Appendix A
Khmaladze's transformation

In the sequel we review, based on Martinussen & Skovgaard (2002), how certain processes may be transformed to Gaussian martingales. The technique is originally developed in Khmaladze (1981), where strict proofs are given. Let us first consider the prototype example of a (time-transformed) Brownian bridge before we turn to the general formulation. Let $B(t) = H(t) - \beta t$, where $H(t)$ is a Gaussian martingale on $[0, 1]$ with variance function $\mathrm{var}H(t) = \int_0^t a(s)^{-1}\, ds$. The quantity β may be stochastic such as $\beta = H(1) - H(0)$ corresponding to the Brownian bridge case. We seek a linear transformation of the process B to a Gaussian martingale. Furthermore we can make the term βt vanish by this transformation. The transformation to the new process $\tilde{B}$ is given as

$$\tilde{B}(t) = B(t) - \int_0^t \left(\int_s^1 a(u)\, du \right)^{-1} \int_s^1 a(u)\, dB(u)\, ds, \qquad \text{(A.1)}$$

which is a Gaussian martingale with the same variance function as H. The idea behind this result of Khmaladze is described below. But first a direct calculation shows that $\tilde{H} = \tilde{B}$ (with $\tilde{H}(t)$ defined by (A.1) replacing B with H), or in other words that the term βt is killed by the transformation. Thus, in particular, the distribution of $\tilde{B}$ is not affected by β being stochastic, even if it depends on the entire process. More generally let $H(t)$ and $B(t)$ be p-vector processes on $[0, \tau]$ so that that $B(t) = H(t) + g(t)\beta$, where $H(t)$ is a Gaussian martingale, $g(t)$ is a (known) $p \times q$-matrix and β is a q-vector that should be thought of as being random so that B is not Gaussian martingale. Let the variance function H be $\mathrm{var}H(t) = \int_0^t a(s)^{-1}\, ds$, which is now a $p \times p$-matrix. Then

$$\begin{aligned}\tilde{H}(t) = H(t) - &\int_0^t a^{-1}(s)a(s)g'(s)\left(\int_s^\tau g'(u)^T a(u)a^{-1}(u)a(u)g'(u)\,du\right)^{-1} \\ &\times \int_s^\tau g'(u)^T a(u)\,dH(u)\,ds \\ = H(t) - &\int_0^t g'(s)\left(\int_s^\tau g'(u)^T a(u)g'(u)\,du\right)^{-1} \\ &\times \int_s^\tau g'(u)^T a(u)\,dH(u)\,ds\end{aligned}$$

is again a Gaussian martingale with the same variance function, and it is directly verified that $\tilde{H}(t) = \tilde{B}(t)$ for all t, so that any component $g(t)\beta$ is killed by conversion from the process $B(t)$ to $\tilde{B}(t)$. We hence also have that

$$\begin{aligned}\tilde{B}(t) = B(t) - &\int_0^t g'(s)\left(\int_s^\tau g'(u)^T a(u)g'(u)\,du\right)^{-1} \\ &\times \int_s^\tau g'(u)^T a(u)\,dB(u)\,ds,\end{aligned} \tag{A.2}$$

is a Gaussian martingale with the same variance function as $H(t)$. The transformation of B to $\tilde{B}$ given by (A.2) is what we usually call *Khmaladze's transformation.*

The two key steps in the derivation of the result are a projection followed by a Doob-Meyer decomposition. First the projection, P_L onto the subspace, L, spanned by the columns of $g(t)$ is

$$(P_L H)(t) = g(t)\left(\int_0^\tau g'(s)^T a(s)g'(s)\,ds\right)^{-1} \int_0^\tau g'(s)^T a(s)\,dH(s),$$

where the inverse variance is used as inner product in the definition of the projection. Application of the orthogonal projection $I - P_L$ to $B(t)$ kills the term $g(t)\beta$ and hence gives the same result as when applied to $H(t)$.

Next we need to adjust $(I - P_L)H(t)$ by subtraction of its compensator given the σ-algebra spanned by the original σ-algebra defining the martingale, $\mathcal{F}_t$ say, and $P_L H$. This is done by calculating the martingale increment by subtraction of the conditional expectation,

$$d\tilde{H}(t) = dH(t) - \operatorname{cov}\left(dH(t), Z(t)\right)\operatorname{var}(Z(t))^{-1}Z(t),$$

where

$$Z(t) = \int_t^\tau g'(s)^T a(s)\,dH(s)$$

is a non-singular representation of the extra information contained in $P_L H$ relative to the σ-algebra $\mathcal{F}_t$. Note that when we subtract the compensator we should in principle start from the process $(I - P_L)H(t)$ but the term $P_L H(t)$ disappears because it is predictable. The expression for $\tilde{H}$ follows directly by calculation of the covariance and variance above.

That the covariance function for $\tilde{H}$ is the same as that for H may be seen by the following calculation. First write

$$\mathrm{cov}\{\tilde{H}(t_1), \tilde{H}(t_2)\} = \int_0^{t_1} \int_0^{t_2} \mathrm{cov}\{d\tilde{H}(t_1), d\tilde{H}(t_2)\}\, ds_2\, ds_1.$$

For $s_1 < s_2$ we rewrite the integrand as

$$\begin{aligned}\mathrm{cov}\{d\tilde{H}(s_1), d\tilde{H}(s_2)\} =& \mathrm{cov}\{dH(s_1), dH(s_2)\} - \mathrm{cov}\{dH(s_1), b(s_2)Z(s_2)\} \\ & - \mathrm{cov}\{b(s_1)Z(s_1), d\tilde{H}(s_2)\},\end{aligned}$$

where $b(s)$ equals the non-random regression coefficient

$$\mathrm{cov}\{dg(s), Z(s)\}\mathrm{var}\{Z(s)\}^{-1}.$$

Using the fact that H has independent increments we see that the second term on the right vanishes because $Z(s_2)$ is a linear function of increments of H over the interval (s_2, τ) which is disjoint from $(s_1, s_1 + ds_1)$. To see that also the third term vanishes note that $Z(s_1) = Z(s_2) + U$ where U is a function of increments over the interval (s_1, s_2). By construction $d\tilde{H}(s_2)$ is independent of the "past" and orthogonalized on $Z(s_2)$ thus completing the argument, which applies similarly to $s_2 < s_1$.

Appendix B
Matrix derivatives

In the following we give some convenient formulae for matrix differentiation. The results are taken from MacRae (1974) where additional results and details can be found.

Consider a $m \times n$-matrix Y, which is a function of the $p \times q$-matrix X. The derivative of matrix Y with respect to X is defined to be an $mp \times nq$-matrix of partial derivatives

$$dY/dX = Y \otimes dX,$$

where $\otimes$ denotes the Kronecher product. Hence

$$dY/dX = \begin{pmatrix} dy_{11}/dX & \cdots & dy_{1n}/dX \\ \vdots & \ddots & \vdots \\ dy_{m1}/dX & \cdots & dy_{mn}/dX \end{pmatrix}.$$

The following results can then be shown.

Theorem B.0.1 *Let Y and Z be matrix functions of X, and let the product YZ be defined. Then*

$$d(YZ)/dX = (dY/dX)(Z \otimes I_q) + (Y \otimes I_p)(dZ/dX). \qquad \text{(Product rule)}$$

Theorem B.0.2 *Let Y be a nonsingular matrix function of X. Then*

$$d(Y^{-1})/dX = -(Y^{-1} \otimes I_p)(dY/dX)(Y^{-1} \otimes I_q). \qquad \text{(Inverse rule)}$$

Appendix C
The Timereg survival package for **R**

This chapter contains a brief description of how to obtain the programs used for the analyses in the book. All programs are available as an add-on-package for the statistical software **R**. The package is available under the general public license (GPL).

The package is available from the Timereg page

```
http:\\biostat.ku.dk\~ts\timereg.html
```

where versions for Linux (Unix) and Windows are available.

After downloading the package and following the instructions given on the homepage you should get a library to use within **R**.

We here give a few extra details in the Linux case. If you do not have super-user permissions you might set up your own local library by the commands

```
R CMD INSTALL timereg --library localdir
```

and then inside **R** write

```
> .libPaths("localdir")
> library(timereg)
This is timereg 0.1-2
```

Manual pages from Timereg package

`aalen` *Fit additive hazards model*

Description

Fits both the additive hazards model of Aalen and the semi-parametric additive hazards model of McKeague and Sasieni. Estimates are unweighted. Time dependent variables and counting process data (multiple events per subject) are possible.

Resampling is used for computing p-values for tests of time-varying effects.

The modeling formula uses the standard survival modeling given in the **survival package.**

Usage

```
aalen(formula,data=sys.parent(),start.time=0,max.time=0,
robust=1,id=NULL,clusters=NULL,residuals=0,n.sim=1000,
weighted.test=0,covariance=0,resample.iid=0)
```

Arguments

`formula`	a formula object with the response on the left of a '~' operator, and the independent terms on the right as regressors. The response must be a survival object as returned by the 'Surv' function.
`data`	a data.frame with the variables.
`start.time`	start of observation period where estimates are computed.
`max.time`	end of observation period where estimates are computed. Estimates thus computed from [start.time, max.time]
`robust`	to compute robust variances and construct processes for resampling. May be set to 0 to save memory.
`id`	For time-varying covariates the variable must associate each record with the id of a subject.
`clusters`	cluster variable for computation of robust variances.
`n.sim`	number of simulations in resampling.

[0]Reproduced with permission from the documentation files in the Timereg package

`weighted.test`	to compute a variance weighted version of the test-processes used for testing time-varying effects.
`residuals`	to returns residuals that can be used for model validation in the function cum.residuals
`covariance`	to compute covariance estimates for nonparametric terms rather than just the variances.
`resample.iid`	to return i.i.d. representation for nonparametric and parametric terms.

Details

The data for a subject is presented as multiple rows or "observations", each of which applies to an interval of observation (start, stop]. For counting process data with the)start,stop] notation is used the 'id' variable is needed to identify the records for each subject. The program assumes that there are no ties, and if such are present random noise is added to break the ties.

