. J. (2012). A Parametric Survival Model When a Covariate is Subject to Left-Censoring. *Biometrics & Biostatistics*, **S3**(2).

Examples

```
## See help file of function "SurvRegCens".
```

`LoglikCens` *Log-likelihood functions for estimation of canonical parameters from a censored sample*

Description

Computes the log-likelihood function for a censored sample, according to a specified distributional assumptions. Available distributions are Normal, Weibull, Logistic, and Gamma.

Usage

```
LoglikNormalCens(x, data, lowerbound, vdelta)
LoglikWeibullCens(x, data, lowerbound, vdelta)
LoglikLogisticCens(x, data, lowerbound, vdelta)
LoglikGammaCens(x, data, lowerbound, vdelta)
```

Arguments

`x`	Two-dimensional vector giving the canonical parameters of the distribution.
`data`	Observed or censored event times.
`lowerbound`	A vector that collect lower bounds for the interval-censored observations. If no lower bound is available then put `NA`.
`vdelta`	A vector which indicates censoring (0: censored, 1: not censored).

Note

Function not intended to be invoked by the user.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

References

Hubeaux, S. (2013). Estimation from left- and/or interval-censored samples. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Lynn, H. S. (2001). Maximum likelihood inference for left-censored HIV RNA data. *Stat. Med.*, **20**, 33–45.

`LoglikNormalDeltaCens` *Log likelihood function to compute mean difference between two normally distributed censored samples.*

Description

Reparametrization of the log likelihood function for a normally distributed censored sample such that the mean difference is a parameter of the function, thus allowing to be made inference on. The mean difference is computed as sample 1 - sample 2.

Usage

```
LoglikNormalDeltaCens(x, data1, lowerbound1, vdelta1, data2,
                      lowerbound2, vdelta2)
```

Arguments

`x`	A vector of four components where the first component corresponds to the mean of the normal distribution of the first sample, the second component corresponds to mean difference between the two samples: sample 1 - sample 2, the third component corresponds to the standard deviation of the normal distribution of the first sample, and the fourth component corresponds to the standard deviation of the normal distribution of the second sample.
`data1`	A vector of data corresponding to the first sample.
`lowerbound1`	A vector which corresponds to the lower bounds for the interval-censored observations of the vector of data corresponding to the first sample. If no lower bound is available then put `NA`.
`vdelta1`	A vector which indicates for censoring for the first sample (0: censored, 1: not censored).
`data2`	A vector of data corresponding to the second sample.
`lowerbound2`	A vector which corresponds to the lower bounds for the interval-censored observations of the vector of data corresponding to the second sample. If no lower bound is available then put `NA`.
`vdelta2`	A vector which indicates for censoring for the second sample (0: censored, 1: not censored).

Note

Function not intended to be invoked by the user.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

References

Hubeaux, S. (2013). Estimation from left- and/or interval-censored samples. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Lynn, H. S. (2001). Maximum likelihood inference for left-censored HIV RNA data. *Stat. Med.*, **20**, 33–45.

`LoglikWeibullSurvRegCens` *Log-likelihood function of a Weibull Survival Regression Model allowing for an interval-censored covariate.*

Description

Computes the log-likelihood function of a Weibull Survival Regression Model allowing for an interval-censored covariate.

Usage

```
LoglikWeibullSurvRegCens(x, data_y, data_delta_loglik, data_cov_noncens = NULL,
                         data_cov_cens, density, data_r_loglik, data_lowerbound,
                          intlimit = 10^-10)
```

Arguments

`x`	Vector of parameters, ordered as follows: Scale parameter, Shape parameter, regression parameters (i.e. β) linked to the non-censored covariates, regression parameter (i.e. β) linked to the censored covariate.
`data_y`	Time-to-event vector.
`data_delta_loglik`	Censored indicator vector of the time-to-event (0: censored, 1: not censored).
`data_cov_noncens`	Matrix where each column represents a non-censored covariate.
`data_cov_cens`	Censored covariate vector.
`density`	Density function of the censored covariate.
`data_r_loglik`	Censored indicator vector of the censored covariate (0: censored, 1: not censored).

`data_lowerbound`
: A vector which corresponds to the lower bounds for the interval-censored observations of the censored covariate. If no lower bound is available then put `NA`.

