Stanford Dataset (Time-Dependent Exposure)

Teeranan Pokaprakarn; Hiroto Udagawa; Danyi Xiong; Nusrat Rabbee (rabbee@berkeley.edu)

Department of Statistics, University of California, Berkeley, CA

Stanford dataset (Time-Dependent Example)

In this section, we partially used Method 1 to simulate the Stanford heart transplant dataset [1] with time-dependent exposure (transplant) of N = 103 subjects and calculate statistical power from the Cox regression model. The time unit in this study is days. We do not consider minimum follow-up time and minimum post-exposure follow-up time in our simulation, as this is the case in the actual dataset (JASA). We make the following estimates from the original dataset:

```
#Exposure rate calculation
exprate<-mean(jasa$transplant)

tdata <- jasa[, -(1:4)]
tdata$futime <- pmax(.5, tdata$futime)
indx <- with(tdata, which(wait.time == futime))
tdata$wait.time[indx] <- tdata$wait.time[indx] - .5
sdata <- tmerge(tdata, tdata, id=1:nrow(tdata),
                death = event(futime, fustat),
                trans = tdc(wait.time))

# getting parameter estimates from jasa (td case)
sfit_td<-survfit(Surv(tstart,tstop,death)~factor(trans),data=sdata)
mod_td<-coxph(Surv(tstart,tstop,death)~factor(trans),data=sdata)
beta_td<-mod_td$coefficients
```

Table 1: Estimates from Stanford time-dependent dataset

Effective Exposure Prop	Beta	Unexposed Median Survival Time	Exposed Median Survival Time
0.6699029	0.125151	148	89

Table 1: Estimates from Stanford time-dependent dataset continued

Overall Events	Unexposed Events	Exposed Events
75	30	45

Further exploratory research on JASA leads to the following summary of follow-up time and censoring:

Table 2: Summary of follow-up time for exposed group

Min.	1st Qu.	Median	3rd Qu.	Max.
4	71	206	619	1799

Table 3: Proportion of censored subjects in each follow-up time interval (exposed group)

[4,71)	[71,206)	[206,619)	[619,1799]
0.0588235	0.1764706	0.5294118	0.6111111

Table 4: Summary of follow-up time for unexposed group

Min.	1st Qu.	Median	3rd Qu.	Max.
0	5	20	46.5	1400

Table 5: Proportion of censored subjects in each follow-up time interval (unexposed group)

[0,5)	[5,20)	[20,47)	[47,1400]
0	0.125	0.1111	0.2222

We considered the power to detect a log hazard ratio ($\hat{\beta}$) different than zero through the Cox regression model if the data arose from a similar dataset. Since the estimated log hazard ratio ($\hat{\beta}$) in the actual JASA data set is 0.12515 (Table 1), we expect power to be at or below $\alpha = 0.025$. In order to detect more $\hat{\beta}$ that are significant and increase the power, we add more flexibility in data generation.

Data Generation: We considered an 1800-day study and let all subjects enter the study at time 0. We generate exposure status by using Binomial distribution probability equals 0.6699029, which was estimated from JASA. Follow-up times are generated separately for exposed and unexposed group from a piecewise exponential distribution approach based on Table 2 and 4. Parameters for follow-up time generation are displayed in Table 6 below.

Table 6: Parameters of piecewise exponential distribution (time-dependent case)