Value

returns an object of type "aalen". With the following arguments:

`cum`	cumulative time-varying regression coefficient estimates are computed within the estimation interval.
`var.cum`	the martingale based pointwise variance estimates for cumulatives.
`robvar.cum`	robust pointwise variances estimates for cumulatives.
`gamma`	estimate of parametric components of model.
`var.gamma`	variance for gamma.
`robvar.gamma`	robust variance for gamma.
`residuals`	list with residuals.
`obs.testBeq0`	observed absolute value of supremum of cumulative components scaled with the variance.
`pval.testBeq0`	p-value for covariate effects based on supremum test.
`sim.testBeq0`	resampled supremum values.
`obs.testBeqC`	observed absolute value of supremum of difference between observed cumulative process and estimate under null of constant effect.
`pval.testBeqC`	p-value based on resampling.
`sim.testBeqC`	resampled supremum values.

`obs.testBeqC.is`
: observed integrated squared differences between observed cumulative and estimate under null of constant effect.

`pval.testBeqC.is`
: p-value based on resampling.

`sim.testBeqC.is`
: resampled supremum values.

`conf.band` resampling based constant to construct robust 95% uniform confidence bands.

`test.procBeqC` observed test-process of difference between observed cumulative process and estimate under null of constant effect over time.

`sim.test.procBeqC`
: list of 50 random realizations of test-processes under null based on resampling.

`covariance` covariances for nonparametric terms of model.

`B.iid` Resample processes for nonparametric terms of model.

`gamma.iid` Resample processes for parametric terms of model.

Author(s)

Thomas Scheike

References

Martinussen and Scheike, Dynamic Regression Models for Survival Data, Springer (2006).

Examples

```
library(survival)
data(sTRACE)
# Fits Aalen model
out<-aalen(Surv(time,status==9)~age+sex+diabetes+chf+vf,
sTRACE,max.time=7,n.sim=500)

summary(out)
par(mfrow=c(2,3))
plot(out)

# Fits semi-parametric additive hazards model
out<-aalen(Surv(time,status==9)~const(age)+const(sex)+const(diabetes)+chf
+vf,sTRACE,max.time=7,n.sim=500)

summary(out)
par(mfrow=c(2,3))
plot(out)
```

`cd4` *The multicenter AIDS cohort study*

Description

Format

This data frame contains the following columns:

obs a numeric vector. Number of observations.

id a numeric vector. Id of subject.

visit a numeric vector. Timings of the visits in years.

smoke a numeric vector code. 0: non-smoker, 1: smoker.

age a numeric vector. Age of the patient at the start of the trial.

cd4 a numeric vector. CD4 percentage at the current visit.

cd4.prev a numeric vector. CD4 level at the preceding visit.

precd4 a numeric vector. Post-infection CD4 percentage.

lt a numeric vector. Gives the starting time for the time-intervals.

rt a numeric vector. Gives the stopping time for the time-interval.

Source

MACS Public Use Data Set Release PO4 (1984-1991). See reference.

References

Kaslow et al. (1987), The multicenter AIDS cohort study: rational, organization and selected characteristics of the participants. Am. J. Epidemiology 126, 310–318.

Examples

```
data(cd4)
names(cd4)
```

`const` *Identifies parametric terms of model*

Description

Specifies which of the regressors that have constant effect.

Author(s)

Thomas Scheike

`cox` *Identifies proportional excess terms of model*

Description

Specifies which of the regressors that lead to proportional excess hazard

Author(s)

Thomas Scheike

`cox.aalen` *Fit Cox-Aalen survival model*

Description

Fits an Cox-Aalen survival model. Time dependent variables and counting process data (multiple events per subject) are possible.

Resampling is used for computing p-values for tests of time-varying effects. Test for proportionality is considered by considering the score processes for the proportional effects of model.

The modeling formula uses the standard survival modeling given in the **survival package.**

Usage

```
cox.aalen(formula=formula(data),data=sys.parent(),beta=0,
Nit=10,detail=0,start.time=0,max.time=0,id=NULL,
clusters=NULL,n.sim=500,residuals=0,robust=1,
weighted.test=0,covariance=0,resample.iid=0,weights=NULL)
```

Arguments

`formula`	a formula object with the response on the left of a '~' operator, and the independent terms on the right as regressors. The response must be a survival object as returned by the 'Surv' function.
`data`	a data.frame with the variables.
`start.time`	start of observation period where estimates are computed.
`max.time`	end of observation period where estimates are computed. Estimates thus computed from [start.time, max.time]
`robust`	to compute robust variances and construct processes for resampling. May be set to 0 to save memory.
`id`	For time-varying covariates the variable must associate each record with the id of a subject.
`clusters`	cluster variable for computation of robust variances.
`n.sim`	number of simulations in resampling.
`weighted.test`	to compute a variance weighted version of the test-processes used for testing time-varying effects.
`residuals`	to returns residuals that can be used for model validation in the function cum.residuals
`covariance`	to compute covariance estimates for nonparametric terms rather than just the variances.
`resample.iid`	to return i.i.d. representation for nonparametric and parametric terms.
`beta`	starting value for relative risk estimates
`Nit`	number of iterations for Newton-Raphson algorithm.
`detail`	if 0 no details is printed during iterations, if 1 details are given.
`weights`	weights for weighted analysis.

Details

The data for a subject is presented as multiple rows or "observations", each of which applies to an interval of observation (start, stop]. For counting process data with the)start,stop] notation is used the 'id' variable is needed to identify the records for each subject. The program assumes that there are no ties, and if such are present random noise is added to break the ties.

Value

returns an object of type ”cox.aalen”. With the following arguments:

`cum`	cumulative time-varying regression coefficient estimates are computed within the estimation interval.
`var.cum`	the martingale based pointwise variance estimates.
`robvar.cum`	robust pointwise variances estimates.
`gamma`	estimate of parametric components of model.
`var.gamma`	variance for gamma.
`robvar.gamma`	robust variance for gamma.
`residuals`	list with residuals.
`obs.testBeq0`	observed absolute value of supremum of cumulative components scaled with the variance.
`pval.testBeq0`	p-value for covariate effects based on supremum test.
`sim.testBeq0`	resampled supremum values.
`obs.testBeqC`	observed absolute value of supremum of difference between observed cumulative process and estimate under null of constant effect.
`pval.testBeqC`	p-value based on resampling.
`sim.testBeqC`	resampled supremum values.
`obs.testBeqC.is`	observed integrated squared differences between observed cumulative and estimate under null of constant effect.
`pval.testBeqC.is`	p-value based on resampling.
`sim.testBeqC.is`	resampled supremum values.
`conf.band`	resampling based constant to construct robust 95% uniform confidence bands.
`test.procBeqC`	observed test-process of difference between observed cumulative process and estimate under null of constant effect over time.
`sim.test.procBeqC`	list of 50 random realizations of test-processes under null based on resampling.
`covariance`	covariances for nonparametric terms of model.
`B.iid`	Resample processes for nonparametric terms of model.
`gamma.iid`	Resample processes for parametric terms of model.

`loglike` approximate log-likelihood for model, similar to Cox's partial likelihood.

`D2linv` inverse of the derivative of the score function.

`score` value of score for final estimates.

`test.procProp` observed score process for proportional part of model.

`pval.Prop` p-value based on resampling.

`sim.supProp` re-sampled absolute supremum values.

`sim.test.procProp` list of 50 random realizations of test-processes for proportionality under the model based on resampling.

Author(s)

Thomas Scheike

References

Martinussen and Scheike, Dynamic Regression Models for Survival Data, Springer (2006).

Examples

```
library(survival)
data(sTRACE)
# Fits Cox model
out<-cox.aalen(Surv(time,status==9)~prop(age)+prop(sex)+
prop(vf)+prop(chf)+prop(diabetes),sTRACE,max.time=7,n.sim=500)

# makes Lin, Wei, Ying test for proportionality
summary(out)
par(mfrow=c(2,3))
plot(out,score=1)

# Fits Cox-Aalen model
out<-cox.aalen(Surv(time,status==9)~prop(age)+prop(sex)+
vf+chf+prop(diabetes),sTRACE,max.time=7,n.sim=500)

# plots the additive part of the model. To obtain more sensible
# plots center covariates in proportional part of model
summary(out)
par(mfrow=c(2,3))
plot(out)
```

`csl` *CSL liver cirrhosis data*

Description

Format

This data frame contains the following columns:

id a numeric vector. Id of subject.

time a numeric vector. Time of measurement.

prot a numeric vector. Prothrombin level at measurement time.

dc a numeric vector code. 0: censored observation, 1: died at eventT.

eventT a numeric vector. Time of event (death).

treat a numeric vector code. 0: active treatment of prednisone, 1: placebo treatment.

sex a numeric vector code. 0: female, 1: male.

age a numeric vector. Age of subject at inclusion time subtracted 60.

prot.base a numeric vector. Prothrombin base level before entering the study.

prot.prev a numeric vector. Level of prothrombin at previous measurement time.

lt a numeric vector. Gives the starting time for the time-intervals.

rt a numeric vector. Gives the stopping time for the time-intervals.

Source

P.K. Andersen

References

Schlichting, P., Christensen, E., Andersen, P., Fauerholds, L., Juhl, E., Poulsen, H. and Tygstrup, N. (1983), The Copenhagen Study Group for Liver Diseases, Hepatology 3, 889–895

Examples

```
data(csl)
names(csl)
```

`cum.residuals` *Model validation based on cumulative residuals*

Description

Computes cumulative residuals and approximative p-values based on resampling techniques.

Usage

```
cum.residuals(object,data=sys.parent(),modelmatrix=0,
cum.resid=0,n.sim=500,weighted.test=1,start.design=1)
```

Arguments

`object`	an object of class 'aalen', 'timecox', 'cox.aalen' where the residuals are returned ('residuals=1')
`data`	data frame based on which residuals are computed.
`modelmatrix`	specifies a grouping of the data that is used for cumulating residuals. Must have same size as data and be ordered in the same way.
`n.sim`	number of simulations in resampling.
`weighted.test`	to compute a variance weighted version of the test-processes used for testing constant effects of covariates.
`cum.resid`	to compute residuals versus each of the continuous covariates in the model.
`start.design`	if '1' the groupings specified in modelmatrix changes over time, i.e. in the case with time-dependent covariates.

