Survival HW4

David Coomes

Due March 9, 2020

Question 1

(a)

Table 1. Time ratios for Accelerated Failure Time model

	Estimate	95% CI Lower	95% CI Upper
dose	1.025	1.016	1.034
clinic	2.032	1.493	2.765
prison	0.795	0.627	1.007

My results from the AFT model agree qualitatively with my results from the Cox PH model. In the AFT model we see that there is a slight (and significant) increase in time to exit from maintenance associated with an increase in methadone dose, which corresponds to a slight decrease in HR in the Cox model, adjusting for clinic and history of previous incarceration for both models. Similarily, we see a large increase in time to exit associated with clinic 2 as compared to clinic 1, and a large reduction in HR. Finally, there is a small (and non-significant) decrease in time to exit from maintenance associated with a history of previous incarceration, which corresponds to an increase in the HR (also non-significant).

When using a generalized gamma baseline distribution, the dose time ratio estimate decreases by 0.68%, the clinic time ratio increases by 0.79%, and the incarceration time ratio increases by 4.91%.

(b) The ratio of mean time to exit comparing an individual from clinic 2 without a history of incarceration and administered a dose of 40 mg/day to individuals from clinic 1 with a history of incarceration and administered a dose of 100 mg/day is **0.590 (95% CI: 0.305-1.143).**

The median time until exit for a person from clinic 2 without a history of incarceration who is administered a dose of 40 mg/day is **509 days (95% CI: 334-684).** The median time unitl exit for a person from clinic 1 with a history of incarceration who is administered a dose of 100 mg/day is **862 days (95% CI: 511-1213).**

(c) The time ratio comparing patients who were administered a dose of 80 mg/day to those administered 60 mg/day *with* a history of previous incarceration and from the same clinic is **1.507 (95% CI: 1.135-2.002).** The time ratio comparing patients who were administed a dose of 80 mg/day to those administered 60 mg/day *wihout* a history of previous incarceration and from the same clinic is **1.712 (95% CI: 1.362-2.152).** These two subgroup-specific time ratios are not significantly different from each other **(p=0.46)**.

Question 2

(a) **Figure 1. Kaplan Meier survival curve showing age at blindness due to diabetic retinopathy in untreated eyes, accounting for delayed entry.** The median age at blindness is 24.4 (95% CI: 20.5-26.5).

