PhD course in 'Advanced survival analysis'

Analysis of recurrent events

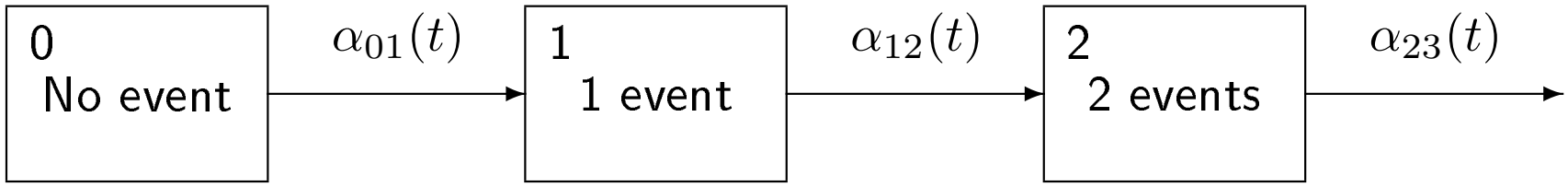
Per Kragh Andersen: 27 October 2021

Summary

- Multi-state models for recurrent events
  - Examples
  - Competing risks or not?
  - Do events have a 'duration'?
- Models for intensities
  - Jacod's likelihood
  - Hazard models from survival analysis
  - Accounting for intra-individual dependence
  - Exercise 1
- Marginal models
  - No competing risks: Marginal parameters (including $E(N(t))$); regression models
  - Exercise 2
  - Models in the presence of competing risks
  - Exercise 3

Reference: Cook and Lawless (2007). *The Statistical Analysis of Recurrent Events*, Springer.

Recurrent events: mortality negligible, 'no duration'



$N_{hj}(t)$ counting process with intensity process $\alpha_{hj}(t \mid \mathcal{F}_{t-})Y_h(t)$.

Transition intensities ('hazards') $\alpha_{hj}(t)$ may depend on the past $(\mathcal{F}_t)$, e.g. on number and/or times of previous events.

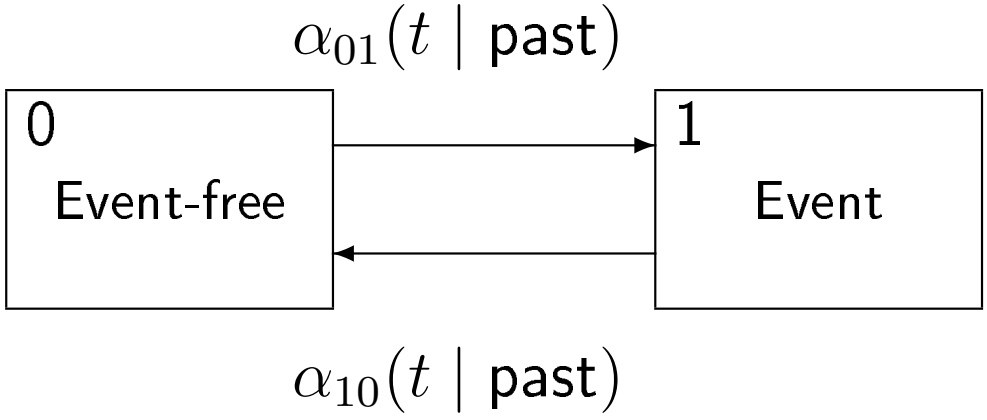
Example: tumors in rats

Data from Cook and Lawless 2007 Springer-book "The Statistical Analysis of Recurrent Events" ("orginally" from Gail et al., 1980, *Biometrics*). Data set `rats2` in the `survival` package of `R`.

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`
- `time1` ('start') , `time2` ('stop'), `status` (tumor, 1 or not, 0)
- `obs` (record no.)
- `trt` (treatment indicator 1 or 0)

Recurrent events 'with duration', mortality negligible



Here, there are 'gaps' after occurrence of the event where the subject is not at risk for a new event. In both states there may be a past to consider when modeling the intensities.

Example: Pulmonary exacerbations

Data from Cook and Lawless (2007) book

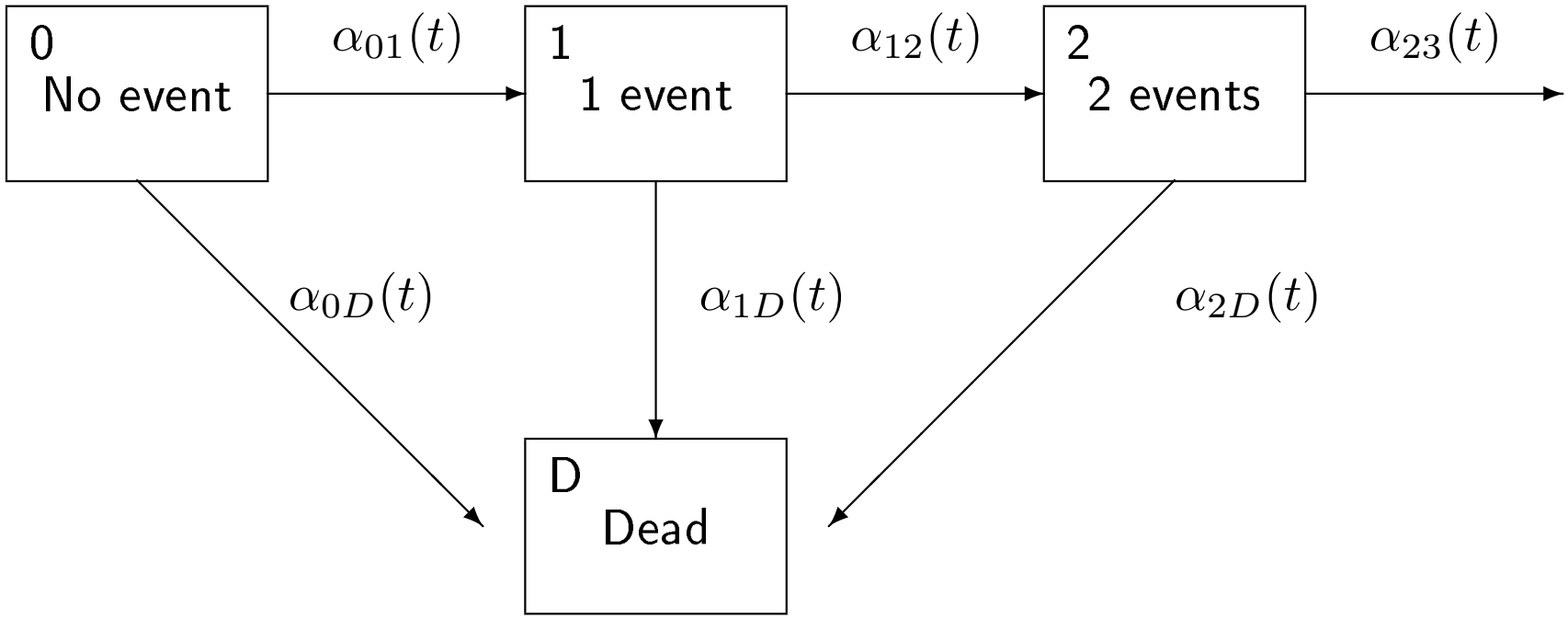
(orginally from Fuchs et al., *NEJM*, 1994; also used by Therneau and Hamilton, 1997, *Statist. in Med.*).

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`
- `trt` (treatment indicator), `meanfev` (at baseline)
- `start`, `stop`, `status`
- `startold` (= `start` of previous episode)
- `etype` (1 if 'at risk', 2 if 'under treatment')
- `enum` (record no.), `enum1` (gap time no.), `enum2` (treatment period no.)

Data set `rhDNase` in the `survival` package of `R` (with a slightly different structure).