Value

returns an object of type "cum.residuals" with the following arguments:

`cum`	cumulative residuals versus time for the groups specified by modelmatrix.
`var.cum`	the martingale based pointwise variance estimates.
`robvar.cum`	robust pointwise variances estimates of cumulatives.
`obs.testBeq0`	observed absolute value of supremum of cumulative components scaled with the variance.
`pval.testBeq0`	p-value covariate effects based on supremum test.
`sim.testBeq0`	resampled supremum value.

`conf.band`	resampling based constant to construct robust 95% uniform confidence bands for cumulative residuals.
`obs.test`	absolute value of supremum of observed test-process.
`pval.test`	p-value for supremum test statistic.
`sim.test`	resampled absolute value of supremum cumulative residuals.
`proc.cumz`	observed cumulative residuals versus all continuous covariates of model.
`sim.test.proccumz`	list of 50 random realizations of test-processes under model for all continuous covariates.

Author(s)

Thomas Scheike

References

Martinussen and Scheike, Dynamic Regression Models for Survival Data, Springer (2006).

Examples

```
library(survival)
data(sTRACE)
# Fits Aalen model and returns residuals
fit<-aalen(Surv(time,status==9)~age+sex+diabetes+chf+vf,
sTRACE,max.time=7,n.sim=0,residuals=1)

# constructs and simulates cumulative residuals versus age groups
fit.mg<-cum.residuals(fit,sTRACE,
                      model.matrix(~-1+factor(cut(age,4)),sTRACE))

par(mfrow=c(1,4))
# cumulative residuals with confidence intervals
plot(fit.mg);
# cumulative residuals versus processes under model
plot(fit.mg,score=1);
summary(fit.mg)

# cumulative residuals vs. covariates Lin, Wei, Ying style
fit.mg<-cum.residuals(fit,sTRACE,cum.resid=1)

par(mfrow=c(2,4))
plot(fit.mg,score=2)
summary(fit.mg)
```

`dynreg` *Fit time-varying regression model*

Description

Fits time-varying regression model with partly parametric components. Time-dependent variables for longitudinal data. The model assumes that the mean of the observed responses given covariates is a linear time-varying regression model :

$$E(Z_{ij}|X_{ij}(t)) = \beta^T(t)X_{ij}^1(t) + \gamma^T X_{ij}^2(t)$$

where Z_{ij} is the j'th measurement at time t for the i'th subject with covariates X_{ij}^1 and X_{ij}^2. Resampling is used for computing p-values for tests of time-varying effects.

Usage

```
dynreg(formula,data=sys.parent(),aalenmod,bandwidth=0.5,
id=NULL,bhat=NULL,start.time=0,max.time=0,n.sim=500,
residuals=0,meansub=1,weighted.test=0)
```

Arguments

`formula`	a formula object with the response on the left of a '~' operator, and the independent terms on the right as regressors.
`data`	a data.frame with the variables.
`start.time`	start of observation period where estimates are computed.
`max.time`	end of observation period where estimates are computed. Estimates thus computed from [start.time, max.time]
`id`	For time-varying covariates the variable must associate each record with the id of a subject.
`n.sim`	number of simulations in resampling.
`weighted.test`	to compute a variance weighted version of the test-processes used for testing time-varying effects.
`residuals`	to returns residuals that can be used for model validation in the function 'cum.residuals'.
`aalenmod`	Aalen model for measurement times. Specified as a survival model (see aalen function).

`bandwidth` bandwidth for local iterations. Default is 50% of the range of the considered observation period.

`bhat` initial value for estimates. If NULL local linear estimate is computed.

`meansub` if '1' then the mean of the responses is subtracted before the estimation is carried out.

Details

The data for a subject is presented as multiple rows or "observations", each of which applies to an interval of observation (start, stop]. For counting process data with the)start,stop] notation is used the 'id' variable is needed to identify the records for each subject. The program assumes that there are no ties, and if such are present random noise is added to break the ties.

Value

returns an object of type "dynreg". With the following arguments:

`cum` the cumulative regression coefficients. This is the efficient estimator based on an initial smoother obtained by local linear regression :

$$\hat{B}(t) = \int_0^t \tilde{\beta}(s)ds +$$

$$\int_0^t X^-(Diag(z) - Diag(X^T(s)\tilde{\beta}(s)))dp(ds \times dz),$$

where $\tilde{\beta}(t)$ is an initial estimate either provided or computed by local linear regression. To plot this estimate use type="eff.smooth" in the plot() command.

`var.cum` the martingale based pointwise variance estimates.

`robvar.cum` robust pointwise variances estimates.

`gamma` estimate of semi-parametric components of model.

`var.gamma` variance for gamma.

`robvar.gamma` robust variance for gamma.

`cum0` simple estimate of cumulative regression coefficients that does not use use an initial smoothing based estimate

$$\hat{B}_0(t) = \int_0^t X^- Diag(z)dp(ds \times dz).$$

To plot this estimate use type="0.mpp" in the plot() command.

`var.cum0` the martingale based pointwise variance estimates of cum0.

`cum.ms` estimate of cumulative regression coefficients based on initial smoother (but robust to this estimator).

$$\hat{B}_{ms}(t) = \int_0^t X^-(Diag(z) - f(s))dp(ds \times dz),$$

where f is chosen as the matrix

$$f(s) = Diag(X^T(s)\tilde{\beta}(s))(I - X_\alpha(s)X_\alpha^-(s)),$$

where X_α is the design for the sampling intensities.

This is also an efficient estimator when the initial estimator is consistent for $\beta(t)$ and then asymptotically equivalent to cum, but small sample properties appear inferior. Its variance is estimated by var.cum.

To plot this estimate use type="ms.mpp" in the plot() command.

`cum.ly` estimator where local averages are subtracted. Special case of cum.ms.

To plot this estimate use type="ly.mpp" in plot.

`var.cum.ly` the martingale based pointwise variance estimates.

`gamma0` estimate of parametric component of model.

`var.gamma0` estimate of variance of parametric component of model.

`gamma.ly` estimate of parametric components of model.

`var.gamma.ly` estimate of variance of parametric component of model.

`gamma.ms` estimate of variance of parametric component of model.

`var.gamma.ms` estimate of variance of parametric component of model.

`residuals` list of residuals.

`obs.testBeq0` observed absolute value of supremum of cumulative components scaled with the variance.

`pval.testBeq0` p-value for covariate effects based on supremum test.

`sim.testBeq0` resampled supremum values.

`obs.testBeqC` observed absolute value of supremum of difference between observed cumulative process and estimate under null of constant effect.

`pval.testBeqC` p-value based on resampling.

`sim.testBeqC` resampled supremum values.

`obs.testBeqC.is`
: observed integrated squared differences between observed cumulative and estimate under null of constant effect.

`pval.testBeqC.is`
: p-value based on resampling.

`sim.testBeqC.is`
: resampled supremum values.

`conf.band`
: resampling based constant to construct robust 95% uniform confidence bands.

`test.procBeqC`
: observed test-process of difference between observed cumulative process and estimate under null of constant effect.

`sim.test.procBeqC`
: list of 50 random realizations of test-processes under null based on resampling.

`covariance`
: covariances for nonparametric terms of model.

Author(s)

Thomas Scheike

References

Martinussen and Scheike, Dynamic Regression Models for Survival Data, Springer (2006).

Examples

```
library(survival)
data(csl)
indi.m<-rep(1,length(csl$lt))

# Fits time-varying regression model on time-range from 0 to 3 years.
out<-dynreg(prot~treat+prot.prev+sex+age,csl,
Surv(lt,rt,indi.m)~+1,start.time=0,max.time=3,id=csl$id,
n.sim=500,bandwidth=0.3,meansub=0)
summary(out)
par(mfrow=c(2,3))
plot(out)

# Fits time-varying semi-parametric regression model.
outS<-dynreg(prot~treat+const(prot.prev)+const(sex)+const(age),csl,
Surv(lt,rt,indi.m)~+1,start.time=0,max.time=3,id=csl$id,
n.sim=500,bandwidth=0.3,meansub=0)
summary(outS)
```

`melanoma` *The Melanoma Survival Data*

Description

The melanoma data frame has 205 rows and 7 columns. It contains data relating to survival of patients after operation for malignant melanoma collected at Odense University Hospital by K.T. Drzewiecki.

Format

This data frame contains the following columns:

no a numeric vector. Patient code.

status a numeric vector code. Survival status. 1: dead from melanoma, 2: alive, 3: dead from other cause.

days a numeric vector. Survival time.

ulc a numeric vector code. Ulceration, 1: present, 0: absent.

thick a numeric vector. Tumor thickness (1/100 mm).

sex a numeric vector code. 0: female, 1: male.

Source

Andersen, P.K., Borgan Ø., Gill R.D., Keiding N. (1993), *Statistical Models Based on Counting Processes*, Springer-Verlag.

Drzewiecki, K.T., Ladefoged, C., and Christensen, H.E. (1980), Biopsy and prognosis for cutaneous malignant melanoma in clinical stage I. Scand. J. Plast. Reconstru. Surg. 14, 141-144.

Examples

```
data(melanoma)
names(melanoma)
```

`mela.pop` *Melanoma data and Danish population mortality by age and sex*

Description

Melanoma data with Danish population mortality rates by age and sex.

Format

This data frame contains the following columns:

id a numeric vector. Gives patient id.

sex a numeric vector. Gives sex of patient.

start a numeric vector. Gives the starting time for the time-interval for which the covariate rate is representative.

stop a numeric vector. Gives the stopping time for the time-interval for which the covariate rate is representative.

status a numeric vector code. Survival status. 1: dead from melanoma, 0: alive or dead from other cause.

age a numeric vector. Gives the age of the patient at removal of tumor.

rate a numeric vector. Gives the population mortality for the given sex and age. Based on Table A.2 in Andersen et al. (1993).

Source

Andersen, P.K., Borgan Ø, Gill R.D., Keiding N. (1993), *Statistical Models Based on Counting Processes*, Springer-Verlag.

Examples

```
data(mela.pop)
names(mela.pop)
```

`plot.aalen` *Plots estimates and test-processes*

Description

This function plots the non-parametric cumulative estimates for the additive risk model or the test-processes for the hypothesis of time-varying effects with re-sampled processes under the null.