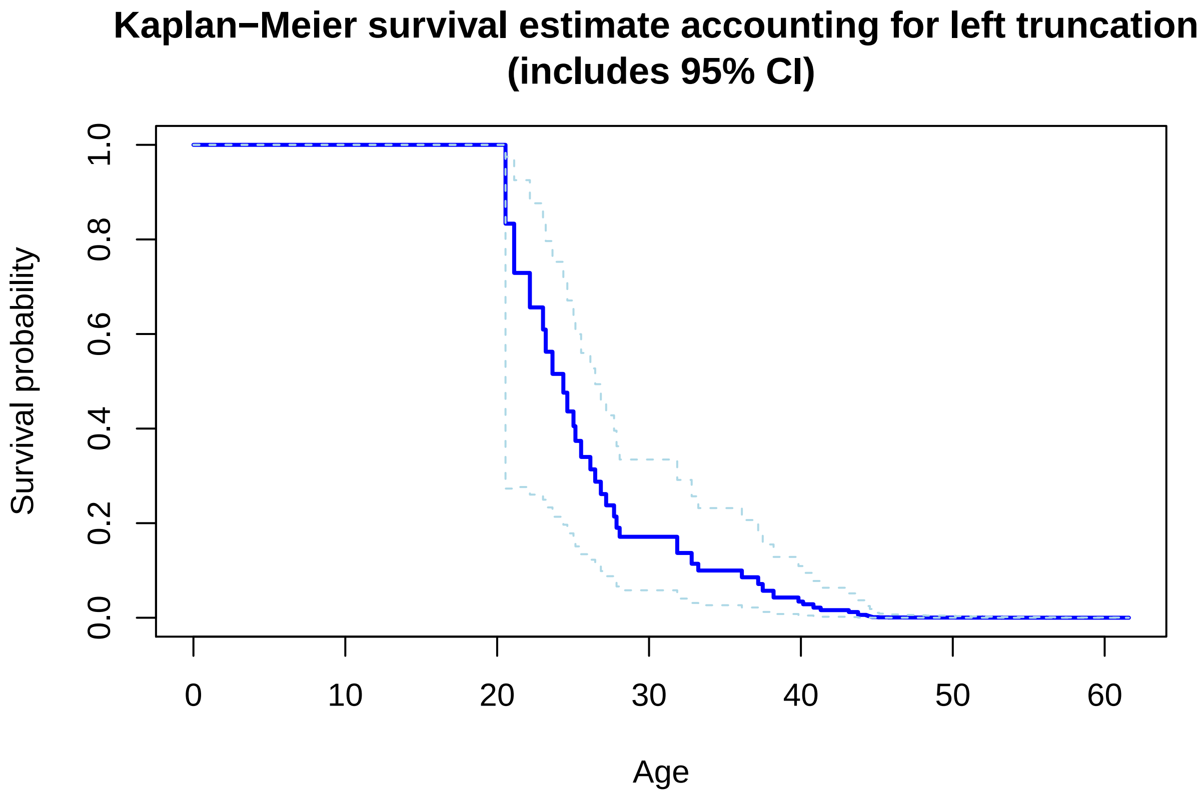
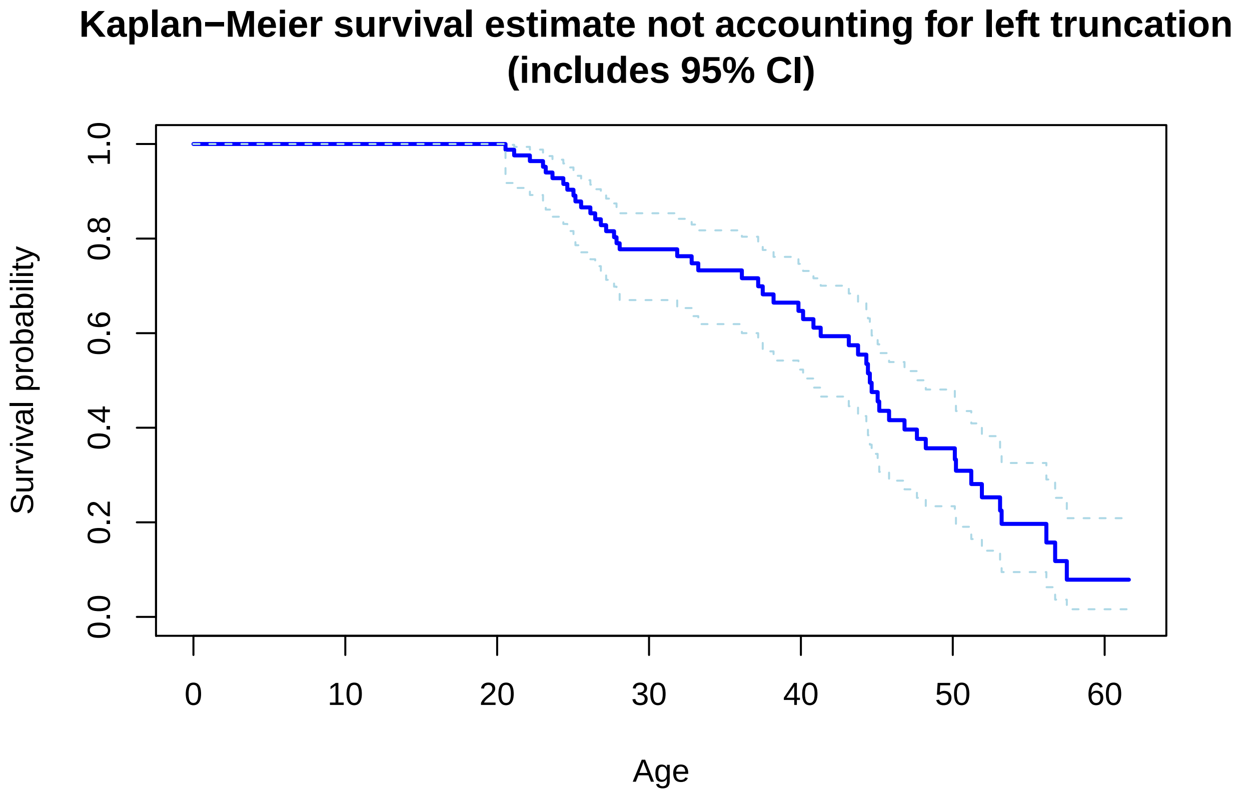