Recurrent events and mortality, 'no duration'

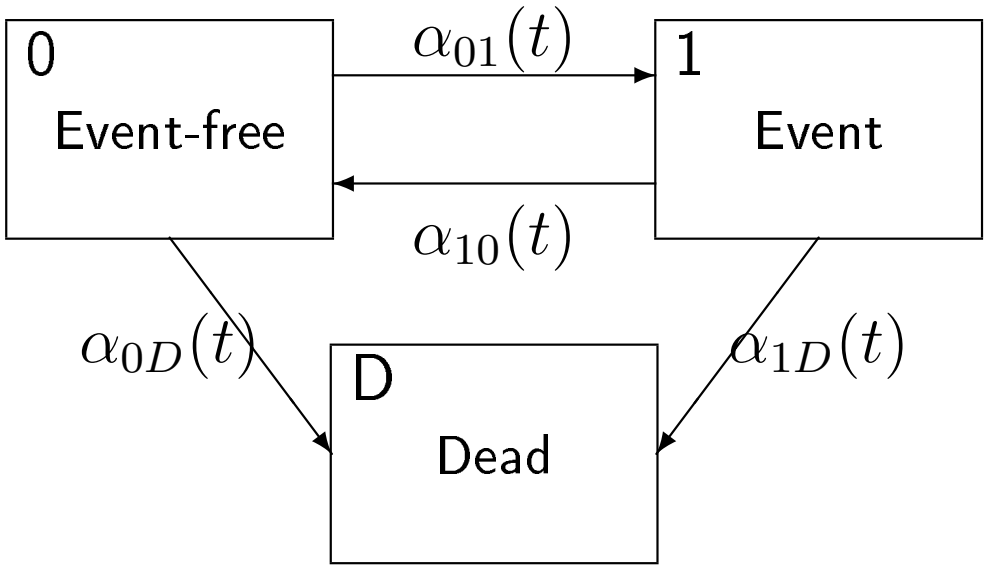


Example: clinical trial in stage I bladder cancer

Trial conducted by the Veterans Administration Cooperative Urological Research Group (Byar, 1980) - famous text book example. Data set `bladder1` in the `survival` package of `R`

- 118 patients with stage I bladder cancer, 30 died during follow-up
- randomized to pyridoxine (32/7), thiotepa (38/12), or placebo (48/11)
- followed for the occurrence of superficial bladder tumors
- Data set includes, among a few other variables:
  - `id`: person-id
  - `enum`: record number per `id`
  - `start`: start time in months
  - `stop`: stop time in months
  - `status`: 0 = alive, no new tumor, 1 = alive, new tumor, 2/3 = dead
  - `number`: number of tumors at time 0
  - `size`: largest tumor at time 0 (*cm*)
  - `treatment`: placebo, pyridoxine or thiotepa.

Recurrent events 'with duration' and mortality



(or similar, 'un-folded' diagram)

Again, all intensities may depend on the past.

Example: Psychiatric admissions (1)

Kessing, Olsen and Andersen, *Amer. J. Epidemiol.* (1999) studied data from the Danish Central Psychiatric Registry *DCPR* which includes:

- All psychiatric admissions 1 April 1970-
- Discharge diagnoses
- ICD-8 until 31 december 1993
- ICD-10 from 1 January 1994

Data were linked to the Danish Civil Registration System *CPR* to obtain information on vital status and emigration.

The 'terminating events' are death or a diagnosis of schizophrenia.

Sample from register data

- All patients identified in DCPR before 1 January 1994 with a bipolar ('manic') diagnosis at first discharge
- Restrict attention to patients younger than 52 years at diagnosis, 1307 men and 1655 women
- Followed to 31 December 1999 w.r.t. later psychiatric admissions, death, diagnosis of schizophrenia or emigration

Purpose of study: Evaluate theory of 'sensitization' (or 'kindling').

According to this theory, mood episodes themselves stress the brain so that its sensitivity to biologic and psychosocial stressors increases. This leads to shorter and shorter intervals between successive episodes.

Example: Psychiatric admissions (2)

Clinical data collected by Swiss psychiatrist Jules Angst (!) in Zürich 1959-62 (Kessing et al., 2004). Here, we restrict attention to prospectively collected data on patients with a first diagnosis after 1958:

- 119 patients with unipolar ('depression', 98) or bipolar ('manic', 21) disorder, 85 died during follow-up (until 1985)
- dates on admission to and discharge from psychiatric hospital
- covariates include
  - sex
  - first diagnosis (unipolar or bipolar)
  - age at first diagnosis
  - year at first diagnosis

Example: Psychiatric admissions - 'cycles'

Patients are (obviously) not at risk for an admission while in hospital.

However, some times, 'cycles' (times from admission to re-admission, i.e. disregarding that some time is spent in hospital) are considered instead of times from discharge to re-admission. That brings the example into the framework of recurrent events with 'no duration'.

Same remarks for the data on pulmonary exacerbations.

A small sidetrack

It should be mentioned that one simple standard way out of a recurrent events problem is to ignore it (!) by restricting attention to the *first* occurrence of the event (maybe in competition with death).

This is, obviously, an *inefficient* solution since it disregards parts of the data that could contain important and useful information.

Simulation studies have been conducted that illustrate the loss of efficiency.

Likelihood

A multi-state model, say $X(t)$ involves different *states*: $h = 1, ..., k$ and *types of direct transition*, $h \to j, h, j = 1, ..., k; h \neq j$.

The *counting process* $N_{hj}(t)$ counts the number of direct $h \to j$ transitions in $[0, t]$. Let $Y_h(t)$ be the number of subjects observed in state h at time $t-$. Then the *intensity process* for $N_{hj}(t)$ is:

$$\alpha_{hj}(t)Y_h(t) \approx P(X(t + dt) = j \mid X(t-) = h, \mathcal{F}_{t-})/dt,$$

for some function $\alpha_{hj}(\cdot)$ of time and past.

According to *Jacod's formula*, the likelihood for the α's based on observation of $(N_{hj}(t), Y_h(t), 0 \leq t \leq \tau)$ (censoring allowed) is

$$\prod_t \prod_{h,j;h\neq j} \left(\alpha_{hj}(t)Y_h(t)dt\right)^{dN_{hj}(t)} \times \exp(-\sum_{h,j} \int_0^\tau \alpha_{hj}(u)Y_h(u)du)$$

$$= \prod_{h,j;h\neq j} \left(\prod_t \left(\alpha_{hj}(t)Y_h(t)dt\right)^{dN_{hj}(t)} \exp(-\int_0^\tau \alpha_{hj}(u)Y_h(u)du) \right).$$

Models for intensities

The likelihood *factorizes* into factors, each depending only on one type of transition ($h \to j$). Furthermore, each factor looks exactly like the likelihood for censored survival data.

This means that, in all of the recurrent events scenarios studied (i.e., with or without duration and with or without mortality), standard hazard-based methods for survival data may be used as intensity models, e.g.:

- The Nelson-Aalen estimator for $A_{hj}(t) = \int_0^t \alpha_{hj}(u)du$,

$$\widehat{A}_{hj}(t) = \int_0^t \frac{dN_{hj}(u)}{Y_h(u)},$$

- The log-rank test for comparing two (or more) intensities
- Various versions of the Cox regression model

A time variable for analysis must be chosen - more below.

Recurrent events, the PWP model

One option is to use separate models (possibly with shared covariate effects) for $\alpha_{h,h+1}(t), h = 0, 1, ...$

This is known as the *PWP model* (Prentice, Williams and Peterson, *Biometrika*, 1981):

$$\alpha_{h,h+1}(t \mid X) = \alpha_{h0}(t) \exp(\beta^{\mathsf{T}} X), h = 0, 1, ...$$

This is a Cox model with *time-dependent strata* and may be fitted using delayed entry for later events.

The effect of X might vary with the number of previous events, i.e. β_h instead of β.

Recurrent events, the AG model

Assuming proportionality between the different $\alpha_{h0}(t)$, a special case of the *AG* model (Andersen and Gill, *Ann. Statist*, 1982) is obtained.

In this model there is a single event intensity:

$$\alpha(t) = \alpha_0(t) \exp(\beta^{\mathsf{T}} X(t))$$

that is allowed to depend on the past (e.g., number of previous events $N(t-)$) via *time-dependent covariates* which may also interact with other covariates.

The model is fitted using delayed entry for later events.

If there are only time-fixed covariates, this is an *inhomogeneous Poisson process* – an example of a *Markov* process. With *no covariates*, the Nelson-Aalen estimator may be used for estimating $A_0(t)$.

Recurrent events, gap time models

Gap time models are models where the baseline intensity depends, not on t, but on time $t - T_{N(t-)}$ since last event, e.g.:

$$\alpha_{h,h+1}(t) = \alpha_{h0}(t - T_h) \exp(\beta_h^{\mathsf{T}} X),$$

where $N(t)$ is the recurrent events counting process. If there are no covariates and successive gap times are i.i.d., this is a *renewal process*.

