

Piece-wise exponential models

Philipp Kopper

02 June 2020

Data preprocessing

As outlined in the previous section, the **PEM data structure** deals with the survival problem as a count data problem. The main feature of this representation is the partition of the follow-up time into a finite number of intervals. Within these intervals hazard rates are assumed to be piece-wise constant.

Survival times & Events

Below, I show how to transform regular survival data into PEM format where the intervals are equidistant and of size 0.4. Still, it should be chosen with great care. The data set contains a case id, a variable indicating the observed event time (obs_times) and the observed event (status) As usual in survival analysis, the status 1 refers to the event while 0 to a censoring. A priori the selected interval cut points defining the baseline hazard in each interval are arbitrary.

```
library(pammtools)

## The add_term function has changed in pammtools 0.1.12.
##   The argument `relative` is deprecated and replaced by `reference`.
##   See help page for details.

##
## Attaching package: 'pammtools'

## The following object is masked from 'package:stats':
##
##   filter

library(tidy)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

data_set <- data.frame(id = 1:3, obs_times = c(1, 0.5, 2), status = c(0, 1, 1))
data_set

##   id obs_times status
## 1  1      1.0      0
```

```
## 2 2      0.5      1
## 3 3      2.0      1

ped <- as_ped(data = data_set, Surv(obs_times, status) ~ ., id = "id",
              cut = seq(0, max(data_set$obs_times), 0.4))
ped
```

##	id	tstart	tend	interval	offset	ped_status
## 1	1	0.0	0.4	(0,0.4]	-0.9162907	0
## 2	1	0.4	0.8	(0.4,0.8]	-0.9162907	0
## 3	1	0.8	1.2	(0.8,1.2]	-1.6094379	0
## 4	2	0.0	0.4	(0,0.4]	-0.9162907	0
## 5	2	0.4	0.8	(0.4,0.8]	-2.3025851	1
## 6	3	0.0	0.4	(0,0.4]	-0.9162907	0
## 7	3	0.4	0.8	(0.4,0.8]	-0.9162907	0
## 8	3	0.8	1.2	(0.8,1.2]	-0.9162907	0
## 9	3	1.2	1.6	(1.2,1.6]	-0.9162907	0
## 10	3	1.6	2.0	(1.6,2]	-0.9162907	1

The offset reflects information on the observed survival time. Re-arranging the Poisson **likelihood** of the PEM one can derive the following definition: