SAS exercises

Henrik Ravn

Novo Nordisk

DSBS Course
Survival Analysis in Clinical Trials – Part 2
January 2018

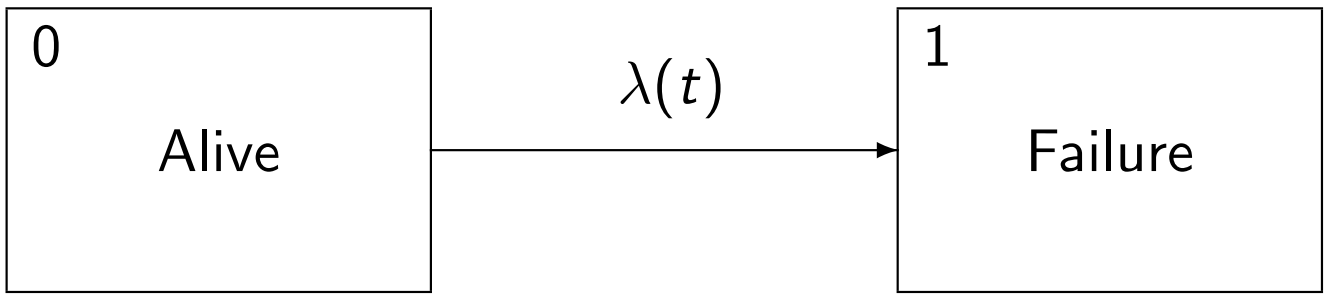
PBC3 data – description

PBC3 was a multi-centre randomized clinical trial conducted in six European hospitals. Between 1 Jan. 1983 and 1 Jan. 1987, 349 patients with the liver disease primary biliary cirrhosis (PBC) were randomized to either treatment with Cyclosporin A (CyA, 176 patients) or placebo (173 patients). The purpose of the trial was to study the effect of treatment on the survival time. However, during the course of the trial an increased use of liver transplantation for patients with this disease made the investigators redefine the main response variable to be time to "failure of medical treatment" defined as either death or liver transplantation. Patients were then followed from randomization until treatment failure, drop-out or 1 Jan, 1989; 61 patients died (CyA: 30, placebo: 31), another 29 were transplanted (CyA: 14, placebo: 15) and 4 patients were lost to follow-up before 1 Jan. 1989. At entry a number of clinical, biochemical and histological variables, including serum bilirubin, serum albumin, sex, age were recorded.

PBC3 data – list of variables

`ptno`	patient identification
`tment`	treatment (0: placebo, 1: CyA)
`sex`	(1: males, 0: females)
`age`	years
`alb`	albumin (g/L)
`bili`	bilirubin (micromoles/L)
`years`	observation time in years
`status`	status at exit (0: censored, 1: liver transpl, 2 : dead)
`biligroup`	Bilirubin categorised by quintiles

PBC3 – two-state model



PBC3 exercise – goodness of fit

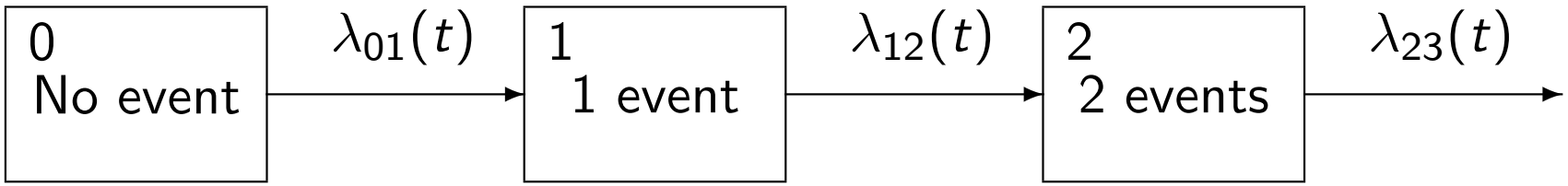
1. Fit a Cox model with treatment (`tment`) as the only covariate
2. Add (to the previous model) the quantitative covariates bilirubin and albumin (`bili` and `alb`). What happened to the effect of treatment?
3. Check proportional and log-linear assumptions for the three covariates in the previous model (use `ASSESS` statement)
4. Try to use logarithm of bilirubin instead. What happened to the effect of treatment?