Figure 2. Kaplan Meier survival curve showing age at blindness due to diabetic retinopathy in untreated eyes *without* accounting for delayed entry. The median age at blindness is 44.5 (95% CI: 40.8-48.2).



(b) The results that estimate age at blindness due to DB while accounting for left-truncation do cause some concern for me. Of particular concern is the steep reduction in the survival curve at age 20, and the fact that the survival falls to 0 by age 45 despite the fact that there are some individuals in our data set that do not join the study until after this age. This causes me to wonder whether the age at blindness may be associated with the age at onset of DB - which would violate the independence assumption required to valid delayed entry adjustment.

Appendix

```r
knitr::opts_chunk$set(echo = FALSE, warning=FALSE)

if (!require(kableExtra)) install.packages("kableExtra")
if (!require(flexsurv)) install.packages("flexsurv")
if (!require(survMisc)) install.packages("survMisc")
if (!require(msm)) install.packages("msm")
if (!require(tidyverse)) install.packages("tidyverse")
if (!require(tableone)) install.packages("tableone")
if (!require(compareGroups)) install.packages("compareGroups")
if (!require(table1)) install.packages("table1")


library(tidyverse)
library(knitr)
library(kableExtra)
#source("getmedianres.R")
library(survival)
library(flexsurv)
library(survMisc) # for Wilcoxon-Gehan-Breslow test
library(msm)
library(tableone)
library(compareGroups)
library(table1)



link = "https://github.com/dmccoomes/Survival/raw/master/Homework%201/addicts.csv"
adix <- read.csv(link)

surv.adix <- Surv(time=adix$time, event=adix$event, type="right")
survfit.adix <- survfit(surv.adix ~ 1, data=adix, conf.type = "log-log")



weibull.adix.1 <- flexsurvreg(surv.adix ~ dose + clinic + prison, data=adix, dist="weibull")
#weibull.adix.1

weibull.adix.est <- round(exp(weibull.adix.1$res[3:5, 1:3]), 3)

weibull.adix.est %>%
  kable(col.names = c("Estimate", "95% CI Lower", "95% CI Upper")) %>%
  kable_styling(full_width = F,
                position="left")


cox.1 <- coxph(surv.adix ~ dose + clinic + prison, data=adix)
cox.1


gengamma.adix.1 <- flexsurvreg(surv.adix ~ dose + clinic + prison, data=adix, dist="gengamma")
#gengamma.adix.1
```

```
gengamma.adix.est <- round(exp(gengamma.adix.1$res[4:6, 1:3]), 3)
#gengamma.adix.est

perc_diff <- (gengamma.adix.est[,1]-weibull.adix.est[,1])/weibull.adix.est[,1]

perc_diff


AFT.point <- (40-100) * coef(weibull.adix.1)["dose"] +
                  (2-1) * coef(weibull.adix.1)["clinic"] +
                  (0-1) * coef(weibull.adix.1)["prison"]

AFT.se <- deltamethod(g = ~(40-100) * x1 + (2-1) * x2 + (0-1) * x3,
                      mean = coef(weibull.adix.1)[c("dose", "clinic", "prison")],
                      cov=vcov(weibull.adix.1)[c("dose", "clinic", "prison"),
                                               c("dose", "clinic", "prison")],
                      ses=TRUE)

#point estimate
exp(AFT.point)

#confindene interval
exp(c(AFT.point - 1.96*AFT.se,