`intlimit`
: In computation of integrals, values of the function to be integrated below `intlimit` are set to 0. This makes integration results more accurate and speeds up integration. If the data is such that the absolute values of the underlying baseline Weibull density are very small, i.e. in the range of `intlimit`, it is advisable to rescale the time variable, e.g. change the scaling from days to years. A very small value of the estimated λ is indicative of that situation.

Note

Function not intended to be invoked by the user.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

References

Hubeaux, S. (2013). Parametric Surival Regression Model with left- and/or interval-censored covariate. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Sattar, A., Sinha, S. K. and Morris, N. J. (2012). A Parametric Survival Model When a Covariate is Subject to Left-Censoring. *Biometrics & Biostatistics*, **S3**(2).

`NormalMeanDiffCens` *Maximum Likelihood Estimator for the mean difference between two censored normally distributed samples*

Description

Computes estimates of the parameters of two censored Normal samples, as well as the mean difference between the two samples.

Usage

```
NormalMeanDiffCens(censdata1, censdata2, conf.level = 0.95,
     null.values = c(0, 0, 1, 1))
```

Arguments

censdata1	Observations of first sample, format as specified by code = interval2 in Surv.
censdata2	Observations of second sample, as specified by code = interval2 in Surv.
conf.level	Confidence level for confidence intervals.
null.values	Fixed values for hypothesis tests. Four-dimensional vector specifying the hypothesis for μ_1, μ_2, σ_1, σ_2.

Value

A table with estimators and inference for the means and standard deviations of both samples, as well as the difference Δ between the mean of the first and second sample. Hypothesis tests are for the values in null.values and for the null hypothesis of no mean difference.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

References

Hubeaux, S. (2013). Estimation from left- and/or interval-censored samples. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Lynn, H. S. (2001). Maximum likelihood inference for left-censored HIV RNA data. *Stat. Med.*, **20**, 33–45.

Examples

```
## example with interval-censored Normal samples
n <- 500
prop.cens <- 0.35
mu <- c(0, 2)
sigma <- c(1, 1)

set.seed(2013)

## Sample 1:
LOD1 <- qnorm(prop.cens, mean = mu[1], sd = sigma[1])
x1 <- rnorm(n, mean = mu[1], sd = sigma[1])
s1 <- censorContVar(x1, LLOD = LOD1)

## Sample 2:
LOD2 <- qnorm(0.35, mean = mu[2], sd = sigma[2])
x2 <- rnorm(n, mean = mu[2], sd = sigma[2])
s2 <- censorContVar(x2, LLOD = LOD2)

## inference on distribution parameters and mean difference:
NormalMeanDiffCens(censdata1 = s1, censdata2 = s2)
```

`ParamSampleCens` *Maximum Likelihood Estimator of parameters from a censored sample*

Description

Computes maximum likelihood estimators of the canonical parameters for several distributions, based on a censored sample.

Usage

```
ParamSampleCens(censdata, dist = c("normal", "logistic", "gamma", "weibull")[1],
     null.values = c(0, 1), conf.level = 0.95, initial = NULL)
```

Arguments

`censdata`	Dataframe that contains censored data, format as specified by `code = interval2` in `Surv`.
`dist`	Assumed distribution of the sample.
`null.values`	Fixed values for hypothesis tests.
`conf.level`	Confidence level of confidence intervals.
`initial`	Initial values for the maximization.

Value

`coeff`	Estimators, standard errors, confidence intervals, and 2-sided p-values for the null hypothesis as given in `null.values`. All this inference is based on maximum likelihood theory, i.e. standard errors are computed using the inverse expected Hessian at the maximum likelihood estimator.
`percent.cens`	Percentage of censored observations.
`loglik`	Log likelihood function value at the estimator.
`info.converg`	Convergence information provided by the function `optim`.
`info.converg.message`	Message provided by the function `optim`.

Note

Functions with similar functionality are provided in the package **fitdistrplus**.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

References

Hubeaux, S. (2013). Estimation from left- and/or interval-censored samples. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Lynn, H. S. (2001). Maximum likelihood inference for left-censored HIV RNA data. *Stat. Med.*, **20**, 33–45.