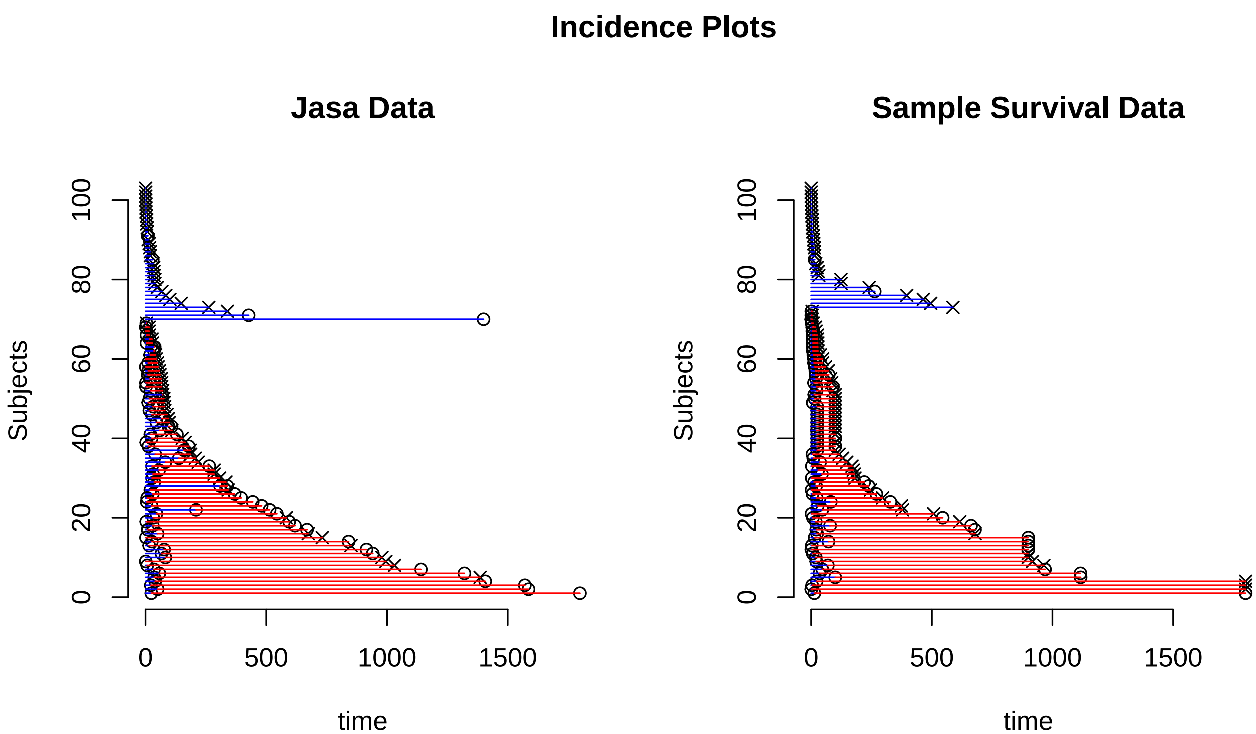
follow-up time <= 206 (rate1)	206 < follow-up time <= 900 (rate2)	follow-up time > 900 (rate3)
0.0123109	0.0020593	0.0007962

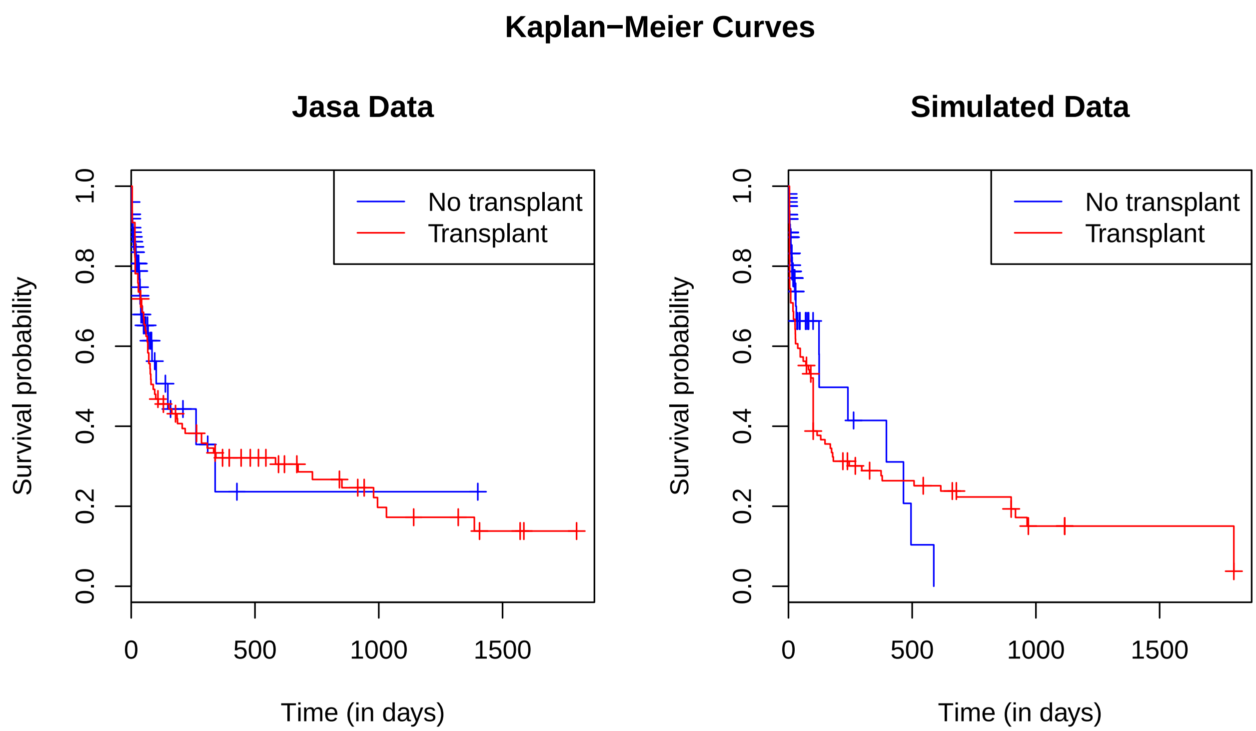
Table 6: Parameters of piecewise exponential distribution (time-dependent case) continued