Usage

```
plot.aalen(object,pointwise.ci=1,hw.ci=0,sim.ci=0,robust=0,
specific.comps=FALSE,level=0.05, start.time=0,stop.time=0,
add.to.plot=FALSE,mains=TRUE,xlab="Time",
ylab="Cumulative coefficients",score=FALSE)
```

Arguments

`object`	the output from the "aalen" function.
`pointwise.ci`	if >1 pointwise confidence intervals are plotted with lty=pointwise.ci
`hw.ci`	if >1 Hall-Wellner confidence bands are plotted with lty=hw.ci. Only 0.95 % bands can be constructed.
`sim.ci`	if >1 simulation based confidence bands are plotted with lty=sim.ci. These confidence bands are robust to non-martingale behaviour.
`robust`	robust standard errors are used to estimate standard error of estimate, otherwise martingale based standard errors are used.
`specific.comps`	all components of the model is plotted by default, but a list of components may be specified, for example first and third "c(1,3)".
`level`	gives the significance level.
`start.time`	start of observation period where estimates are plotted.
`stop.time`	end of period where estimates are plotted. Estimates thus plotted from [start.time, max.time].
`add.to.plot`	to add to an already existing plot.
`mains`	add names of covariates as titles to plots.
`xlab`	label for x-axis.

`ylab`	label for y-axis.
`score`	to plot test processes for test of time-varying effects along with 50 random realization under the null-hypothesis.

Author(s)

Thomas Scheike

References

Martinussen and Scheike

Examples

```
library(survival)
data(sTRACE)
# Fits Aalen model
out<-aalen(Surv(time,status==9)~age+sex+diabetes+chf+vf,
sTRACE,max.time=7,n.sim=500)

par(mfrow=c(2,3))
# plots estimates
plot(out)
# plots tests-processes for time-varying effects
plot(out,score=TRUE)
```

`plot.cum.residuals`

Plots cumulative residuals

Description

This function plots the output from the cumulative residuals function ”cum.residuals”. The cumulative residuals are compared with the performance of similar processes under the model.

Usage

```
plot.cum.residuals(object,pointwise.ci=1,hw.ci=0,sim.ci=0,
robust=1, specific.comps=FALSE,level=0.05,start.time=0,
stop.time=0,add.to.plot=FALSE,mains=TRUE,xlab="Time",
ylab ="Cumulative Residuals",ylim=NULL,score=0)
```

Arguments

`object`	the output from the ”cum.residuals” function.
`pointwise.ci`	if >1 pointwise confidence intervals are plotted with lty=pointwise.ci

`hw.ci`	if >1 Hall-Wellner confidence bands are plotted with lty=hw.ci. Only 95% bands can be constructed.
`sim.ci`	if >1 simulation based confidence bands are plotted with lty=sim.ci. These confidence bands are robust to non-martingale behaviour.
`robust`	if "1" robust standard errors are used to estimate standard error of estimate, otherwise martingale based estimate are used.
`specific.comps`	all components of the model is plotted by default, but a list of components may be specified, for example first and third "c(1,3)".
`level`	gives the significance level. Default is 0.05.
`start.time`	start of observation period where estimates are plotted. Default is 0.
`stop.time`	end of period where estimates are plotted. Estimates thus plotted from [start.time, max.time].
`add.to.plot`	to add to an already existing plot. Default is "FALSE".
`mains`	add names of covariates as titles to plots.
`xlab`	label for x-axis. Default is "Time".
`ylab`	label for y-axis. Default is "Cumulative Residuals".
`ylim`	limits for y-axis.
`score`	if '0' plots related to modelmatrix are specified, thus resulting in grouped residuals, if '1' plots for modelmatrix but with random realizations under model, if '2' plots residuals versus continuous covariates of model with random realizations under the model.

Author(s)

Thomas Scheike

References

Martinussen and Scheike, Dynamic Regression Models for Survival Data, Springer (2006).

Examples

```
library(survival)
data(sTRACE)
# Fits Aalen model and returns residuals
out<-aalen(Surv(time,status==9)~age+sex+diabetes+chf+vf,
sTRACE,max.time=7,n.sim=0,residuals=1)
```

```
# constructs and simulates cumulative residuals versus age groups
out.mg<-cum.residuals(out,sTRACE,
                      model.matrix(~-1+factor(cut(age,4)),sTRACE))

par(mfrow=c(1,4))
# cumulative residuals with pointwise confidence intervals
plot(out.mg);
# cumulative residuals versus processes under model
plot(out.mg,score=1);

# cumulative residuals against covariates Lin, Wei, Ying style
out.mg<-cum.residuals(out,sTRACE,cum.resid=1)
par(mfrow=c(2,4))
plot(out.mg,score=2)
```

`plot.dynreg` *Plots estimates and test-processes*

Description

This function plots the non-parametric cumulative estimates for the additive risk model or the test-processes for the hypothesis of constant effects with re-sampled processes under the null.

Usage

```
plot.dynreg(object,type="eff.smooth",pointwise.ci=1,hw.ci=0,
sim.ci=0,robust=0, specific.comps=FALSE,level=0.05,
start.time=0,stop.time=0,add.to.plot=FALSE,mains=TRUE,
xlab="Time",ylab="Cumulative coefficients",score=FALSE)
```

Arguments

`object`	the output from the "dynreg" function.
`type`	the estimator plotted. Choices "eff.smooth", "ms.mpp", "0.mpp" and "ly.mpp". See the dynreg function for more on this.
`pointwise.ci`	if >1 pointwise confidence intervals are plotted with lty=pointwise.ci
`hw.ci`	if >1 Hall-Wellner confidence bands are plotted with lty=hw.ci. Only 0.95 % bands can be constructed.
`sim.ci`	if >1 simulation based confidence bands are plotted with lty=sim.ci. These confidence bands are robust to non-martingale behavior.

`robust` robust standard errors are used to estimate standard error of estimate, otherwise martingale based estimate are used.

`specific.comps` all components of the model is plotted by default, but a list of components may be specified, for example first and third "c(1,3)".

`level` gives the significance level.

`start.time` start of observation period where estimates are plotted.

`stop.time` end of period where estimates are plotted. Estimates thus plotted from [start.time, max.time].

`add.to.plot` to add to an already existing plot.

`mains` add names of covariates as titles to plots.

`xlab` label for x-axis.

`ylab` label for y-axis.

`score` to plot test processes for test of time-varying effects along with 50 random realization under the null-hypothesis.

Author(s)

Thomas Scheike

References

Martinussen and Scheike, Dynamic Regression Models for Survival Data, Springer (2006).

Examples

```
library(survival)
data(csl)
indi.m<-rep(1,length(csl$lt))

# Fits time-varying regression model on time-range from 0 to 3 years.
out<-dynreg(prot~treat+prot.prev+sex+age,csl,
Surv(lt,rt,indi.m)~+1,start.time=0,max.time=3,id=csl$id,
n.sim=500,bandwidth=0.3,meansub=0)

par(mfrow=c(2,3))
# plots estimates
plot(out)
# plots tests-processes for time-varying effects
plot(out,score=TRUE)
```

`print.aalen` *Prints call*

Description

Prints call for object. Lists nonparametric and parametric terms of model

Usage

```
print.aalen(object)
```

Arguments

`object` an aalen object

Author(s)

Thomas Scheike

`prop` *Identifies the multiplicative terms in Cox-Aalen model and proportional excess risk model*

Description

Specifies which of the regressors that belong to the multiplicative part of the Cox-Aalen model or the proportional excess risk model.

Usage

```
see cox.aalen or prop.excess
```

Author(s)

Thomas Scheike

prop.excess	*Fits Proportional excess hazards model*

Description

Fits proportional excess hazards model.

The models are written using the survival modeling given in the survival package.

Usage

```
prop.excess(formula=formula(data),data=sys.parent(),
excess=1,tol=0.0001,max.time=0,n.sim=1000,alpha=1,frac=1)
```

Arguments

`formula`	a formula object, with the response on the left of a '~' operator, and the terms on the right. The response must be a survival object as returned by the 'Surv' function.
`data`	a data.frame with the variables.
`excess`	specifies for which of the subjects the excess term is present. Default is that the term is present for all subjects.
`tol`	tolerance for numerical procedure.
`max.time`	stopping considered time-period if different from 0. Estimates thus computed from [0,max.time] if max.time>0.
`n.sim`	number of simulations in re-sampling.
`alpha`	tuning parameter in Newton-Raphson procedure. Value smaller than one may give more stable convergence.
`frac`	number between 0 and 1. Is used in supremum test where observed jump times t1, ..., tk is replaced by t1, ..., tl with l=round(frac*k).

Details

The program assumes that there are no ties, and if such are present random noise is added to break the ties.

Value

Returns an object of type "prop.excess". With the following arguments:

`cum`	estimated cumulative regression functions. First column contains the jump times, then follows the estimated components of additive part of model and finally the excess cumulative baseline.
`var.cum`	robust pointwise variance estimates for estimated cumulatives.
`gamma`	estimate of parametric components of model.
`var.gamma`	robust variance estimate for gamma.
`pval`	p-value of Kolmogorov-Smirnov test (variance weighted) for excess baseline and Aalen terms, H: B(t)=0.
`pval.HW`	p-value of supremum test (corresponding to Hall-Wellner band) for excess baseline and Aalen terms, H: B(t)=0. Reported in summary.
`pval.CM`	p-value of Cramer von Mises test for excess baseline and Aalen terms, H: B(t)=0.
`quant`	95 percent quantile in distribution of resampled Kolmogorov-Smirnov test statistics for excess baseline and Aalen terms. Used to construct 95 percent simulation band.
`quant95HW`	95 percent quantile in distribution of resampled supremum test statistics corresponding to Hall-Wellner band for excess baseline and Aalen terms. Used to construct 95 percent Hall-Wellner band.
`simScoreProp`	observed scoreprocess and 50 resampled scoreprocesses (under model). List with 51 elements.

Author(s)

Torben Martinussen

References

Martinussen and Scheike, Dynamic Regression Models for Survival Data, Springer (2006).