Examples

```
n <- 500
prop.cens <- 0.35

## example with a left-censored Normally distributed sample
set.seed(2013)
mu <- 3.5
sigma <- 1
LOD <- qnorm(prop.cens, mean = mu, sd = sigma)
x1 <- rnorm(n, mean = mu, sd = sigma)
s1 <- censorContVar(x1, LLOD = LOD)
ParamSampleCens(censdata = s1)


## example with an interval-censored Normal sample
set.seed(2013)
x2 <- rnorm(n, mean = mu, sd = sigma)
LOD <- qnorm(prop.cens / 2, mean = mu, sd = sigma)
UOD <- qnorm(1 - prop.cens / 2, mean = mu, sd = sigma)
s2 <- censorContVar(x2, LLOD = LOD, ULOD = UOD)
ParamSampleCens(censdata = s2)


## Not run:
## compare to fitdistrplus
library(fitdistrplus)
s2 <- as.data.frame(s2)
colnames(s2) <- c("left", "right")
summary(fitdistcens(censdata = s2, distr = "norm"))

## End(Not run)
```

`print.src` *Print result of Weibull regression with an interval-censored covariate*

Description

`print` method for class `"src"`.

Usage

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

Arguments

x	An object of class "src", usually a result of a call to SurvRegCens.
...	Further arguments.

Value

The function print.src returns the estimated parameters of the Weibull regression, incl. AIC, when calling SurvRegCens.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

References

Hubeaux, S. (2013). Parametric Surival Regression Model with left- and/or interval-censored covariate. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Hubeaux, S. and Rufibach, K. (2014). **SurvRegCensCov**: Weibull Regression for a Right-Censored Endpoint with a Censored Covariate. Preprint, http://arxiv.org/abs/1402.0432.

Sattar, A., Sinha, S. K. and Morris, N. J. (2012). A Parametric Survival Model When a Covariate is Subject to Left-Censoring. *Biometrics & Biostatistics*, **S3**(2).

Examples

```
## See help file of function "SurvRegCens".
```

---

summary.src *Summarizing Weibull regression with an interval-censored covariate*

Description

summary method for class "src".

Usage

```
## S3 method for class 'src'
summary(object, ...)
```

Arguments

object	An object of class "src", usually a result of a call to SurvRegCens.
...	Further arguments.

Value

The function summary.src returns the estimated parameters, incl. statistical inference, of the Weibull regression, incl. AIC, when calling SurvRegCens.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

References

Hubeaux, S. (2013). Parametric Surival Regression Model with left- and/or interval-censored covariate. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Hubeaux, S. and Rufibach, K. (2014). **SurvRegCensCov**: Weibull Regression for a Right-Censored Endpoint with a Censored Covariate. Preprint, http://arxiv.org/abs/1402.0432.

Sattar, A., Sinha, S. K. and Morris, N. J. (2012). A Parametric Survival Model When a Covariate is Subject to Left-Censoring. *Biometrics & Biostatistics*, **S3**(2).

Examples

```
## See help file of function "SurvRegCens".
```

SurvRegCens *Weibull Survival Regression Model with a censored covariate*

Description

Computes estimators for the shape and scale parameter of the Weibull distribution, as well as for the vector of regression parameters in a parametric survival model with potentially right-censored time-to-event endpoint distributed according to a Weibull distribution. The regression allows for one potentially interval-censored and an arbitrary number of non-censored covariates.

Usage

```
SurvRegCens(formula, data = parent.frame(), Density, initial, conf.level = 0.95,
          intlimit = 10^-10, namCens = "VarCens", trace = 0, reltol = 10^-8)
```

Arguments

formula	A formula expression as for other regression models. The response has to be a survival object for right-censored data, as returned by the Surv function. The censored covariate is equally specified using Surv. See the examples below and the documentation for Surv, lm and formula for details.
data	A data frame in which to interpret the variables named in the formula argument.
Density	Density function of the censored covariate.

initial	Initial values for the parameters to be optimized over, ordered according to Scale parameter, Shape parameter, regression parameters (i.e. β) linked to the non-censored covariates, regression parameter (i.e. β) linked to the censored covariate. A straightforward initial vector is based on ignoring the censoring of the censored covariate and using survreg, see the example below for an illustration.
conf.level	Confidence level of confidence intervals.
intlimit	In computation of integrals, values of the function to be integrated below intlimit are set to 0. This makes integration results more accurate and speeds up integration. If the data is such that the absolute values of the underlying baseline Weibull density are very small, i.e. in the range of intlimit, it is advisable to rescale the time variable, e.g. change the scaling from days to years. A very small value of the estimated λ is indicative of that situation.
namCens	Name of censored covariate, to tidy outputs.
trace	trace argument in optim, indicates whether to show optimization progress.
reltol	reltol argument in optim. By changing this one can modify the relative tolerance in maximization of the likelihood function.

Details

The time-to-event distributed according to a Weibull distribution, i.e. time-to-event $\sim$ Weibull(λ, γ), has conditional density given by,

$$f_{Y_i}(t|\mathbf{x}_i, \boldsymbol{\beta}) = \gamma\lambda t^{\gamma-1} \exp(\mathbf{x}_i\boldsymbol{\beta}) \exp(-\lambda t^{\gamma} \exp(\mathbf{x}_i\boldsymbol{\beta})),$$

conditional hazard function given by,

$$h_i(t|\mathbf{x}_i, \boldsymbol{\beta}) = \lambda\gamma t^{\gamma-1} \exp(\mathbf{x}_i\boldsymbol{\beta}),$$

and conditional survival function given by,

$$S_i(t|\mathbf{x}_i, \boldsymbol{\beta}) = \exp(-\lambda t^{\gamma} \exp(\mathbf{x}_i\boldsymbol{\beta})),$$

where $\mathbf{x}_i$ collects the values of each covariate for observation i and $\boldsymbol{\beta}$ represents the regression parameters.

Value

SurvRegCens returns an object of class "src", a list containing the following components:

coeff	Estimators, confidence intervals, p-values for the for the null hypothesis: {Estimators is equal to 0}, and this for each of the parameters of the Weibull survival regression model.
percent.cens	Percentage of censored observations in the censored covariate.
loglik	Log-likelihood function value at the estimators.
info.converg	Convergence information provided by the function optim.
info.converg.message	Message provided by optim.

The methods print.src, summary.src, coef.src, and logLik.src are used to print or obtain a summary, coefficients, or the value of the log-likelihood at the maximum.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

Kaspar Rufibach, <kaspar.rufibach@gmail.com>
http://www.kasparrufibach.ch

References

Hubeaux, S. (2013). Parametric Surival Regression Model with left- and/or interval-censored covariate. Technical report, Biostatistics Oncology, F. Hoffmann-La Roche Ltd.

Hubeaux, S. and Rufibach, K. (2014). **SurvRegCensCov**: Weibull Regression for a Right-Censored Endpoint with a Censored Covariate. Preprint, http://arxiv.org/abs/1402.0432.

Sattar, A., Sinha, S. K. and Morris, N. J. (2012). A Parametric Survival Model When a Covariate is Subject to Left-Censoring. *Biometrics & Biostatistics*, **S3**(2).

Examples

```
## Not run:
## ------------------------------------------------------------
## 1 censored-covariate and 2 non-censored covariates
## no censoring, to compare result with survival::survreg
## modify prop.cens to introduce left-censoring of covariate
## ------------------------------------------------------------

set.seed(158)
n <- 100
lambda <- exp(-2)
gamma <- 1.5

## vector of regression parameters: the last entry is the one for the censored covariate
beta <- c(0.3, -0.2, 0.25)
true <- c(lambda, gamma, beta)

## non-censored covariates
var1 <- rnorm(n, mean = 4, sd = 0.5)
var2 <- rnorm(n, mean = 4, sd = 0.5)

## Generate censored covariate.
## For generation of Weibull survival times, do not left-censor it yet.
var3 <- rnorm(n, mean = 5, sd = 0.5)

## simulate from a Weibull regression model
time <- TimeSampleWeibull(covariate_noncens = data.frame(var1, var2),
          covariate_cens = var3, lambda = lambda, gamma = gamma, beta = beta)

## left-censor covariate
## prop.cens specifies the proportion of observations that should be left-censored
prop.cens <- 0
LOD <- qnorm(prop.cens, mean = 5, sd = 0.5)
var3.cens <- censorContVar(var3, LLOD = LOD)
```

```
## censor survival time
event <- matrix(1, nrow = n, ncol = 1)
time.cens <- rexp(n, rate = 0.5)
ind.time <- (event >= time.cens)
event[ind.time] <- 0
time[ind.time] <- time.cens[ind.time]

## specify the density for the censored covariate:
## For simplicity, we take here the "true" density we simulate from. In an application,
## you might want to use a density with parameters estimated from the censored covariate,
## e.g. using the function ParamSampleCens. See example in Hubeaux & Rufibach (2014).
DensityCens <- function(value){return(dnorm(value, mean = 5, sd = 0.5))}

## use Weibull regression where each censored covariate value is set
## to LOD ("naive" method)
naive <- survreg(Surv(time, event) ~ var1 + var2 + var3.cens[, 2], dist = "weibull")
initial <- as.vector(ConvertWeibull(naive)$vars[, 1])

## use new method that takes into account the left-censoring of one covariate
data <- data.frame(time, event, var3.cens, var1, var2)
formula <- formula(Surv(time, event) ~  Surv(time = var3.cens[, 1], time2 = var3.cens[, 2],
                      type = "interval2") + var1 + var2)
cens1 <- SurvRegCens(formula = formula, data = data, Density = DensityCens, initial = initial,
                      namCens = "biomarker")
summary(cens1)
coef(cens1)
logLik(cens1)

## compare estimates
tab <- data.frame(cbind(true, initial, cens1$coeff[, 1]))
colnames(tab) <- c("true", "naive", "Weibull MLE")
rownames(tab) <- rownames(cens1$coeff)
tab

## compare confidence intervals
ConvertWeibull(naive)$HR[, 2:3]
cens1$coeff[, 7:8]


## ------------------------------------------------------------
## model without the non-censored covariates
## ------------------------------------------------------------
naive2 <- survreg(Surv(time, event) ~ var3.cens[, 2], dist = "weibull")
initial2 <- as.vector(ConvertWeibull(naive2)$vars[, 1])

## use new method that takes into account the left-censoring of one covariate
formula <- formula(Surv(time, event) ~ Surv(time = var3.cens[, 1], time2 = var3.cens[, 2],
                      type = "interval2"))
cens2 <- SurvRegCens(formula = formula, data = data, Density = DensityCens, initial = initial2,
                      namCens = "biomarker")
summary(cens2)

## compare estimates
```

```
tab <- data.frame(cbind(true[c(1, 2, 5)], initial2, cens2$coeff[, 1]))
colnames(tab) <- c("true", "naive", "Weibull MLE")
rownames(tab) <- rownames(cens2$coeff)
tab

## compare confidence intervals
ConvertWeibull(naive2)$HR[, 2:3]
cens2$coeff[, 7:8]

## End(Not run)
```

`TimeSampleWeibull` *Generate time-to-event data according to a Weibull regression model*

Description

Generates time-to-event data using the transform inverse sampling method, and such that the time-to-event is distributed according to a Weibull distribution induced by censored and/or non-censored covariates. Can be used to set up simulations.

Usage

```
TimeSampleWeibull(covariate_noncens = NULL, covariate_cens, lambda, gamma, beta)
```

Arguments

`covariate_noncens`	Censored covariate vector.
`covariate_cens`	Matrix where each column represents a non-censored covariate.
`lambda`	Scale parameter.
`gamma`	Shape parameter.
`beta`	Regression parameters, ordered as β linked to the non-censored covariates, β linked to the censored covariate.

Note

The use of this function is illustrated in `SurvRegCens`.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

`WeibullDiag` *Diagnostic Plot of Adequacy of Weibull Distribution*

Description

This function constructs a diagnostic plot of the adequacy of the Weibull distribution for survival data with respect to one categorical covariate. If the Weibull distribution fits the data well, then the lines produced should be linear and parallel.

Usage

```
WeibullDiag(formula, data = parent.frame(), labels = names(m$strata))
```

Arguments

`formula`	A formula containing a `Surv` object, should only contain one categorical predictor, or a set of indicators describing only one predictor.
`data`	Data set.
`labels`	A vector containing labels for the plotted lines.

Details

As discussed in Klein and Moeschberger (2003), one method for checking the adequacy of the Weibull model with a categorical covariate is to produce stratified Kaplan-Meier estimates (KM), which can be transformed to estimate the log cumulative hazard for each stratum. Then in a plot of $\log(t)$ versus $\log(-\log(KM))$, the lines should be linear and parallel. This can be seen as the log cumulative hazard for the Weibull distribution is

$$\log H(t) = \log \lambda + \alpha \log t.$$

Value

Produces a plot of log Time vs. log Estimated Cumulative Hazard for each level of the predictor (similarly to what can be obtained using `plot.survfit` and the `fun = "cloglog"` option), as well as a data set containing that information.