Rats data – recurrent events with no gaps and no terminal event

Data from Cook and Lawless 2007 Springer-book "The Statistical Analysis of Recurrent Events" ("orginally" from Gail et al., 1980, Biometrics). 76 female rats were exposed to a carcinogen and then given retinyl acetate to prevent cancer for 60 days. 48 rats, still tumor-free, were randomized to either continued treatment (23) or control (25) and followed for another 122 days. They were examined for tumors twice weekly and times of tumors were noted. The data set includes the variables:

`id`	rat id number
`start`	start time
`stop`	stop time
`status`	tumor (=1) or not (=0) at stop time
`num`	record number for each rat
`trt`	treatment indicator

Rats – multi-state model



Rats exercise

Intensity model

1. Fit an intensity-based model using only treatment as a covariate. What is the interpretation of the hazard ratio for the treatment effect?
2. Fit a stratified intensity-based model stratified by number of events (PWP-model) to estimate the treatment effect. What is the interpretation of the hazard ratio for the treatment effect?
3. Fit an (un-stratified) intensity-based model using the number of previous events as a quantitative covariate
4. Fit an (un-stratified) intensity-based model using the number of previous events as a categorical covariate. How do all these models differs?

Frailty model

1. Fit a gamma-frailty model and estimate the treatment effect. What is the interpretation of the hazard ratio for the treatment effect?
2. Does frailty seem to be present?

Rats exercise – continued

Marginal model

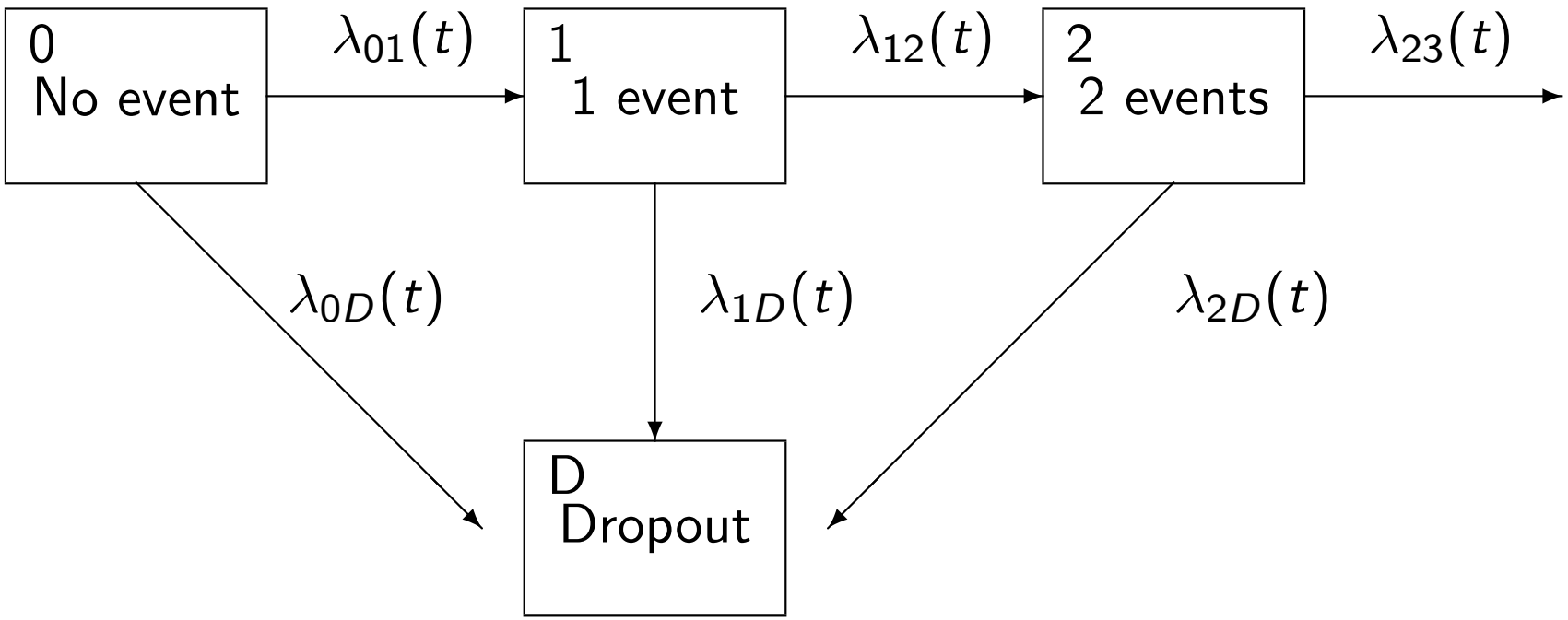
1. Plot the estimated (non-parametrically) cumulative marginal mean function (CMF) for each treatment
2. Compare the estimates SAS calculates for `CUMHAZ` and `CMF` in the `BASELINE` statement (maybe the PROC PHREG manual is needed here)
3. Fit a model for the marginal mean function to estimate the treatment effect using robust SE. How do you interpret the exponentiated regression coefficient?