Examples

```
library(survival)
data(melanoma)
attach(melanoma)
lt<-log(thick)          # log-thickness
```

```
excess<-(thick>=210)    # excess risk for thick tumors

# Fits Proportional Excess hazards model
fit<-prop.excess(Surv(days/365,status==1)~sex+ulc+cox(sex)+cox(ulc)
                 +cox(lt),excess=excess,n.sim=2000)
summary(fit)
par(mfrow=c(2,3))
plot(fit)
```

`prop.odds` *Fit Semiparametric Proportional Odds Model*

Description

Fits a semiparametric proportional odds model:

$$logit(1 - S_Z(t)) = log(G(t)) + \beta^T Z$$

where G(t) is increasing but otherwise unspecified. Model is fitted by maximizing the modified partial likelihood. A goodness-of-fit test by considering the score functions is also computed by resampling methods.

The modeling formula uses the standard survival modeling given in the `survival package`.

Usage

```
prop.odds(formula,data=sys.parent(),beta=0,Nit=10,
detail=0,start.time=0,max.time=0,id=NULL,n.sim=500,
weighted.test=0,profile=1,sym=0)
```

Arguments

`formula`	a formula object, with the response on the left of a ' ' operator, and the terms on the right. The response must be a survival object as returned by the 'Surv' function.
`data`	a data.frame with the variables.
`start.time`	start of observation period where estimates are computed.
`max.time`	end of observation period where estimates are computed. Estimates thus computed from [start.time, max.time]. This is very useful to obtain stable estimates, especially for the baseline.
`id`	For time-varying covariates the variable must associate each record with the id of a subject.

`n.sim`	number of simulations in resampling.
`weighted.test`	to compute a variance weighted version of the test-processes used for testing time-varying effects.
`beta`	starting value for relative risk estimates
`Nit`	number of iterations for Newton-Raphson algorithm.
`detail`	if 0 no details is printed during iterations, if 1 details are given.
`profile`	if profile is 1 then modified partial likelihood is used, profile=0 fits by simple estimating equation. The modified partial likelihood is recommended.
`sym`	to use symmetrized second derivative in the case of the estimating equation approach (profile=0). This may improve the numerical performance.

Details

The data for a subject is presented as multiple rows or "observations", each of which applies to an interval of observation (start, stop]. The program essentially assumes no ties, and if such are present a little random noise is added to break the ties.

Value

returns an object of type 'cox.aalen'. With the following arguments:

`cum`	cumulative time-varying regression coefficient estimates are computed within the estimation interval.
`var.cum`	the martingale based pointwise variance estimates.
`robvar.cum`	robust pointwise variances estimates.
`gamma`	estimate of proportional odds parameters of model.
`var.gamma`	variance for gamma.
`robvar.gamma`	robust variance for gamma.
`residuals`	list with residuals.
`obs.testBeq0`	observed absolute value of supremum of cumulative components scaled with the variance.
`pval.testBeq0`	p-value for covariate effects based on supremum test.
`sim.testBeq0`	resampled supremum values.
`obs.testBeqC`	observed absolute value of supremum of difference between observed cumulative process and estimate under null of constant effect.
`pval.testBeqC`	p-value based on resampling.

`sim.testBeqC` resampled supremum values.

`obs.testBeqC.is`
: observed integrated squared differences between observed cumulative and estimate under null of constant effect.

`pval.testBeqC.is`
: p-value based on resampling.

`sim.testBeqC.is`
: resampled supremum values.

`conf.band` resampling based constant to construct robust 95% uniform confidence bands.

`test.procBeqC` observed test-process of difference between observed cumulative process and estimate under null of constant effect over time.

`loglike` modified partial likelihood, pseudo profile likelihood for regression parameters.

`D2linv` inverse of the derivative of the score function.

`score` value of score for final estimates.

`test.procProp` observed score process for proportional odds regression effects.

`pval.Prop` p-value based on resampling.

`sim.supProp` re-sampled supremum values.

`sim.test.procProp`
: list of 50 random realizations of test-processes for constant proportional odds under the model based on resampling.

Author(s)

Thomas Scheike

References

Martinussen and Scheike, Dynamic Regression Models for Survival Data, Springer (2006).

Examples

```
library(survival)
data(sTRACE)
# Fits Proportional odds model
out<-prop.odds(Surv(time,status==9)~age+diabetes+chf+vf+sex,
sTRACE,max.time=7,n.sim=500)
```

```
summary(out)

par(mfrow=c(2,3))
plot(out,sim.ci=2)
plot(out,score=1)
```

`summary.aalen` *Prints summary statistics*

Description

Computes p-values for test of significance for nonparametric terms of model, p-values for test of constant effects based on both supremum and integrated squared difference.

Returns parameter estimates and their standard errors.

Usage

```
summary.aalen(aalen.object,digits=3)
```

Arguments

`aalen.object`	an aalen object.
`digits`	number of digits in printouts.

Author(s)

Thomas Scheike

Examples

```
library(survival)
data(sTRACE)
# Fits Aalen model
out<-aalen(Surv(time,status==9)~age+sex+diabetes+chf+vf,
sTRACE,max.time=7,n.sim=500)

summary(out)
```

`summary.cum.residuals` *Prints summary statistics for goodness-of-fit tests based on cumulative residuals*

Description

Computes p-values for extreme behaviour relative to the model of various cumulative residual processes.

Usage

```
summary.cum.residuals(cum.residuals.object,digits=3
```

Arguments

`cum.resids.object`
: an cum.residuals object.

`digits`
: number of digits in printouts.

Author(s)

Thomas Scheike

Examples

```
library(survival)
data(sTRACE)
# Fits Aalen model and returns residuals
out<-aalen(Surv(time,status==9)~age+sex+diabetes+chf+vf,
sTRACE,max.time=7,n.sim=0,residuals=1)

# constructs and simulates cumulative residuals versus age groups
# and versus covariates of model
out.mg<-cum.residuals(out,sTRACE,
modelmatrix=model.matrix(~-1+factor(cut(age,4)),sTRACE),cum.resid=1)

summary(out.mg)
```

`timecox` *Fit Cox model with partly time-varying effects.*

Description

Fits proportional hazards model with some effects time-varying and some effects constant. Time dependent variables and counting process data (multiple events per subject) are possible.

Resampling is used for computing p-values for tests of time-varying effects.

The modeling formula uses the standard survival modeling given in the **survival package.**

Usage

```
timecox(formula=formula(data),data=sys.parent(),
start.time=0,max.time=0,id=NULL,clusters=NULL,n.sim=1000,
residuals=0,robust=1,Nit=20,bandwidth=0.5,method="basic",
weighted.test=0,degree=1,covariance=0)
```

Arguments

`formula`	a formula object with the response on the left of a '~' operator, and the independent terms on the right as regressors. The response must be a survival object as returned by the 'Surv' function.
`data`	a data.frame with the variables.
`start.time`	start of observation period where estimates are computed.
`max.time`	end of observation period where estimates are computed. Estimates thus computed from [start.time, max.time]
`robust`	to compute robust variances and construct processes for resampling. May be set to 0 to save memory.
`id`	For time-varying covariates the variable must associate each record with the id of a subject.
`clusters`	cluster variable for computation of robust variances.
`n.sim`	number of simulations in resampling.
`weighted.test`	to compute a variance weighted version of the test-processes used for testing time-varying effects.
`residuals`	to returns residuals that can be used for model validation in the function cum.residuals
`covariance`	to compute covariance estimates for nonparametric terms rather than just the variances.
`Nit`	number of iterations for score equations.
`bandwidth`	bandwidth for local iterations. Default is 50 % of the range of the considered observation period.
`method`	Method for estimation. This refers to different parametrizations of the baseline of the model. Options

are "basic" where the baseline is written as $\lambda_0(t) = \exp(\alpha_0(t))$ or the "breslow" version where the baseline is parametrised as $\lambda_0(t)$.

`degree` gives the degree of the local linear smoothing, that is local smoothing. Possible values are 1 or 2.

Details

The data for a subject is presented as multiple rows or "observations", each of which applies to an interval of observation (start, stop]. When counting process data with the)start,stop] notation is used the 'id' variable is needed to identify the records for each subject. The program assumes that there are no ties, and if such are present random noise is added to break the ties.

Value

Returns an object of type "timecox". With the following arguments:

`cum` cumulative time-varying regression coefficient estimates are computed within the estimation interval.

`var.cum` the martingale based pointwise variance estimates.

`robvar.cum` robust pointwise variances estimates.

`gamma` estimate of parametric components of model.

`var.gamma` variance for gamma.

`robvar.gamma` robust variance for gamma.

`residuals` list with residuals.

`obs.testBeq0` observed absolute value of supremum of cumulative components scaled with the variance.

`pval.testBeq0` p-value for covariate effects based on supremum test.

`sim.testBeq0` resampled supremum values.

`obs.testBeqC` observed absolute value of supremum of difference between observed cumulative process and estimate under null of constant effect.

`pval.testBeqC` p-value based on resampling.

`sim.testBeqC` resampled supremum values.

`obs.testBeqC.is`
observed integrated squared differences between observed cumulative and estimate under null of constant effect.

`pval.testBeqC.is`
p-value based on resampling.

`sim.testBeqC.is`
: resampled supremum values.

`conf.band`
: resampling based constant to construct robust 95% uniform confidence bands.

`test.procBeqC`
: observed test-process of difference between observed cumulative process and estimate under null of constant effect over time.

`sim.test.procBeqC`
: list of 50 random realizations of test-processes under null based on resampling.

`schoenfeld.residuals`
: Schoenfeld residuals are returned for "breslow" parametrization.

Author(s)

Thomas Scheike

References

Martinussen and Scheike, Dynamic Regression Models for Survival Data, Springer (2006).

Examples

```
library(survival)
data(sTRACE)
# Fits time-varying Cox model
out<-timecox(Surv(time/365,status==9)~age+sex+diabetes+chf+vf,
sTRACE,max.time=7,n.sim=500)

summary(out)
par(mfrow=c(2,3))
plot(out)
par(mfrow=c(2,3))
plot(out,score=TRUE)

# Fits semi-parametric time-varying Cox model
out<-timecox(Surv(time/365,status==9)~const(age)+const(sex)+
const(diabetes)+chf+vf,sTRACE,max.time=7,n.sim=500)

summary(out)
par(mfrow=c(2,3))
plot(out)
```

`TRACE`	*The TRACE study group of myocardial infarction*

Description

The TRACE data frame contains 1877 patients and is a subset of a data set consisting of approximately 6000 patients. It contains data relating survival of patients after myocardial infarction to various risk factors.

sTRACE is a subsample consisting of 300 patients.

tTRACE is a subsample consisting of 1000 patients.