These models are some times called *semi-Markov*.

To fit the model, no delayed entry is needed: for each transition $h \to h+1$, subjects are at risk from the 'new time zero'.

Putter et al. (*Statist. in Med.*, 2007), in their tutorial on multi-state models, called such models 'clock reset' (in contrast to Markov models which were called 'clock forward').

Example: Rats

AG-type Cox model:

Rats: $\exp(\widehat{\beta}) = 0.442, \quad (0.328, 0.595),$

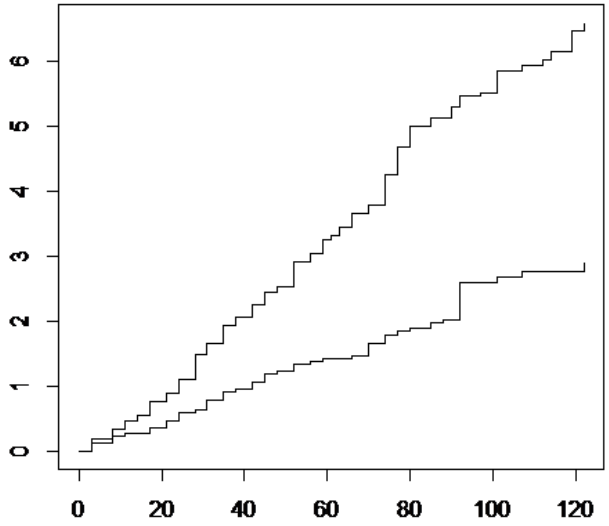
logrank statistic 30.54, $P < 0.001$.

```
coxph(Surv(start,stop,status==1)~trt,
data=rats2,ties='breslow')
```

Nelson-Aalen plot (next page):

```
plot(survfit(Surv(start,stop,status==1)~trt,
data=rats2),fun='cumhaz')
```

Nelson-Aalen plot for rats data



Taking intra-individual dependence into account

Repeated events within the same subject are likely dependent.

This may be accounted for in a number of ways:

- *Explicitly modelling* how the intensity depends on the past
- Using *robust standard errors*
- Using *random effects* ('frailty') models

Random effects ('frailty') models

Possible model for intensity, $\alpha_i(t)$, for subject i:

$$\alpha_{hji}(t \mid Z_i = z_i) = z_i \cdot (\alpha_{hj0}(w) \text{ or } \alpha_{hj0}(t)) \exp(\beta_{hj}^{\mathsf{T}} X_i(t))$$

- $w = t - T_{ih}$ gap time at time t,
- $X_i(t)$: observed explanatory variables
- Z_i: random, unobserved *frailty* assumed to follow some distribution with mean 1 and SD σ across the population, e.g. the gamma distribution.

  The frailty accounts for dependence between successive episodes in each subject.

Example: Rats

PWP model: $\exp(\widehat{\beta}) = 0.586, \quad (0.418, 0.820),$

(stratified) logrank statistic 9.87, $P = 0.002.$

Adjusted AG model: $\exp(\widehat{\beta}) = 0.561, \quad (0.407, 0.773).$
$(\exp(\widehat{\beta}_N) = 1.139, \quad (1.075, 1.207))$

AG model, robust SD: $\exp(\widehat{\beta}) = 0.442, \quad (0.300, 0.652).$

R code

```
coxph(Surv(start,stop,status==1)~trt+strata(obs),
data=rats2,ties='breslow')

coxph(Surv(start,stop,status==1)~trt+obs,
data=rats2,ties='breslow')

coxph(Surv(start,stop,status==1)~trt+cluster(id),
data=rats2,ties='breslow')
```

For the frailty model, add `'frailty(id)'` to model formula.

Psychiatric admissions (1)

A gap time model was used because of the 'renewal features' of the data: When evaluating the tendency to re-admission it is more relevant to consider time since last event than time since diagnosis.

A simple such model (without covariates) would be:

$$\alpha_{hji}(t) = \alpha_{h0}(t - T_{ih})$$

where $h = N_i(t-)$ is the number of episodes before time t since diagnosis.

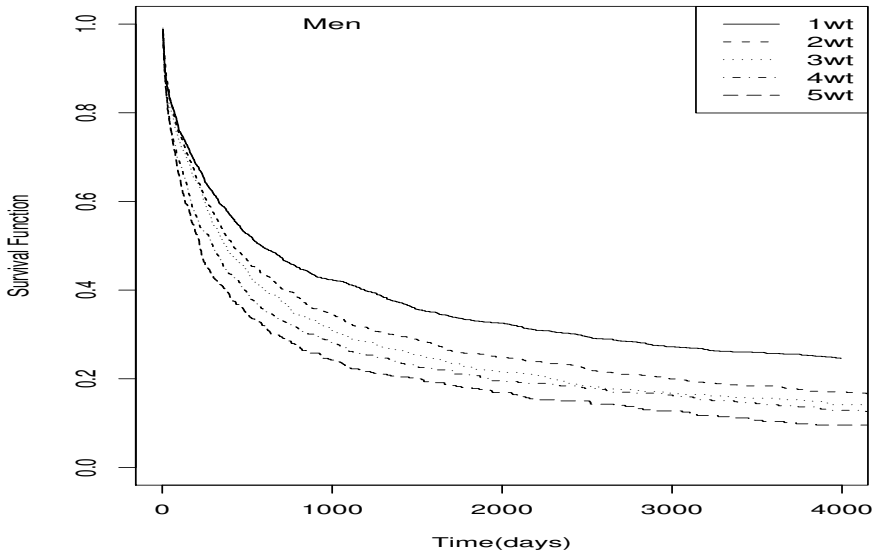
For each patient, we then have a process allowing for different gap time distributions for different numbers of previous episodes (same distribution across patients) and we can try to estimate the survival function:

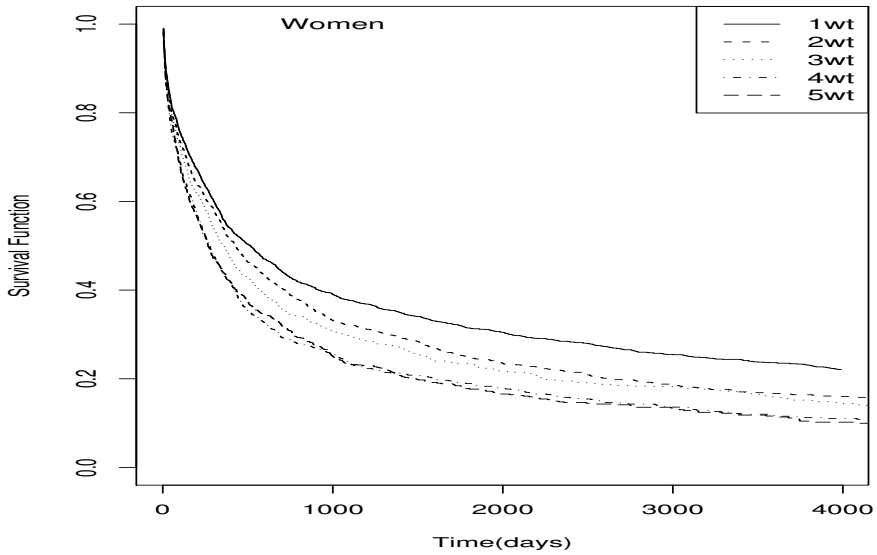
$$S_h(w) = \exp(-\int_0^w \alpha_{h0}(u)du)$$

by the Kaplan-Meier estimator.

Kaplan-Meier estimates

Kaplan-Meier estimates of the survival functions for 1st, 2nd and later gap times.





Using Kaplan-Meier curves

The Kaplan-Meier curves do not properly address the problem of sensitization due to selection/heterogeneity – those with 2 episodes is a select subgroup of those with 1 etc.

Another way of stating the problem is via *dependent censoring*: e.g., censoring of the second gap time depends on the first gap time and if successive gap times are correlated then censoring of the second gap time depends on the second gap time:

U_i: censoring time for subject i, W_{i1}, W_{i2}: first and second gap time for subject i
U_{i2}: censoring time for W_{i2} is $U_{i2} = U_i - w_{i1}$.

If W_{i1} and W_{i2} are dependent, then W_{i2} and U_{i2} are also dependent and the Kaplan-Meier estimator does not work for $S_2(w)$.

Random effects ('frailty') model for younger bipolar patients

	Men		Women	
Episode (j)	Rate ratio	95% c.i.	Rate ratio	95% c.i.
1	1.00	ref.	1.00	ref.
2	1.18	1.03-1.34	1.22	1.08-1.37
3	1.46	1.27-1.69	1.47	1.29-1.67
4	1.72	1.49-2.02	1.61	1.40-1.85
5+	2.35	2.07-2.67	2.19	1.07-2.45
σ^2	0	(no frailty)	0	(no frailty)
1	1.00	ref.	1.00	ref.
2	0.99	0.84-1.15	1.07	0.93-1.22
3	1.10	0.91-1.33	1.16	0.99-1.36
4	1.16	0.93-1.45	1.17	0.98-1.39
5+	1.30	1.04-1.64	1.25	1.04-1.50
σ^2	0.45	0.26-0.63	0.41	0.28-0.54

Joint frailty models

The frailty accounts for unobserved risk factors in the recurrent events process, and inference builds on integrating the frailty out of the likelihood.

If there is a non-negligible mortality and if mortality rates depend on the same frailty then a *joint frailty model* for both the recurrent events process and the mortality rate is needed.

Rondeau et al. (*Biostatistics*, 2007) developed such a model which is implemented in the `R` package `frailtypack`.

See also Liu et al. (*Biometrics*, 2004) and Huang and Wang (*JASA*, 2004).

Joint frailty models

The frailties $Z_1, ..., Z_n$ are assumed to be i.i.d. gamma variates with mean 1 and SD σ. The model is then:

$$\alpha_i(t \mid Z_i = z_i) = z_i \cdot \alpha_0(t) \exp(\beta_1^{\mathsf{T}} X_i(t))$$

for the recurrent events intensity and

$$\alpha_{Di}(t \mid Z_i = z_i) = z_i^{\gamma} \alpha_{D0}(t) \exp(\beta_2^{\mathsf{T}} X_i(t))$$

for the mortality rate. Thus, the frailty is shared but its effect on recurrent events and mortality could be different, modelled by the power γ. The sets of covariates could differ between the two models.

The likelihood is not extremely nice and the authors approximate the baseline intensities by cubic splines and use a penalized likelihood.

Joint frailty models

If there are 'gaps' between successive episodes then one must either

1. assume that their distribution is independent of the frailty Z (psychiatric admissions)

   or

2. model the dependence, e.g., by using $\frac{1}{Z}$ as the frailty for that transition (O'Keefe et al., *Appl. Statist.*, 2018).

The second option is not available in `frailtypack`.

Exercise 1

- Attach the data set `rhdnase-rev.txt` from the course home page and use the subset where `etype==1`
- Estimate, non-parametrically, and plot the cumulative exacerbation intensities in the two treatment groups
- Fit an AG type Cox model for the intensity with treatment as the only covariate
- Add the covariate `meanfev`. What happened to the treatment effect?
- Add, instead, some function of previous exacerbations as covariate (e.g., $I(N(t-) > 0)$). What happened to the treatment effect?
- Fit a PWP type Cox model for the intensity with treatment as the only covariate (i.e., *stratify* on `enum1`)
- Estimate the (unadjusted) treatment effect on the intensity with robust standard errors. What happened to the treatment effect and the confidence interval?
- Estimate the treatment effect on the intensity taking a possible *frailty* into account. What are the assumptions for the frailty distribution?

Randomized trials

Transition *intensities* are the basic parameters in multi-state models and, in principle, all other parameters are functions of these. However, they are *conditional on the past development*.

In a randomized trial the past may, obviously, be influenced by treatment and may, therefore, be *intermediate* between treatment and an ultimate outcome like death or relapse.

In epidemiology, it is well-known that one should avoid conditioning on intermediate variables since this may mask exposure effects, and the same holds true in RCTs.

Marginal parameters for recurrent events

For analysis of *recurrent events*, it has therefore been argued (e.g, by Cook and Lawless, 2007) that intensity-based models that condition on previous events are not optimal when analyzing RCTs, and *marginal* parameters have been put forward instead (cf. rats data).

Marginal parameters in general multi-state models include:

- State occupation probabilities $Q_h(t) = P(X(t) = h)$
- Average time spent in a given state
- Times (from 0) to entry into a given state

  In the special case of recurrent events, one could (also) focus on
- Expected number of events in $[0, t]$: $\mu(t) = E(N(t))$

As we shall see, such parameters may some times be analyzed directly, i.e. without going via the intensities.

Times to entry into given states

For recurrent events, this is the distribution of the event times $T_1, T_2, ...$

For this purpose, the model of Wei, Lin and Weisfeld (*JASA*, 1989) (the 'WLW' model) may be used. This is a series of Cox models for $T_1, T_2, ...$, e.g.,

$$\alpha_h(t \mid X) = \alpha_{h0}(t) \exp(\beta_h^{\mathsf{T}} X)$$

for the hazard function for the *marginal distribution* of time to event number h. Technically, the model is fitted as separate Cox models for each event time pretending that, at any time where a subject is alive and uncensored, he or she is *at risk for all event numbers*, e.g. at risk for the second event before the occurrence of the first.

Intra-individual dependence is treated by using the robust sandwich estimator for the variance of parameter estimates.

A remark on the marginal Cox model

While the WLW model (the 'marginal Cox model') may not be the best model choice for a recurrent events analysis (we will introduce more useful methods below), it is useful for other kinds of 'clustered survival data'.

This could be life times for related individuals (families, litters, ..) and in *cluster-randomized studies*.

By using a marginal model in such cases it is possible to draw inference on the (marginal) distribution of survival times in presence of an intra-cluster correlation and without having to specify the form of the dependence.

This is the general idea behind 'GEE' methods (generalized estimating equations).

Estimating the mean via plug-in in intensity-based models

Let $N_i(t)$ count recurrent events for subject i. We have the Doob-Meyer decomposition:

$$N_i(t) = \int_0^t \alpha(u) Y_i(u) du + M_i(t).$$

Thereby,

$$E(N_i(t)) = \int_0^t \alpha(u) Q(u) du,$$

where

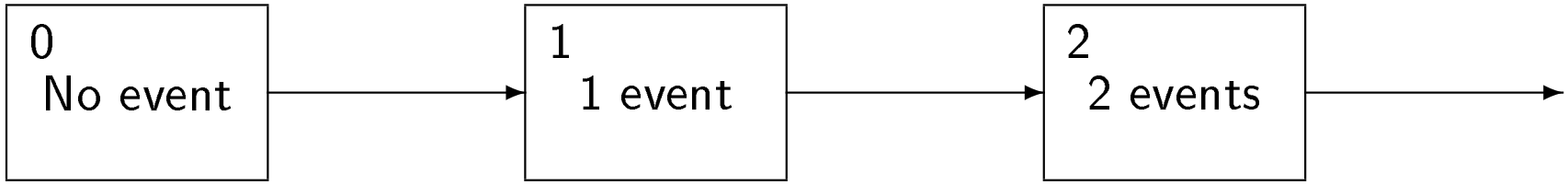
$$Q(t) = E(Y_i(t)) = P(i \text{ at risk at time } t).$$

This implies that the mean may be estimated by combining estimators for $A(t) = \int_0^t \alpha(u) du$ and $Q(t)$, e.g. the Nelson-Aalen and 'Aalen-Johansen' estimators.

Model with no mortality and no gaps

Plug-in methods are, however, not what is being used most in practice. This is probably due to the fact that, like in 'standard survival analyses' simple *non-parametric* (and semi-parametric) methods are available and these are less prone to model mis-specification.

The simplest case is when there are no 'gaps' between events and when mortality is negligible:



Model with no mortality and no gaps

Instead of focussing on the transition intensities, we go directly for $\mu(t) = E(N(t))$ or its derivative $\mu'(t)$ which is some times denoted the *marginal rate function*. It has the interpretation:

$$\mu'(t) \approx E(dN(t))/dt$$

in contrast to the intensity that is conditional on the past

$$\lambda(t) \approx E(dN(t) \mid \mathcal{F}_{t-})/dt.$$

Let $Y_i(t) = I(U_i \geq t)$ be the at-risk indicator for subject i and assume first that censoring U is independent of the recurrent events process $N(t)$. Then the estimating equation:

$$\sum_i Y_i(t)\big(dN_i(t) - d\mu(t)\big) = 0$$

is *unbiased*.