Hypo – hypoglycaemic recurrent events with dropout as terminal event

RCT data on time to treatment emergent hypoglycaemic event (recurrent) in a 1-year RCT (N=559) of active drug vs placebo. A treatment emergent hypoglycaemic event is defined as an event that has onset date on or after the first day of randomised treatment and no later than 14 days after the last day of randomised treatment. The event process is interrupted by dropout (drug discontinuation) before week 52.

`id`	Subject id
`enum`	record number per id
`start`	Start time in weeks
`stop`	Stop time in weeks
`status`	Status at stop time: 0 = Complete on drug 1 = hypo event 2 = dropout
`drug`	0 = Placebo 1 = Active

Hypo – multi-state model



Hypo exercise

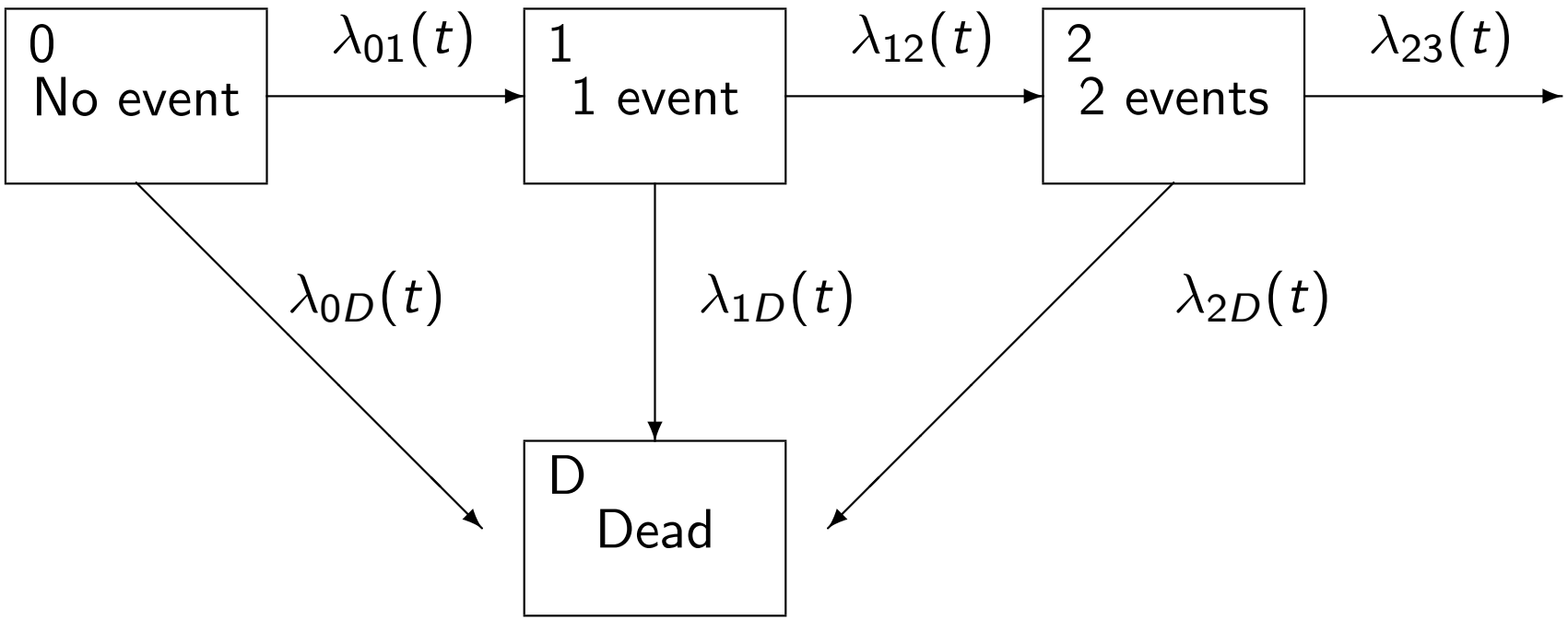
1. Estimate non-parametrically the cumulative marginal mean function for the recurrent hypos for each treatment group by censoring for dropout. What does it estimate?
2. Estimate non-parametrically the cumulative marginal mean function taking into account the terminal event (dropout) for each treatment group using PHREG. What does it estimate?
3. Fit an intensity-based model for hypo events, by censoring for dropout with `drug` as covariate. What is the interpretation of the hazard ratio?
4. Fit a model for the marginal mean ($E(N(t))$ for hypo events taking into account the terminal event (dropout) with `drug` as covariate.