Format

This data frame contains the following columns:

id a numeric vector. Patient code.

status a numeric vector code. Survival status. 9: dead from myocardial infarction, 0: alive, 7: dead from other causes.

time a numeric vector. Survival time in years.

chf a numeric vector code. Clinical heart pump failure, 1: present, 0: absent.

diabetes a numeric vector code. Diabetes, 1: present, 0: absent.

vf a numeric vector code. Ventricular fibrillation, 1: present, 0: absent.

wmi a numeric vector. Measure of heart pumping effect based on ultrasound measurements where 2 is normal and 0 is worst.

sex a numeric vector code. 1: female, 0: male.

age a numeric vector code. Age of patient.

Source

The TRACE study group.

Jensen, G.V., Torp-Pedersen, C., Hildebrandt, P., Kober, L., F. E. Nielsen, Melchior, T., Joen, T. and P. K. Andersen (1997), Does in-hospital ventricular fibrillation affect prognosis after myocardial infarction?, European Heart Journal 18, 919–924.

Examples

```
data(TRACE)
names(TRACE)
```

Bibliography

AALEN, O. O. (1975). *Statistical inference for a family of counting processes.* PhD thesis, Univ. of California, Berkeley.

AALEN, O. O. (1978a). Nonparametric estimation of partial transition probabilities in multiple decrement models. *Ann. Statist.* **6**, 534–545.

AALEN, O. O. (1978b). Nonparametric inference for a family of counting processes. *Ann. Statist.* **6**, 701–726.

AALEN, O. O. (1980). A model for non-parametric regression analysis of counting processes. In Klonecki, W., Kozek, A., & Rosinski, J., editors, *Lecture Notes in Statistics-2: Mathematical Statistics and Probability Theory*, pages 1–25. Springer-Verlag New York.

AALEN, O. O. (1989). A linear regression model for the analysis of life times. *Statist. Med.* **8**, 907–925.

AALEN, O. O. (1993). Further results on the non-parametric linear regression model in survival analysis. *Statist. Med.* **12**, 1569–1588.

AALEN, O. O. & JOHANSEN, S. (1978). An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. *Scand. J. Statist.* **5**, 141–150.

ANDERSEN, E. W. (2005). Two-stage estimation in copula models used in family studies. *Lifetime Data Anal.* **11**, 333 – 350.

ANDERSEN, P. K., BORGAN, Ø., GILL, R. D., & KEIDING, N. (1993). *Statistical Models Based on Counting Processes.* Springer, New York.

ANDERSEN, P. K. & GILL, R. D. (1982). Cox's regression model for counting processes: A large sample study. *Ann. Statist.* **10**, 1100–1120.

ANDERSEN, P. K. & VÆTH, M. (1989). Simple parametric and nonparametric models for excess and relative mortality. *Biometrics* **45**, 523–535.

ARJAS, E. (1988). A graphical method for assessing goodness of fit in Cox's proportional hazards model. *J. Amer. Statist. Assoc.* **83**, 204–212.

BAGDONAVICIUS, V. & NIKULIN, M. (1999). Generalised proportional hazards model based on modified parital likelihood. *Lifetime Data Anal.* **5**, 329–350.

BAGDONAVICIUS, V. & NIKULIN, M. (2001). *Accelerated life models: Modelling and statistical analysis.* Chapman & Hall, London.

BARKER, P. & HENDERSON, R. (2004). Modelling converging hazards in survival analysis. *Lifetime Data Anal.* **10**, 263–281.

BERAN, R. (1981). Nonparametric regression with randomly censored survival data. Tech. rep., University of California, Berkeley.

BICKEL, P., KLASSEN, C., RITOV, Y., & WELLNER, J. (1993). *Efficient and Adaptive Estimation for Semiparametric Models.* Springer-Verlag New York.

BORGAN, Ø. (1984). Maximum likelihood estimation in parametric counting process models, with applications to censored failure time data. *Scand. J. Statist.* **11**, 1–16. Correction: **11**, 275.

BRÉMAUD, P. (1981). *Point Processes and Queues: Martingale Dynamics.* Springer-Verlag, New York.

BROOKMEYER, R. & GAIL, M. H. (1987). Biases in prevalent cohorts. *Biometrics* **43**, 739–749.

BRUMBACK, B. & RICE, J. (1998). Smoothings spline models for the analysis of nested and crossed samples of curves. *J. Amer. Statist. Assoc.* **93**, 961–994.

BUCKLEY, J. & JAMES, I. R. (1979). Linear regression with censored data. *Biometrika* **66**, 429–436.

CAI, J. & PRENTICE, R. L. (1995). Estimating equations for hazard ratio parameters based on correlated failure time data. *Biometrika* **82**, 151–164.

Cai, J. & Prentice, R. L. (1997). Regression estimation using multivariate failure time data and a common baseline hazard function model. *Lifetime Data Anal.* **3**, 197–213.

Cai, T., Wei, L. J., & Wilcox, M. (2000). Semiparametric regression analysis for clustered failure time data. *Biometrika* **87**, 867–878.

Cai, Z. & Sun, Y. (2003). Local linear estimation for time-dependent coefficients in Cox's regression models. *Scand. J. Statist.* **30**, 93–112.

Chen, K., Jin, Z., & Ying, Z. (2002). Semiparametric analysis of transformation models with censored data. *Biometrika* **89**, 659–668.

Chen, Y. & Jewell, N. (2001). On a general class of semiparametric hazards regression models. *Biometrika* **88**, 687–702.

Chen, Y. Q. & Wang, M.-C. (2000). Analysis of accelerated hazards models. *J. Amer. Statist. Assoc.* **95**, 608–618.

Cheng, S. C., Fine, J. P., & Wei, L. J. (1998). Prediction of cumulative incidence function under the proportional hazards model. *Biometrics* **54**, 219–228.

Cheng, S. C. & Wei, L. J. (2000). Inferences for a semiparametric model with panel data. *Biometrika* **87**, 89–97.

Cheng, S. C., Wei, L. J., & Ying, Z. (1995). Analysis of transformation models with censored data. *Biometrika* **82**, 835–845.

Cheng, S. C., Wei, L. J., & Ying, Z. (1997). Prediction of survival probabilities with semi-parametric transformation models. *J. Amer. Statist. Assoc.* **92**, 227–235.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141–151.

Collett, D. (2003). *Modelling Survival Data in Medical Research.* Chapman Hall, London.

Cox, D. R. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. Ser. B* **34**, 187–220.

Cox, D. R. (1975). Partial likelihood. *Biometrika* **62**, 269–276.

Dabrowska, D. M. (1997). Smoothed Cox regression. *Ann. Statist.* **25**, 1510–1540.

Dauxois, J.-Y. & Kirmani, S. N. U. A. (2003). Testing the proportional odds model under random censoring. *Biometrika* **90**, 913–922.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B* **39**, 1–38. With discussion.

Fan, J. & Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications.* Chapman Hall London, New York.

Fan, J. & Zhang, J. (2000a). Statistical estimation in varying-coefficient models. *J. Roy. Stat. Soc. Ser. B* **62**, 303–322.

Fan, J. & Zhang, W. (1999). Statistical estimation in varying-coefficient models. *Ann. Statist.* **27**, 1491–1518.

Fan, J. & Zhang, W. (2000b). Simultaneous confidence bands and hypothesis testing in varying-coefficient models. *Scand. J. Statist.* **27**, 715–727.

Fine, J., Ying, Z., & Wei, L. J. (1998). On the linear transformation model with censored data. *Biometrika* **85**, 980–986.

Fine, J. P. & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *J. Amer. Statist. Assoc.* **94**, 496–509.

Fleming, T. R. (1978a). Asymptotic distribution results in competing risks estimation. *Ann. Statist.* **6**, 1071–1079.

Fleming, T. R. (1978b). Nonparametric estimation for nonhomogeneous Markov processes in the problem of competing risks. *Ann. Statist.* **6**, 1057–1079.

Fleming, T. R. & Harrington, D. P. (1991). *Counting Processes and Survival Analysis.* Wiley, New York.

Fygenson, M. & Ritov, Y. (1994). Monotone estimating equations for censored data. *Ann. Statist.* **22**, 732–746.

Genest, C. & MacKay, R. J. (1986). The joy of copulas: Bivariate distributions with uniform marginals. *Amer. Statist.* **40**, 280–283.

Genest, C. & MacKay, R. J. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552.

Gill, R. D. (1983). Discussion of the papers by Helland and Kurtz. *Bull. Internat. Statist. Inst.* **50**, 239–243.

Gill, R. D. (1985). Discussion of the paper by D. Clayton and J. Cuzick. *J. Roy. Statist. Soc. Ser. A* **148**, 108–109.

Gill, R. D. & Johansen, S. (1990). A survey of product-integration with a view towards application in survival analysis. *Ann. Statist.* **18**, 1501–1555.

Glidden, D. V. (1999). Checking the adequacy of the gamma frailty model for multivariate failure times. *Biometrika* **86**, 381–393.

Glidden, D. V. (2000). A two-stage estimator of the dependence parameter for the Clayton-Oakes model. *Lifetime Data Anal.* **6**, 141–156.

Grambsch, P. M. & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals (corr: 95v82 p668). *Biometrika* **81**, 515–526.

Gray, R. J. (1988). A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Ann. Statist.* **16**, 1141–1154.

Greenwood, M. (1926). The natural duration of cancer. In *Reports on Public Health and Medical Subjects* **33**, pages 1–26. His Majesty's Stationery Office, London.

Greenwood, P. E. & Wefelmeyer, W. (1990). Efficiency of estimators for partially specified filtered models. *Stoch. Proc. and their Appl.* **36**, 353–370.

Greenwood, P. E. & Wefelmeyer, W. (1991). Efficient estimating equations for nonparametric filtered models. In Prabhu, N. U. & Basawa, I. V., editors, *Statistical Inference in Stochastic Processes*, pages 107–141, New York. Marcel Dekker.

Hall, W. J. & Wellner, J. A. (1980). Confidence bands for a survival curve from censored data. *Biometrika* **67**, 133–143.

Harboe, I. S. & Scheike, T. H. (2001). Time-varying coefficient model for longitudinal data: Uniform confidence bands and hypothesis tests applied to growth curves. *unpublished manuscript* .

Harrington, D. P. & Fleming, T. R. (1982). A class of rank test procedures for censored survival data. *Biometrika* **69**, 133–143.

Hastie, T. & Tibshirani, R. (1993). Varying-coefficient models. *J. Roy. Stat. Soc. Ser. B* **55**, 757–796.

Härdle, W. (1990). *Applied Nonparametric Regression.* Cambridge University press.

Hoover, P., Rice, J., Wu, C., & Yang, L.-P. (1998). Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. *Biometrika* **85**, 809–822.

Horvitz, D. & Thompson, D. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Stastist. Assoc.* **47**, 663–685.

Hougaard, P. (2000). *Analysis of multivariate survival data.* Statistics for Biology and Health. Springer-Verlag, New York.

Huang, J. Z., Wu, C. O., & Zhou, L. (2002). Varying-coefficient models and basis function approximations for the analysis of repeated measurements. *Biometrika* **89**, 111–128.

Huffer, F. W. & McKeague, I. W. (1991). Weighted least squares estimation for Aalen's additive risk model. *J. Amer. Statist. Assoc.* **86**, 114–129.

Huster, W. J., Brookmeyer, R., & Self, S. G. (1989). Modelling paired survival data with covariates. *Biometrics* **45**, 145–156.

Jacobsen, M. (1982). *Statistical Analysis of Counting Processes*, volume 12 of *Lecture Notes in Statistics.* Springer-Verlag, New York.

Jacobsen, M. (1989). Existence and unicity of MLEs in discrete exponential family distributions. *Scand. J. Statist.* **16**, 335–349.

Jensen, G. V., Torp-Pedersen, C., Hildebrandt, P., Kober, L., Nielsen, F. E., Melchior, T., Joen, T., & Andersen, P. K. (1997). Does in-hospital ventricular fibrillation affect prognosis after myocardial infarction? *European Heart Journal* **18**, 919–924.

Jespersen, N. C. B. (1986). Dichotomizing a continuous covariate in the Cox regression model. Research Report 86/2, Statistical Research Unit, University of Copenhagen.

Jin, Z., Lin, D. Y., Wei, L. J., & Ying, Z. (2003). Rank-based inference for the accelerated failure time model. *Biometrika* **90**, 341–353.

Johansen, S. (1978). The product limit estimator as maximum likelihood estimator. *Scand. J. Statist.* **5**, 195–199.

Kalbfleisch, J. D. & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data.* Wiley, New York.

Kaplan, E. L. & Meier, P. (1958). Non-parametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53**, 457–481.

Kaslow, R. A., Ostrow, D. G., Detels, R., Phair, J. P., Polk, B. F., & Rinaldo, C. R. (1987). The Multicenter AIDS Cohort Study: rationale, organization and selected characteristics of the participants. *Am. J. Epidem.* **126**, 310–318.

Khmaladze, E. V. (1981). Martingale approach to the goodness of fit tests. *Theory Probab. Appl.* **26**, 246–265.

Klein, J. P. (1992). Semiparametric estimation of random effects using the Cox model based on the EM algorithm. *Biometrics* **48**, 795–806.

Klein, J. P. & Moeschberger, M. L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data.* Springer-Verlag Inc.

Koul, H., Susarla, V., & Van Ryzin, J. (1981). Regression analysis with randomly right censored data. *Ann. Statist.* **9**, 1276–1288.

Lai, T. L. & Ying, Z. (1991a). Large sample theory of a modified Buckley-James estimator for regression analysis with censored data. *Ann. Statist.* **19**.

Lai, T. L. & Ying, Z. (1991b). Rank regression methods for left-truncated and right-censored data. *Ann. Statist.* **19**.

Last, G. & Brandt, A. (1995). *Marked Point Processes on the Real Line. The Dynamic Approach.* Springer-Verlag New York.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data.* Wiley, New York.

Lawless, J. F. & Nadeau, C. (1995). Some simple robust methods for the analysis of recurrent events. *Technometrics* **37**, 158–168.

Lee, E. W., Wei, L. J., & Amato, D. A. (1992). Cox-type regression analysis for large numbers of small groups of correlated failure time observations (disc: P247). In *Survival Analysis*, pages 237– 247.

Liang, K. Y. & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Lin, D. Y. (1991). Goodness of fit for the Cox regression model based on a class of parameter estimators. *J. Amer. Statist. Assoc.* **86**, 725–728.

Lin, D. Y. (1994). Cox regression analysis of multivariate failure time data: The marginal approach. *Statist. Med.* **13**, 2233–2247.

Lin, D. Y., Oakes, D., & Ying, Z. (1998a). Additive hazards regression with current status data. *Biometrika* **81**, 289–298.

Lin, D. Y., Wei, L., Yang, I., & Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *J. Roy. Stat. Soc. Ser. B* **62**, 711–730.

Lin, D. Y. & Wei, L. J. (1989). The robust inference for the Cox proportional hazards model. *J. Amer. Statist. Assoc.* **84**, 1074–1078.

Lin, D. Y., Wei, L. J., & Ying, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* **80**, 557–572.

Lin, D. Y., Wei, L. J., & Ying, Z. (1998b). Accelerated failure time models for counting processes. *Biometrika* **85**, 605–618.

Lin, D. Y. & Ying, Z. (1994). Semiparametric analysis of the additive risk model. *Biometrika* **81**, 61–71.

Lin, D. Y. & Ying, Z. (1995). Semiparametric analysis of general additive-multiplicative hazard models for counting processes. *Ann. Statist.* **23**, 1712–1734.

Lin, D. Y. & Ying, Z. (2001). Semiparametric and nonparametric regression analysis of longitudinal data. *J. Amer. Statist. Assoc.* **96**, 103–126.

Little, R. & Rubin, D. (1987). *Statistical Analysis with Missing Data.* Wiley , New York.

Lynn, H. (2001). Maximum likelihood inference for left-censored hiv rna data. *Statist. Med.* **20**, 33–45.

MacRae, E. C. (1974). Matrix derivatives with an application to an adaptive linear decision problem. *Ann. Statist.* **2**, 337–346.

Martinussen, T. (1997). *Statistical analysis based on incomplete longitudinal data with mortality follow-up.* PhD thesis, Department of Biostatistics, University of Copenhagen.

Martinussen, T. & Pipper, C. B. (2005). Estimation in the positive stable shared frailty Cox proportional hazards model. *Lifetime Data Anal.* **11**, 99–115.

Martinussen, T. & Scheike, T. H. (1999). A semi-parametric additive regression model for longitudinal data. *Biometrika* **86**, 691–702.

Martinussen, T. & Scheike, T. H. (2000). A non-parametric dynamic additive regression model for longitudinal data. *Ann. Statist.* **28**, 1000–1025.

Martinussen, T. & Scheike, T. H. (2001). Sampling adjusted analysis of dynamic additive regression models for longitudinal data. *Scand. J. Statist.* **28**, 303–323.

Martinussen, T. & Scheike, T. H. (2002a). Efficient estimation in additive hazards regression with current status data. *Biometrika* **89**, 649–658.

Martinussen, T. & Scheike, T. H. (2002b). A flexible additive multiplicative hazard model. *Biometrika* **89**, 283–298.

MARTINUSSEN, T., SCHEIKE, T. H., & SKOVGAARD, I. M. (2002). Efficient estimation of fixed and time-varying covariate effects in multiplicative intensity models. *Scand. J. Statist.* **28**, 57–74.

MARTINUSSEN, T. & SKOVGAARD, I. M. (2002). Goodness-of-fit of semiparametric regression models. Technical report, Royal Veterinary and Agricultural University, Copenhagen.

MARTINUSSEN, T. & SØRENSEN, T. I. A. (1998). Age-dependent U-shaped risk functions and Aalen's additive risk model. *Biometrics* **54**, 989–1001.

MCKEAGUE, I. W. (1988). Asymptotic theory for weighted least squares estimators in Aalen's additive risk model. *Contemp. Math.* **80**, 139–152.

MCKEAGUE, I. W. & SASIENI, P. D. (1994). A partly parametric additive risk model. *Biometrika* **81**, 501–514.

MCKEAGUE, I. W. & UTIKAL, K. J. (1990a). Identifying nonlinear covariate effects in semimartingale regression models. *Probab. Theory and Related Fields* **87**, 1–25.

MCKEAGUE, I. W. & UTIKAL, K. J. (1990b). Inference for a nonlinear counting process regression model. *Ann. Statist.* **18**, 1172–1187.

MEYER, P. A. (1976). *Séminaire de Probabilités X.*, volume 511 of *Lecture Notes in Mathematics*, pages 245–400. Springer-Verlag, Berlin.

MILLER, R. G. (1976). Least squares regression with censored data. *Biometrika* **63**, 449–464.

MURPHY, S., ROSSINI, A., & VAN DER VAART, A. (1997). Maximum likelihood estimation in the proportional odds model. *J. Amer. Statist. Assoc.* **92**, 968–976.

MURPHY, S. A. (1993). Testing for a time dependent coefficient in Cox's regression model. *Scand. J. Statist.* **20**, 35–50.

MURPHY, S. A. & SEN, P. K. (1991). Time-dependent coefficients in a Cox-type regression model. *Stochastic Proccess. Appl.* **39**, 153–180.

NÆS, T. (1982). The asymptotic distribution of the estimator for the regression parameter in Cox's regression model. *Scand. J. Statist.* **9**, 107–115.

NELSON, W. (1969). Hazard plotting for incomplete failure data. *J. Qual. Technol.* **1**, 27–52.

NELSON, W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics* **14**, 945–965.

Nielsen, G. G., Gill, R. D., Andersen, P., & Sørensen, T. I. A. (1992). A counting process approach to maximum likelihood estimation in frailty models. *Scand. J. Statist.* **19**, 25–43.