The Nelson-Aalen estimator

The equation leads to the consistent estimator

$$d\widehat{\mu}(t) = \frac{\sum_i dN_i(t)Y_i(t)}{\sum_i Y_i(t)}$$

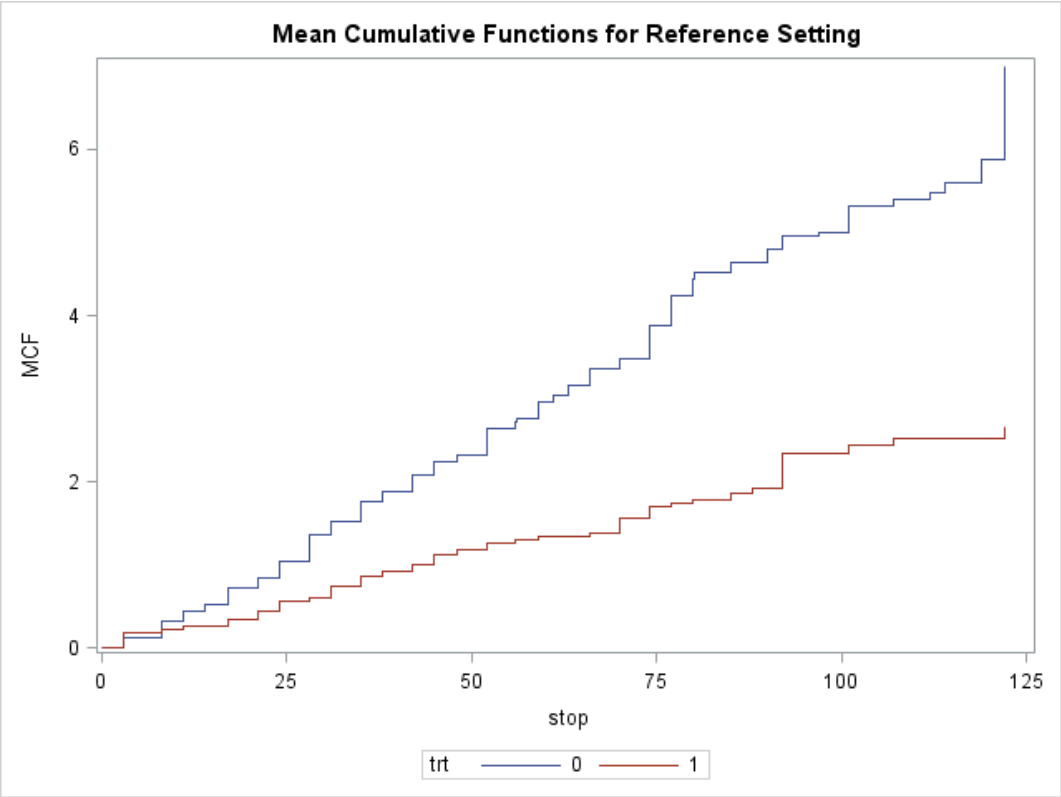
which is simply the *Nelson-Aalen estimator*

$$\widehat{\mu}(t) = \int_0^t \frac{dN(u)}{Y(u)}$$

with $N(t) = \sum_i N_i(t)$ and $Y(t) = \sum_i Y_i(t)$. The estimator was discussed by Lawless and Nadeau (*Technometrics*, 1995) building on earlier work by Nelson. Use robust standard errors.

For the special case where $N(t)$ is an inhomogeneous Poisson process we have $P(dN(t) = 1 \mid \mathcal{F}_{t-}) = P(dN(t) = 1)$ and the mean then equals the cumulative intensity (which now has a nice interpretation!) but it is important to notice that the estimator works without the Poisson assumption.

Tumors in rats



A two-sample test

Lawless and Nadeau (1995) also suggested a two-sample test in the spirit of the logrank test:

$$U(t) = \int_0^t \frac{Y_1(u)Y_2(u)}{Y_1(u) + Y_2(u)} \left(d\widehat{\mu}_1(u) - d\widehat{\mu}_2(u) \right)$$

together with a consistent estimator of the variance of $U(\tau)$.

For the rats data, this gives the value 11.20, $P < 0.001$.

The test is obtainable as a *score* test in the Cox-type model to be described below.

A regression model for $\mu(t)$

One may consider a Cox-type regression model for the mean function $\mu(t)$ (Lawless and Nadeau, 1995; asymptotic results by Lin et al., *JRSS(B)*, 2000):

$$\mu(t \mid X) = \mu_0(t) \exp(\beta^{\mathsf{T}} X)$$

with the baseline mean function $\mu_0(t)$ completely unspecified. A set of unbiased estimating equations is:

$$\begin{aligned} \sum_i Y_i(t)\big(dN_i(t) - d\mu_0(t)\exp(\beta^{\mathsf{T}} X_i)\big) &= 0 \\ \sum_i Y_i(t) X_{ij}\big(dN_i(t) - d\mu_0(t)\exp(\beta^{\mathsf{T}} X_i)\big) &= 0, \quad j = 1, ..., p, \end{aligned}$$

(if there are p covariates). The equations come from a *working Poisson model* but apply more generally.

A regression model for $\mu(t)$

Solving the first equation for fixed β gives the 'Breslow-type' estimator:

$$d\widehat{\mu}_0(t) = \frac{\sum_i dN_i(t)Y_i(t)}{\sum_i Y_i(t)\exp(\beta^{\mathsf{T}} X_i)}$$

which may then be inserted into the second equation(s) to give:

$$U_j(\beta) = \sum_i \int_0^\infty Y_i(t)\big(X_{ij} - \frac{\sum_\ell Y_\ell(t)X_{\ell j}\exp(\beta^{\mathsf{T}} X_\ell)}{\sum_\ell Y_\ell(t)\exp(\beta^{\mathsf{T}} X_\ell)}\big)dN_i(t) = 0$$

which is just the score equation from the Cox-model.

This means that standard 'Cox software' can do the job - just use robust variances, e.g. tumors in rats $\exp(\widehat{\beta}) = 0.442 \quad (0.299, 0.654)$.

Non-independent censoring

When censoring is not completely independent of $N(t)$ the estimating equation (without covariates):

$$\sum_i Y_i(t)\big(dN_i(t) - d\mu(t)\big) = 0$$

is no longer necessarily unbiased and to fix this problem, IPCW may be used.

We need an estimator, $\widehat{G}_i(t)$, of the censoring distribution (distribution of U_i).

This may depend on the past of $N_i(t)$.

Lin et al. also discussed goodness of fit of the model based on cumulative residuals using an i.i.d. decomposition.

Non-independent censoring

It can then be shown that (for the situation without covariates) the estimating equation

$$\sum_i \frac{Y_i(t)}{\widehat{G}_i(t)}\big(dN_i(t) - d\mu(t)\big) = 0$$

is approximately unbiased and leads to the estimator

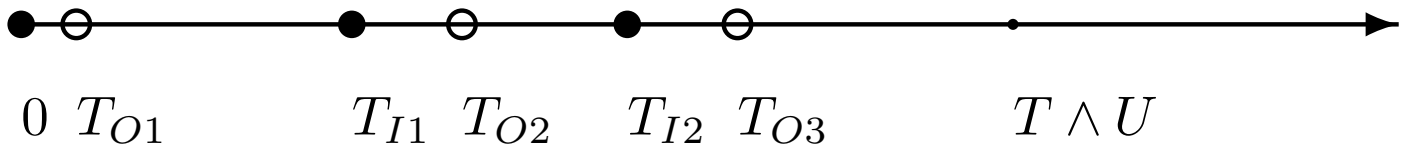
$$d\widehat{\mu}(t) = \frac{\sum_i dN_i(t)Y_i(t)/\widehat{G}_i(t)}{\sum_i Y_i(t)/\widehat{G}_i(t)}.$$

The equations for the Cox-type model may be fixed in a similar way.

In both situations, robust variances are used.