Bladder data – recurrent events with death as terminal event

Trial conducted by the Veterans Administration Cooperative Urological Research Group (Byar, 1980) - famous text book example. 118 patients with stage I bladder cancer randomized to pyridoxine (32), thiotepa (38), or placebo (48) followed for the occurrence of superficial bladder tumors. Here we only look at placebo and thiotepa with death as a terminating event:

`subject`	person-id
`enum`	record no per subject
`start`	start time in months
`stop`	stop time in months
`status`	0 = alive, 1 = new tumor, 2 = dead, 3 = new tumor and dead
`trt`	0 = placebo, 1 = thiotepa
`number`	no. of tumors at time 0
`size`	largest tumor at time 0

Bladder – multi-state model



Bladder exercise

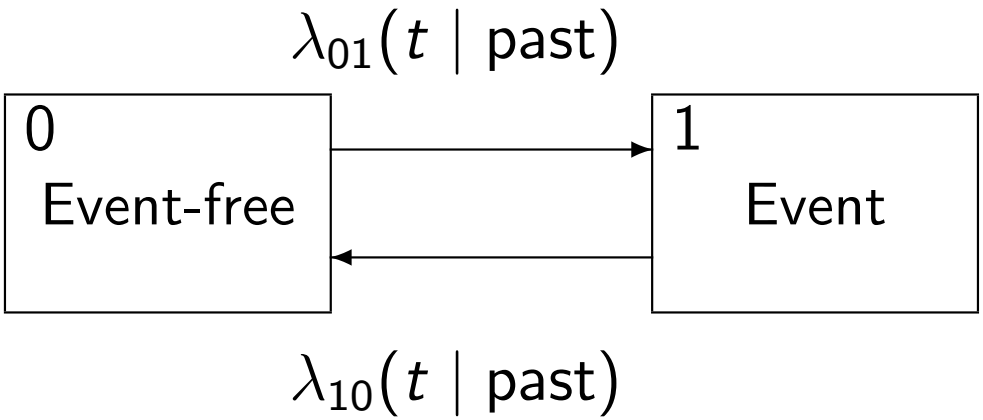
1. Fit an intensity-based model for the recurrent tumors, by censoring for death. What is the interpretation of the hazard ratio?
2. Fit a Cox-model using death as an extra event in the event process. What is the interpretation of the hazard ratio?

rhDNase – recurrent events with gaps

This is a dataset reported by Fuchs et al., 1994, N Eng J Med for a double-blind randomized multicenter clinical trial designed to evaluate the effect of rhDNase, a recombinant deoxyribonuclease I enzyme, versus placebo on the occurrence of respiratory exacerbations among patients with cystic fibrosis. Data on the occurrence and resolution of all exacerbations were recorded for 645 patients with cystic fibrosis randomized to rhDNase (321) or placebo (324) followed from randomization and about 169 days. The data set includes the variables:

`id`	subject id
`trt`	1 = rhDNase, 0 = placebo
`fev, fev2`	baseline measurements
`start`	start time
`stop`	stop time
`status`	exacerbation (=1) or not (=0) at stop time
`etype`	1 if "at risk", 2 if "under treatment" = not at risk for an exacerbation
`enum`	record no.
`enum1`	gap time no.
`enum2`	treatment period number

rhDNase – multi-state model



rhDNase exercise

1. Estimate non-parametrically the cumulative intensity for events of exacerbations for each treatment by removing time not at risk due to treatment of exacerbations (`etype`=2)
2. Fit an intensity-based Cox model for the exacerbations, by removing time not at risk due to treatment of exacerbations (`etype`=2) with randomised treatment (`trt`) and baseline FEV (`fev`) as covariates. What is the interpretation of the hazard ratio?
3. Estimate non-parametrically the cumulative marginal mean function for events of exacerbations for each treatment, *without* any adjustment for days not at risk (etype=2), so-called cycles of exacerbations. What does this estimate?
4. Fit a marginal models for cycles of exacerbations with `fev` and `trt` as covariates. What is the interpretation of the "hazard ratio" in the SAS-output?