Oakes, D. (1982). A model for association in bivariate survival data. *J. Roy. Statist. Soc. Ser. B* **44**, 414–422.

Parner, E. (1998). Asymptotic theory for the correlated gamma-frailty model. *Ann. Statist.* **26**, 183–214.

Pepe, M. & Cai, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *J. Amer. Statist. Assoc.* **88**, 811–820.

Pepe, M. & Couper, D. (1997). Modelling partly conditional means with longitudinal data. *J. Amer. Statist. Assoc.* **92**, 991–998.

Pepe, M. S. (1991). Inference for events with dependent risks in multiple endpoint studies. *J. Amer. Statist. Assoc.* **86**, 770–778.

Pipper, C. B. & Martinussen, T. (2003). A likelihood based estimating equation for the Clayton-Oakes model with marginal proportional hazards. *Scand. J. Statist.* **30**, 509–521.

Pipper, C. B. & Martinussen, T. (2004). An estimating equation for parametric shared frailty models with marginal additive hazards. *J. Roy. Statist. Soc. Ser. B* **66**, 207–220.

Pons, O. (2000). Nonparametric estimation in a varying-coefficient Cox model. *Math. Meth. Statist.* **9**, 376–398.

Prentice, R. L. (1978). Linear rank tests with right censored data. *Biometrika* **65**, 167–179. Correction **70**,304(1983).

Priestley, M. B. & Chao, M. T. (1972). Nonparametric function fitting. *J. Roy. Stat. Soc. Ser. B* **34**, 385–392.

Ramlau-Hansen, H. (1983a). The choice of a kernel function in the graduation of counting process intensities. *Scand. Actuar. J.* pages 165–182.

Ramlau-Hansen, H. (1983b). Smoothing counting process intensities by means of kernel functions. *Ann. Statist.* **11**, 453–466.

Rebolledo, R. (1980). Central limit theorems for local martingales. *Z. Wahrsch. verw. Geb.* **51**, 269–286.

Ritov, Y. (1990). Estimation in a linear regression model with censored data. *Ann. Statist.* **18**, 303–328.

ROBINS, J. M. & ROTNITZKY, A. (1992). Recovery of information and adjustment of dependent censoring using surrogate markers. In *AIDS Epidemiology-Methodological Issues*, pages 24–33.

ROSTHØJ, S., ANDERSEN, P. K., & ABILDSTRØM, S. Z. (2004). SAS macros for estimation of the cumulative incidence function s based on a Cox regressio nmodel for competing risks survival data. *Computer methods and programs in biomedicine* **74**, 69–75.

SASIENI, P. (1992a). Information bounds for the conditional hazard ratio in a nested family of regression models. *J. Roy. Stat. Soc. Ser. B* **54**, 617–635.

SASIENI, P. D. (1992b). Information bounds for the additive and multiplicative intensity models (disc: P263-265). In Klein, J. & Goel, P., editors, *Survival Analysis: State of the Art*, pages 249– 263. Klüwer, Dordrecht.

SASIENI, P. D. (1996). Proportional excess hazards. *Biometrika* **83**, 127–141.

SCHEIKE, T. H. (1994). Parametric regression for longitudinal data with counting process measurement times. *Scand. J. Statist.* **21**, 245–263.

SCHEIKE, T. H. (1996). Nonparametric kernel regression when the regressor follows a counting process. *J. Nonparametr. Statist.* **2**, 1–11.

SCHEIKE, T. H. (2000). Comparison of non-parametric regression functions through their cumulatives. *Statist. Probab. Lett.* **46**, 21–32.

SCHEIKE, T. H. (2002). The additive nonparametric and semiparametric Aalen model as the rate function for a counting process. *Lifetime Data Anal.* **8**, 247–262.

SCHEIKE, T. H. (2004). Time-varying effects in survival analysis. In Balakrishnan, N. & Rao, C. R., editors, *Handbook of Statistics 23*, pages 61–85. Elsevier B.V.

SCHEIKE, T. H. & JENSEN, T. K. (1997). A discrete survival model with random effects: An application to time to pregnancy. *Biometrics* **53**, 318–329.

SCHEIKE, T. H. & MARTINUSSEN, T. (2002). Nonparametric and semiparametric dynamic marginal additive regression models. *Research Report, Department of Mathematical Sciences, University of Aalborg* **2002**.

SCHEIKE, T. H. & MARTINUSSEN, T. (2004). On efficient estimation and tests of time-varying effects in the proportional hazards model. *Scand. J. Statist.* **31**, 51–62.

Scheike, T. H., Petersen, J. H., & Martinussen, T. (1999). Retrospective ascertainment of recurrent events: An application to time to pregnancy. *J. Amer. Statist. Assoc.* **94**, 713–725.

Scheike, T. H. & Zhang, M. (1998). Cumulative regression function tests for longitudinal data. *Ann. Statist.* **26**, 1328–1355.

Scheike, T. H. & Zhang, M.-J. (2002). An additive-multiplicative Cox-Aalen model. *Scand. J. Statist.* **28**, 75–88.

Scheike, T. H. & Zhang, M.-J. (2003). Extensions and applications of the Cox-Aalen survival model. *Biometrics* **59**, 1033–1045.

Scheike, T. H. & Zhang, M.-J. (2004). Predicting cumulative incidence probability: Marginal and cause-specific modelling. *Research Report, Department of Biostatistics, University of Copenhagen* **3**.

Scheike, T. H. & Zhang, M.-J. (2005). Predicting cumulative incidence probability by direct binomial regression. *Research Report, Department of Biostatistics, University of Copenhagen* **2**.

Schlichting, P., Christensen, E., Andersen, P. K., Fauerholdt, L., Juhl, E., Poulsen, H., & Tygstrup, N. for The Copenhagen Study Group for Liver Diseases (1983). Identification of prognostic factors in cirrhosis using Cox's regression model. *Hepatology* **3**, 889–895.

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika* **69**, 239–241.

Schumacher, M. (1984). Two-sample tests of Cramér-von Mises and Kolmogorov-Smirnov type for randomly censored data. *Internat. Statist. Rev.* **52**, 263–281.

Shen, Y. & Cheng, S. C. (1999). Confidence bands for cumulative incidence curves under the additive risk model. *Biometrics* **55**, 1093–1100.

Shih, J. H. & Louis, T. A. (1995). Inference on association parameter in copula models for bivariate survival data. *Biometrics* **51**, 1384–1399.

Simonoff, J. (1996). *Smoothing Methods in Statistics.* Springer-Verlag New York.

Slud, E. & Vonta, F. (2004). Consistency of the NPML estimator in the right-censored transformation model. *Scand. J. Statist.* **31**, 21–41.

Solomon, P. J. (1984). Effect of misspecification of regression models in the analysis of survival data. *Biometrika* **71**, 291–298.

Spiekerman, C. F. & Lin, D. Y. (1996). Checking the marginal Cox model for correlated failure time data. *Biometrika* **83**, 143–156.

SPIEKERMAN, C. F. & LIN, D. Y. (1998). Marginal regression models for multivariate failure time data. *J. Amer. Statist. Assoc.* **93**, 1164–1175.

STRUTHERS, C. A. & KALBFLEISCH, J. D. (1986). Misspecified proportional hazard models. *Biometrika* **73**, 363–369.

SUN, Y. & WU, H. (2003). AUC-based tests for nonparmetric functions with longitudinal data. *Statist. Sinica* **13**, 593–612.

SUN, Y. & WU, H. (2005). Semiparmetric time-varying coefficient regression model for longitudinal data. *Scand. J. Statist.* **32**, 21–47.

THERNEAU, T. & GRAMBSCH, P. (2000). *Modeling Survival Data: Extending the Cox Model.* Springer-Verlag New York.

TIAN, L., ZUCKER, D., & WEI, L. (2005). On the Cox model with time-varying regression coefficients. *J. Amer. Statist. Assoc.* **100**, 172–183.

TSIATIS, A. A. (1981). A large sample study of Cox's regression model. *Ann. Statist.* **9**, 93–108.

TSIATIS, A. A. (1990). Estimating regression parameters using linear rank tests for censored data. *Ann. Statist.* **18**, 354–372.

VAN DER VAART, A. W. & WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics.* Springer-Verlag New York.

WEI, L. J. (1984). Testing goodness of fit for proportional hazards model with censored observations. *J. Amer. Statist. Assoc.* **79**, 649–652.

WEI, L. J., LIN, D., & WEISSFELD, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J. Amer. Statist. Assoc.* **84**, 1065–1073.

WEI, L. J., YING, Z., & LIN, D. Y. (1990). Linear regression analysis of censored survival data based on rank tests. *Biometrika* **77**, 845–852.

WINNETT, A. & SASIENI, P. (2001). A note on scaled Schoenfeld residuals for the proportional hazards model. *Biometrika* **88**, 565–571.

WINNETT, A. & SASIENI, P. (2003). Iterated residuals and time-varying covariate effects in Cox regression. *J. Roy. Stat. Soc. Ser. B* **65**, 473–488.

WU, C. & CHIANG, C. (2000). Kernel smoothing on varying coefficient models with longitudinal dependent variable. *Statist. Sinica* **10**, 433–456.

WU, C., CHIANG, C., & HOOVER, D. R. (1998). Asymptotic confidence regions for kernel smoothing of a varying-coefficient model with longitudinal data. *J. Amer. Statist. Assoc.* **93**, 1388–1402.

Xu, R. & O'Quigley, J. (2000). Estimating average regression effect under non-proportional hazards. *Biostatistics* **1**, 423–439.

Ying, Z. (1993). A large sample study of rank estimation for censored regression data. *Ann. Statist.* **21**.

Ying, Z., Jung, S. H., & Wei, L. J. (1995). Survival analysis with median regression models. *J. Amer. Statist. Assoc.* **90**, 178–184.

Zahl, P. (2003). Regresion analysis with multiplicative and time-varying additive regression coefficients with examples from breast and colon cancer. *Statist. Med.* **22**, 1113–1127.

Zucker, D. M. & Karr, A. F. (1990). Nonparametric survival analysis with time-dependent covariate effects: A penalized partial likelihood approach. *Ann. Statist.* **18**, 329–353.

Index