Gaps between 'at-risk' periods: direct estimation of $\mu(t)$

Here, one may consider 'cycles' instead of times from beginning of the at-risk period to new event, e.g., for psychiatric admissions, time from a psychiatric admission to the next, disregarding that there was an admission period where a new admission was not possible



Direct estimation of $\mu(t)$

The Nelson-Aalen (and Ghosh-Lin – see later) estimators estimate the mean number of events in a world where events have a certain duration - a parameter that is of interest even though it would be different in a world where, say, admission intensities were the same but discharge intensities different.

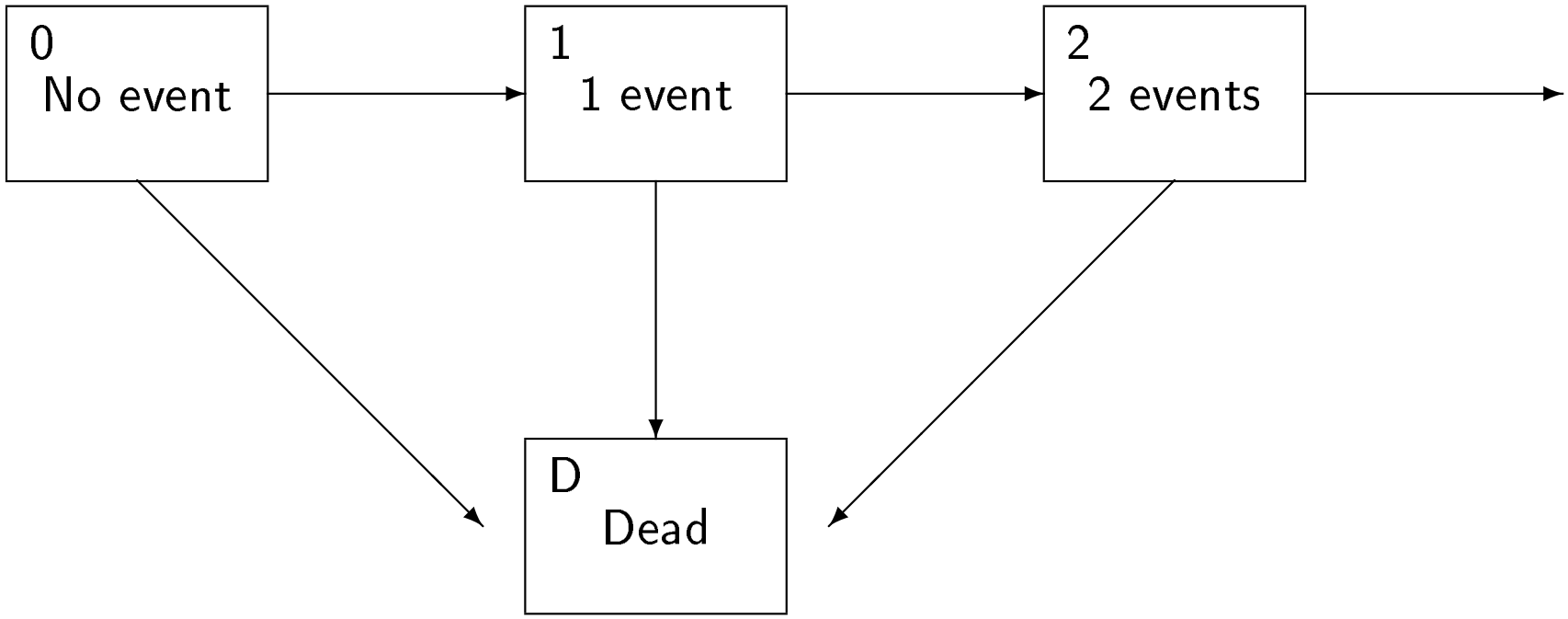
The method works because the estimating equations from earlier are still unbiased for $E(dN(t))$.

The Nelson-Aalen/Ghosh-Lin estimates may be supplemented by estimates of average times in hospital and out of hospital (and of the mortality).

Exercise 2

- Use again the data set `rhdnase-rev.txt` from the course home page and the subset where `etype==1`
- Estimate, non-parametrically, and plot the cumulative mean functions in the two treatment groups (Hint: use `startold` instead of `start`). Compare with the cumulative intensity plot from Exercise 1 and explain the difference
- Fit a 'Cox type model' for the mean function with treatment as the only covariate
- Add the covariate `meanfev`. What happened to the treatment effect?
- Discuss the possibility of including some function of previous exacerbations as covariate

Models with competing risks



Remarks

- The presence of a non-negligible mortality rate complicates the situation
- The resulting model has several similarities with *competing risks*, and we shall see that inference methods for the new model are inspired by those for competing risks
- We shall also see that the same types of bias arise when disregarding mortality in the new model, as those that appear when using '1 - K-M' instead of the correct cumulative incidence estimator with competing risks

An Aalen-Johansen type estimator ('Ghosh-Lin')

When mortality plays a role, the Nelson-Aalen estimator for $\mu(t) = E(N(t))$ will be *upwards biased*. This is because events can only happen as long as the subject is alive and by 'treating death as censoring' we pretend that we have an immortal population where events can happen all the time.

So, we need an estimator for $\mu(t)$ that accounts for mortality and it turns out that a very simple one exists.

Let

$$\rho_i(t) \approx E(dN_i(t) \mid T_i \geq t)/dt$$

be the *marginal rate function given alive* (T_i is the survival time for subject i).

An Aalen-Johansen type estimator

The corresponding cumulative function $R(t) = \int_0^t \rho(u)du$ (assuming that all subjects have the same ρ) can be estimated by the Nelson-Aalen estimator. The function $R(t)$, however, has no nice interpretation but the marginal mean is now simply

$$\mu(t) = E(N(t)) = \int_0^t S(u)dR(u)$$

where $S(\cdot)$ is the survival function (distribution of T) and can be estimated simply by Kaplan-Meier. This leads to the estimator

$$\widehat{\mu}(t) = \int_0^t \widehat{S}(u-)\frac{\sum_i Y_i(u)dN_i(u)}{\sum_i Y_i(u)}.$$

suggested by Cook and Lawless (*Statist in Med.*,1997) and studied in more details (including asymptotic results) by Ghosh and Lin (*Biometrics*, 2000).