Author(s)

Sarah R. Haile, <sarah.haile@gmail.com>

References

Klein, J. and Moeschberger, M. (2003). *Survival analysis: techniques for censored and truncated data*. 2nd edition, Springer.

See Also

Requires **survival**. A similar plot can be produced using `plot.survfit` and the option `fun = "cloglog"`.

Examples

```
data(larynx)
WeibullDiag(Surv(time, death) ~ stage, data = larynx)
```

`WeibullIntegrate` *Function to be integrated in function* `SurvRegCens`

Description

Function to be integrated to compute log-likelihood function for the Weibull survival regression model with a censored covariate.

Usage

```
WeibullIntegrate(x, x_i_noncens = NULL, density, param_y_i,
                 param_delta_i, param_lambda, param_gamma,
                 param_beta, intlimit = 10^-10, ForIntegrate = TRUE)
```

Arguments

`x`	Value of the censored covariate for observation i.
`x_i_noncens`	Vector of values of the non-censored covariates for observation i, i.e. one row of the matrix where each column represents a non-censored covariate.
`density`	Density function of the censored covariate.
`param_y_i`	Value of the time-to-event for observation i.
`param_delta_i`	Censoring indicator of time-to-event for observation i (0: censored, 1: not censored).
`param_lambda`	Scale parameter of the Weibull distribution.
`param_gamma`	Shape parameter of the Weibull distribution.
`param_beta`	Regression parameters (i.e. β): (betaNonCens1, ..., betaNonCens, betaCens)
`intlimit`	In computation of integrals, values of the function to be integrated below `intlimit` are set to 0. This makes integration results more accurate and speeds up integration. If the data is such that the absolute values of the underlying baseline Weibull density are very small, i.e. in the range of `intlimit`, it is advisable to rescale the time variable, e.g. change the scaling from days to years. A very small value of the estimated λ is indicative of that situation.
`ForIntegrate`	`logical` indicating whether the function is to be integrated or not.

Note

Function is not intended to be invoked by the user.

Author(s)

Stanislas Hubeaux, <stanislas.hubeaux@alumni.epfl.ch>

`WeibullReg` *Weibull Regression for Survival Data*

Description

`WeibullReg` performs Weibull regression using the `survreg` function, and transforms the estimates to a more natural parameterization. Additionally, it produces hazard ratios (corresponding to the proportional hazards interpretation), and event time ratios (corresponding to the accelerated failure time interpretation) for all covariates.

Usage

```
WeibullReg(formula, data = parent.frame(), conf.level = 0.95)
```

Arguments

`formula`	A `Surv` formula.
`data`	The dataset containing all variables referenced in `formula`.
`conf.level`	Specifies that $1 - \alpha$ level confidence intervals for the hazard and event time ratios should be produced.

Details

Details regarding the transformations of the parameters and their standard errors can be found in Klein and Moeschberger (2003, Chapter 12). An explanation of event time ratios for the accelerated failure time interpretation of the model can be found in Carroll (2003). A general overview can be found in the `vignette("weibull")` of this package, or in the documentation for `ConvertWeibull`.

Value

`formula`	The formula for the Weibull regression model.
`coef`	The transformed maximum likelihood estimates, with standard errors.
`HR`	The hazard ratios for each of the predictors, with $1-\alpha$ level confidence intervals.
`ETR`	The event time ratios (acceleration factors) for each of the predictors, with $1-\alpha$ level confidence intervals.
`summary`	The summary output from the original `survreg` model.

Author(s)

Sarah R. Haile, <sarah.haile@gmail.com>

References

Carroll, K. (2003). On the use and utility of the Weibull model in the analysis of survival data. *Controlled Clinical Trials*, **24**, 682–701.

Klein, J. and Moeschberger, M. (2003). *Survival analysis: techniques for censored and truncated data*. 2nd edition, Springer.

See Also

Requires the package **survival**. This function depends on ConvertWeibull. See also survreg.

Examples

```
data(larynx)
WR <- WeibullReg(Surv(time, death) ~ factor(stage) + age, data = larynx)
WR
```

Index