follow-up time <=20 (rate4)	follow-up time > 20 (rate5)
0.1374046	0.0051249

As for event status, we initialize all subjects' status to be 1. For exposed subjects, switch their status to censored according to Table 3. For unexposed subjects, switch their status to censored according to Table 5. For each exposed subject, his/her exposure time is generated as the minimum of exponential distribution with rate6 equals 0.02654357 (estimated from waiting time (wait.time in JASA)) and his/her follow-up time/4, in purpose of making that subject getting exposed quickly after entering the study.

Monte Carlo simulations: We repeatedly simulated the data 500 times. We compare our results with jasa in following ways.






```
ret<-getpower.stanford(N=103,duration=1800,b=beta_td,r1=rate1,r2=rate2,r3=rate3,r4=rate4,r5=rate5,
                       er=exprate,type="td",scenario=" ",s1=1,s2=4,fA=0,fB=0,
                       filename="simResults_td.csv",simu.plot=FALSE,r6=rate6)


# repeat for N = 1030
ret2<-getpower.stanford(N=1030,duration=1800,b=beta_td,r1=rate1,r2=rate2,r3=rate3,r4=rate4,r5=rate5,
```

```
er=exprate,type="td",scenario=" ",s1=1,s2=4,fA=0,fB=0,
filename="simResults_fixed.csv",simu.plot=FALSE,r6=rate6)
```

Table 8: Estimates from simulated data

Effective Exposure	Beta Hat	Unexposed Median Survival Time	Exposed Median Survival Time
0.6699029	0.125151	152.7376	72.21829

Table 8: Estimates from simulated data continued

Overall Events	Unexposed Events	Exposed Events	Power
82.07	31.016	51.054	0.046

We repeated the simulation for N = 1030 and have the follow estimates:

Table 9: Estimates from simulated data

Effective Exposure	Beta Hat	Unexposed Median Survival Time	Exposed Median Survival Time
0.6699029	0.125151	156.3287	81.85288

Table 9: Estimates from simulated data continued

Overall Events	Unexposed Events	Exposed Events	Power
791.41	297.818	493.592	0.462

As we expect, after increasing sample size from 103 to 1030, we have approximately 10 times more events in both groups. We also detect approximately 10 times more $\hat{\beta}$ that are significant, which yields a 10 times higher power.

References

[1] T. Therneau and C. Crowson (2015). Using Time Dependent Covariates and Time Dependent Coefficients in the Cox Model. https://cran.rproject.org/web/packages/survival/vignettes/timedep.pdf