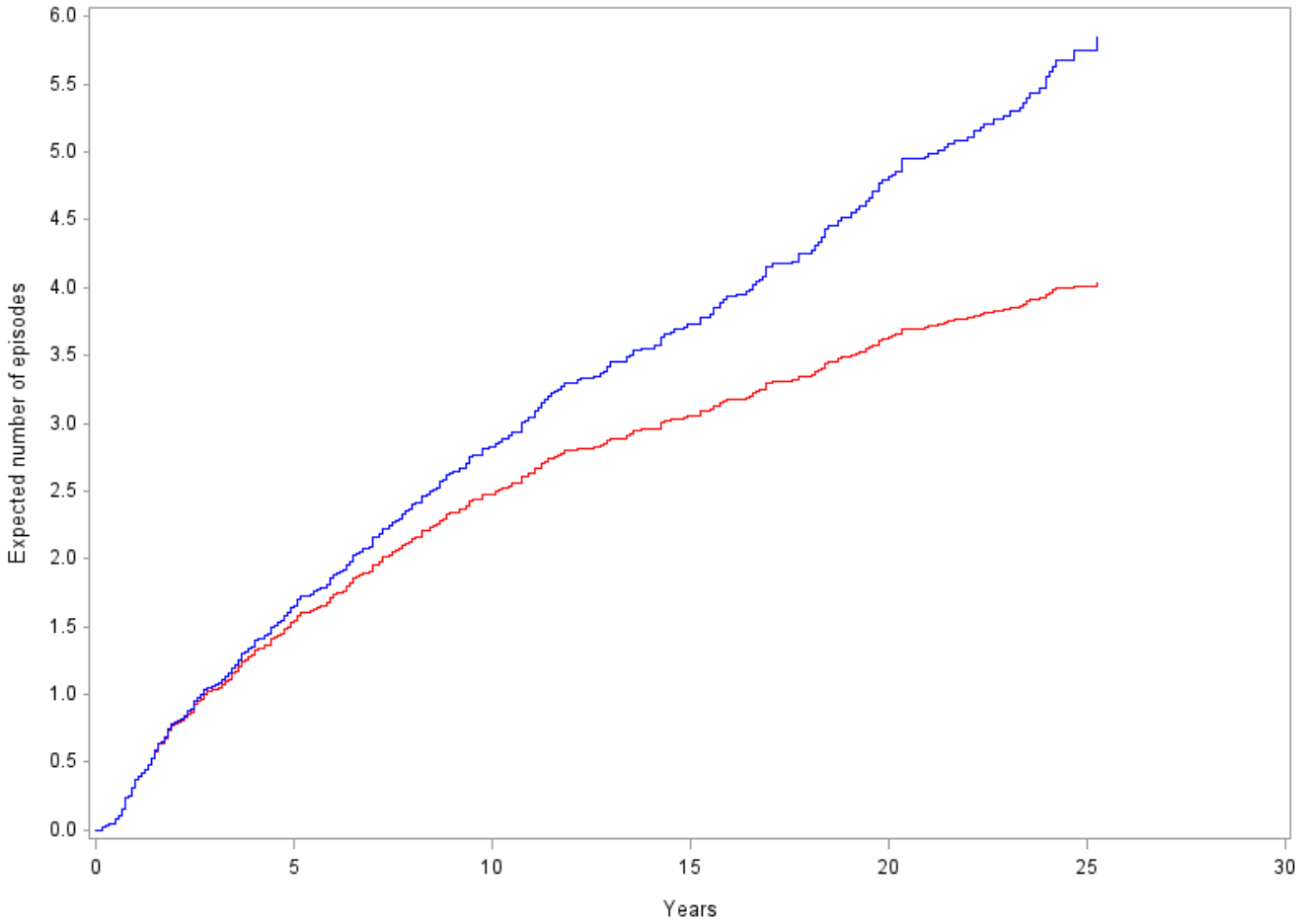
A two-sample test

Ghosh and Lin (2000) also suggested a two-sample test in the spirit of the logrank test (and the Lawless-Nadeau test for cumulative means in the case of no mortality):

$$U(t) = \int_0^t \frac{Y_1(u)Y_2(u)}{Y_1(u) + Y_2(u)} \left(d\widehat{\mu}_1(u) - d\widehat{\mu}_2(u) \right)$$

together with a consistent estimator of the variance of $U(\tau)$.

Next slide shows the (upwards biased) Nelson-Aalen estimate and the (correct) Cook-Lawless/Ghosh-Lin estimate for the cumulative mean function for unipolar patients in the Swiss data set on psychiatric admissions.


Expected number of episodes
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
5
10
15
20
25
30
Years

Regression analysis

In a follow-up paper, Ghosh and Lin (*Statistica Sinica*, 2002) developed a regression model for $\mu(t)$:

$$\mu(t \mid X) = \mu_0(t) \exp(\beta^{\mathsf{T}} X)$$

with asymptotic theory. Estimation follows to a large extent that of the Fine-Gray (1999) model in the sense that they first studied the case where censoring times U_i are known for everybody - even for those who died and, next, an IPCW method was derived. For the first case the estimating equation is:

$$U(\beta) = \sum_i \int_0^\infty \big(X_i - \bar{X}^U(t)\big) I(U_i \geq t) dN_i(t)$$

where $\bar{X}^U(t)$ is the average $\bar{X}^U(t) = \frac{\sum_\ell I(U_\ell \geq t) X_\ell \exp(\beta^{\mathsf{T}} X_\ell)}{\sum_\ell I(U_\ell \geq t) \exp(\beta^{\mathsf{T}} X_\ell)}$.

Regression analysis

In the more common case where censoring times are not known for all, a model for $P(U > t \mid X)$ (possibly without X) is needed. Let $\widehat{G}(t)$ be the estimator for the censoring distribution, e.g. Kaplan-Meier if U is independent of X or a Cox-model-based estimator otherwise.

Ghosh and Lin then introduced the weights

$$w_i(t) = I(U_i \geq T_i \wedge t)\widehat{G}(t)/\widehat{G}(\widetilde{T}_i \wedge t).$$

and showed that:

$$E(w_i(t)) \approx G(t)$$

which is the expectation of $I(U_i \geq t)$ and, therefore, this indicator in the previous estimating equation can be replaced by $w_i(t)$.

Regression analysis

The resulting estimating equation is

$$U(\beta) = \sum_i \int_0^\infty \big(X_i - \bar{X}^G(t)\big) w_i(t) dN_i(t)$$

with

$$\bar{X}^G(t) = \frac{\sum_\ell w_\ell(t) X_\ell \exp(\beta^\mathsf{T} X_\ell)}{\sum_\ell w_\ell(t) \exp(\beta^\mathsf{T} X_\ell)}.$$

Finally, (robust) variances and the baseline mean function can be estimated, the latter by a Breslow-type estimator:

$$\widehat{\mu}_0(t) = \sum_i \int_0^t \frac{w_i(u) dN_i(u)}{\sum_\ell w_\ell(u) \exp(\widehat{\beta}^\mathsf{T} X_\ell)}.$$

Exercise

Consider the Ghosh-Lin estimating equations:

$$U(\beta) = \sum_i \int_0^\infty \big(X_i - \bar{X}^G(t)\big) w_i(t) dN_i(t)$$

in which a weight

$$w_j(t) = I(U_j \geq T_j \wedge t)\widehat{G}(t)/\widehat{G}(\widetilde{T}_j \wedge t)$$

is given to all subjects (j) at an event time t.

There are three kinds of subjects at that time:

1. j is still alive and uncensored at time t
2. j was censored before time t
3. j died before time t

What is the weight $w_j(t)$ in each of these three cases?

Regression analysis - studying both end-points

The model for $\mu(t \mid X)$ may be combined with, e.g. a Cox model for T_i.

Summarizing results from the two models is not obvious - just as summarizing results from Fine-Gray models for two competing causes is not obvious.

However, some way of looking at both end-points of the model is needed

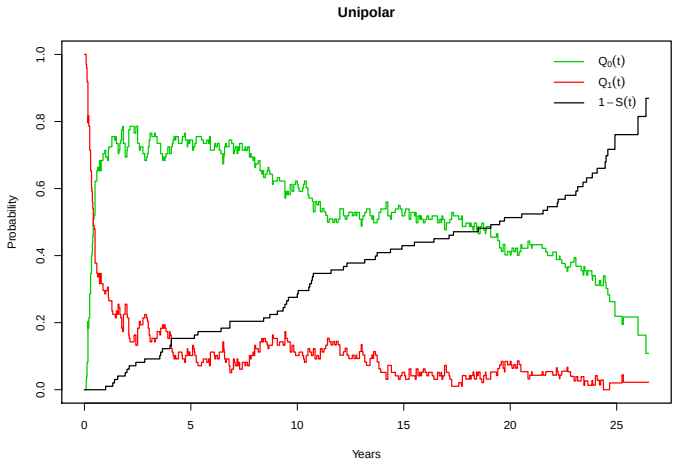
Psychiatric admissions (2)

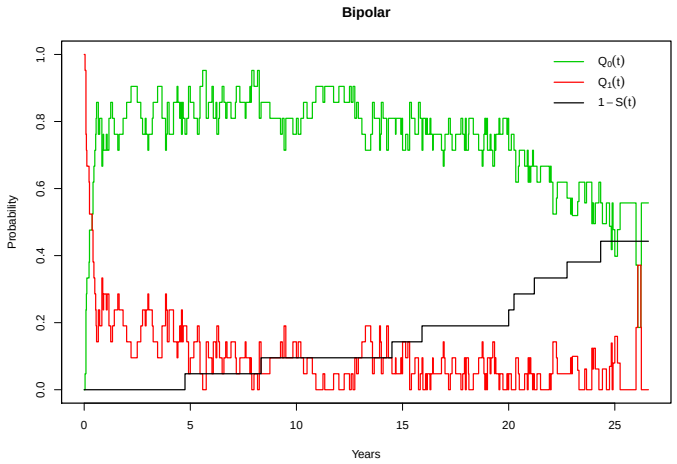
Cox model gives $\exp(\widehat{\beta}) = \exp(0.372) = 1.45 \quad (1.08, 1.94)$ for bipolar vs. unipolar patients.

Adjusted for $N_i(t-)$: $\exp(\widehat{\beta}) = 1.10 \quad (0.83, 1.43)$.

