Time-Dependent Survival Analysis

Weisi Chen

2025-03-13

Table of contents


```
# Load the needed packages
library(ggplot2)
library(dplyr)
library(lubridate)
library(survival)
library(ggsurvfit)
library(gtsummary)
library(here)
library(survminer)
library(broom)
library(forestploter)
library(tidyr)
```

```
# Load the data
data(BMT, package="SemiCompRisks")

head(BMT[, c("T1", "delta1", "TA", "deltaA")])
```

```
#>     T1 delta1   TA deltaA
#> 1 2081      0   67      1
#> 2 1602      0 1602      0
#> 3 1496      0 1496      0
#> 4 1462      0   70      1
```

```
#> 5 1433       0 1433       0
#> 6 1377       0 1377       0
```

Time-dependent covariate approach

This is used when if the value of a covariate is changing over time.

About the sample data

The data comes from the `pbc` and `pdcseq` dataset, available from the *survival* package.

The `pdc` dataset contains baseline data and follow-up status for a set of subjects with primary biliary cirrhosis, while the `pdcseq` dataset contains repeated laboratory values for those subjects.

Some important variables: `status`: status at endpoint, 0/1/2 for censored, transplant, dead `albumin`: serum albumin (g/dl) `ascites`: presence of ascites `bili`: serum bilirunbin (mg/dl) `protime`: standardised blood clotting time `stage`: histologic stage of disease (needs biopsy)

```
# Load the data
pbc <- survival::pbc
pbcseq <- survival::pbcseq
```

Organize the dataset structure for time-dependent survival analysis

```
# baseline info
pbc <- subset(pbc, select=c(id:sex,stage)) |>
  mutate(sex_f = case_when(sex == "f" ~ "Female",
                           sex == "m" ~ "Male"),
         sex_f = factor(sex_f, levels=c("Male", "Female")),
         stage_f = factor(stage))

# set range
pbc2 <- tmerge(pbc, pbc, id=id, death = event(time,status))
# add time-varying covariates
pbc3 <- tmerge(pbc2, pbcseq, id=id, ascites = tdc(day, ascites),
               bili = tdc(day, bili), albumin = tdc(day, albumin),
               protime = tdc(day, protime), alk.phos = tdc(day, alk.phos))
```

Characteristic	HR[1]	95% CI[1]	p-value
log(bili)	3.46	2.86, 4.18	<0.001
log(protime)	53.7	22.9, 126	<0.001

[1]HR = Hazard Ratio, CI = Confidence Interval

Fit the time-dependent cox model

```
tbl_regression(coxph(Surv(tstart, tstop, death == 2) ~ log(bili) + log(protime), data = pbc3)
```

Interpretation: For one-unit increase in log(bili), the hazard of death increases by a factor of 3.46, holding other covariates constant.

-> A 2.718 times increases in serum bilirunbin level, hoding all other covariates constant, increases the hazard of death by a factor of 3.46