Ghosh-Lin regression model: estimated mean ratio is $\exp(0.673) = 1.96 \quad (1.49, 2.57)$.

State occupation probabilities





Cox model for mortality: estimated hazard ratio is

$$\exp(-0.891) = 0.410(0.204, 0.825)$$

Furthermore, bipolar patients tend to have shorter hospitalisations.

Results from Andersen, Angst and Ravn (2019, *LIDA*)

Using the R package mets

prev is the start of previous episode

```
library(mets)

SURV<-phreg(Surv(prev,slut,dc>1)~strata(bip),data=angst,km=TRUE)
REC<-phreg(Surv(prev,slut,dc==1)~strata(bip),data=angst,km=TRUE)

CMF<-recmarg(REC,SURV)
bplot(CMF)

angst$cens<-ifelse(angst$dc==0,1,0)

recreg(EventCens(prev,slut,dc,cens)~factor(bip)+cluster(proband),
data=angst,cause=1,death.code=2,cens.code=1,cens.model~1)
```

A 'composite end-point'

In the competing risks model, the situation can be simplified by considering a *composite end-point*.

In the model for recurrent events with mortality one can do the same by defining the 'recurrent' event:

event or death

That is, in the simplest case, death is considered as an event 'on an equal footing' as the recurrent event, e.g. if death occurs after the first recurrent event then the death event is regarded as the second 'composite event'.

Adding *weights* has been suggested to signify that death is worse than a new event (Mao and Lin, 2016, *Biostatistics*).

A 'composite end-point': inference

Mao and Lin (2016, *Biostatistics*) showed that inference for this (weighted) composite end-point may follow that for the Ghosh-Lin model.

Thus, if $N_i(t)$ is the weighted composite event counting process for subject i, with an assumed multiplictive marginal mean

$$E(N_i(t)) = \mu_0(t)\exp(\beta^{\mathsf{T}} X_i)$$

then, defining the IPC weights $\quad w_i(t) = I(U_i \geq T_i \wedge t)\widehat{G}(t)/\widehat{G}(\widetilde{T}_i \wedge t)$

the estimating equation

$$U(\beta) = \sum_i \int_0^\infty \big(X_i - \bar{X}^G(t)\big) w_i(t) dN_i(t)$$

is unbiased with

$$\bar{X}^G(t) = \frac{\sum_\ell w_\ell(t) X_\ell \exp(\beta^{\mathsf{T}} X_\ell)}{\sum_\ell w_\ell(t)\exp(\beta^{\mathsf{T}} X_\ell)}.$$

Joint models for marginal mean and mortality

Some approaches deserve mentioning (though the methods are not yet much in use).

- Ghosh and Lin (*Biometrics*, 2003) studied joint AFT models for survival and the recurrent events counting process:

  $$(D_i \exp(-\eta^{\mathsf{T}} X_i), N_i(t \exp(\theta^{\mathsf{T}} X_i))) =_d (D_0, N_0(t)), i = 1, \ldots, n, \text{ i.i.d}$$

  (where the bivariate distribution is unspecified) and set up unbiased estimating equations.

- Ye, Kalbfleisch and Schaubel (and Gong) (*Biometrics*, 2007 (and 2013)) studied a joint frailty model for survival and the marginal rate function given survival:

$$\lambda_D(t \mid Z) = Z \cdot \lambda_{0D}(t) \exp(\alpha^{\mathsf{T}} X), \quad \rho(t \mid Z) = Z \cdot \rho_0(t) \exp(\beta^{\mathsf{T}} X)$$

with $Z_i, \ldots, Z_n$ i.i.d. gamma with mean 1 and SD σ and set up unbiased estimating equations.

Joint models for marginal mean and mortality

- Cook et al. (*JASA*, 2009) discussed an alternative ('plug-in') approach to regression analysis of $\mu(t)$ - very much similar to estimating the cumulative incidence with competing risks by plugging-in models for the cause-specific hazards.

  Recall that $\mu(t) = E(N(t)) = \int_0^t S(u)dR(u)$ which means that, given covariates X, we have:

  $$\mu(t \mid X) = E(N(t) \mid X) = \int_0^t S(u \mid X)dR(u \mid X)$$

  and a regression model for $\mu(t)$ may be obtained by combining regression models for $S(\cdot)$, e.g. a simple Cox model, and for the marginal rate, given survival $\rho(\cdot) = dR(\cdot)$, e.g. an AG-type model as discussed above.

  Interpretation suffers from the same problems as those faced in competing risks.

Summary: Two approaches to recurrent events analysis

- *Intensity-based models* are much in line with hazard models for survival data and the same software may be used.
- Methods are applicable both with and without competing risks and both with and without events having a certain duration. Censoring may depend on covariates.
- Intra-individual dependence needs to be addressed and the models are more vulnerable to model-mis-specification.
- *Marginal models* directly target key parameters such as the expected number of events.
- Special attention is needed to address competing risks and covariate-dependent censoring.
- Intra-individual dependence need not be specified and the models are more robust.

Exercise 3

- Attach the data set `bladder1` from the `survival` package and use the subset where treatment is thiotepa or placebo
- Estimate, non-parametrically, and plot the cumulative tumor intensities in the two treatment groups
- Fit an AG type Cox model for the intensity with treatment as the only covariate
- Fit a Cox model for the hazard of death with treatment as the only covariate
- Estimate non-parametrically the cumulative mean function in the two treatment groups using the `recmarg` function in the `mets` package
- Estimate the treatment effect on the mean function in a Ghosh-Lin regression model using the `recreg` function in the `mets` package

References (in order of appearance)

RJ Cook, JF Lawless (2007). Springer Book.

MH Gail, TJ Santner, CC Brown (1980). *Biometrics* **36**, 255-66.

HJ Fuchs et al. (1994). *N. Engl. J. Med.* **331**, 637-42.

TA Therneau, SA Hamilton (1997). *Statist. Med.* **16**, 2029-47.

DP Byar (1980). In: Plenum Book.

L Kessing, EW Olsen, PK Andersen (1999). *Amer. J. Epidemiol.* **149**, 404-411.

LV Kessing, MG Hansen, PK Andersen, J Angst (2004). *Acta Psych. Scand.* **109**, 339-44.

RL Prentice, BJ Williams, AV Peterson (1981). *Biometrika* **68**, 373-9.

PK Andersen, RD Gill (1982). *Ann. Statist.* **10**, 1100-20.

H Putter, M Fiocco, R Geskus (2007). *Statist. Med.* **26**, 2389-430.

V Rondeau et al. (2007). *Biostatistics* **8**, 708-21.

L Liu, RA Wolfe, X Huang (2004). *Biometrics* **60**, 747-56.

Y Huang, MC Wang (2004). *J. Amer. Statist. Assoc.* **99**, 1153-65.

AG O'Keefe, L Su, VT Farewell (2018). *J. Roy. Statist. Soc. C* **67**, 841-60.

LJ Wei, DY Lin, L Weisfeld (1989). *J. Amer. Statist. Assoc.* **84**, 1065-73.

JF Lawless, JC Nadeau (1995). *Technometrics 37*, 158-68.

References (ctd.)

DY Lin, LJ Wei, I Yang, Z Ying (2000). *J. Roy. Statist. Soc. B* **62**, 711-30.

RJ Cook, JF Lawless (1997). *Statist. Med.* **16**, 911-24.

D Ghosh, DY Lin (2000). *Biometrics* **56**, 554-62.

D Ghosh, DY Lin (2002). *Stat. Sinica* **12**, 663-88.

JP Fine, RL Gray (1999). *J. Amer. Statist. Assoc.* **94**, 496-09.

PK Andersen, J Angst, H Ravn (2019). *Lifetime Data Anal.* **25**, 681-695.

L Mao, DY Lin (2016). *Biostatistics* **17**, 390-03.

D Ghosh, DY Lin (2003). *Biometrics* **59**, 877-85.

Y Ye, JD Kalbfleisch, DE Schaubel (2007). *Biometrics* **63**, 78-87.

JD Kalbfleisch, DE Schaubel, Y Ye, Q Gong (2013). *Biometrics* **69**, 366-74.

RJ Cook, JF Lawless, L Lakhal-Chaieb, KA Lee (2009). *J. Amer. Statist. Assoc.* **104**, 60-75.