      AFT.point + 1.96*AFT.se))


#median for person from clinic 2 on 40 mg/day with no history of incarceration
AFT.median.1 <- exp(40 * coef(weibull.adix.1)["dose"] +
                      2 * coef(weibull.adix.1)["clinic"]) *
        log(2) ^ (1/exp(coef(weibull.adix.1)["shape"])) *
        exp(coef(weibull.adix.1)["scale"])

AFT.median.1.se <- deltamethod(g = ~exp(40*x1 + 2*x2)
                               * log(2) ^ (1/exp(x4)) * exp(x5),
                               mean = coef(weibull.adix.1)[c("dose", "clinic", "prison", "shape", "scal
                               cov=vcov(weibull.adix.1)[c("dose", "clinic", "prison", "shape", "scale")
                                                        c("dose", "clinic", "prison", "shape", "scale")]
                               ses=TRUE)

#point estimate
AFT.median.1
#confidence interval
c(AFT.median.1 - 1.96*AFT.median.1.se,
  AFT.median.1 + 1.96*AFT.median.1.se)


#median for person from clinic 1 on 100 mg/day with a history of incarceration
AFT.median.2 <- exp(100 * coef(weibull.adix.1)["dose"] +
                      1 * coef(weibull.adix.1)["clinic"] +
                      1 * coef(weibull.adix.1)["prison"]) *
        log(2) ^ (1/exp(coef(weibull.adix.1)["shape"])) *
        exp(coef(weibull.adix.1)["scale"])
```

```
AFT.median.2.se <- deltamethod(g = ~exp(100*x1 + 1*x2 + 1*x3)
                               * log(2) ^ (1/exp(x4)) * exp(x5),
                               mean = coef(weibull.adix.1)[c("dose", "clinic", "prison", "shape", "scal
                               cov=vcov(weibull.adix.1)[c("dose", "clinic", "prison", "shape", "scale")
                                                        c("dose", "clinic", "prison", "shape", "scale")
                               ses=TRUE)

#point estimate
AFT.median.2
#confidence interval
c(AFT.median.2 - 1.96*AFT.median.2.se,
  AFT.median.2 + 1.96*AFT.median.2.se)



#fit new weibull AFT model
weibull.adix.2 <- flexsurvreg(surv.adix ~ dose + clinic + prison + dose*prison, data=adix, dist="weibul

#weibull.adix.3 <- flexsurvreg(surv.adix ~ dose*prison + clinic, data=adix, dist="weibull")

#get point estimate and CI comparing 80 to 60 mg/day for someone with a history of incarceration and fr
AFT.point.2 <- (80-60) * coef(weibull.adix.2)["dose"] +
                  (80-60) * coef(weibull.adix.2)["dose:prison"]

AFT.se.2 <- deltamethod(g = ~(80-60) * x1 + (80-60) * x4,
                        mean = coef(weibull.adix.2)[c("dose", "clinic", "prison", "dose:prison")],
                        cov=vcov(weibull.adix.2)[c("dose", "clinic", "prison", "dose:prison"),
                                                 c("dose", "clinic", "prison", "dose:prison")],
                        ses=TRUE)

#point estimate
exp(AFT.point.2)

#confidence interval
exp(c(AFT.point.2 - 1.96*AFT.se.2,
       AFT.point.2 + 1.96*AFT.se.2))


#get point estimate and CI comparing 80 to 60 mg/day for someone without a history of incarceration and
AFT.point.3 <- (80-60) * coef(weibull.adix.2)["dose"]

AFT.se.3 <- deltamethod(g = ~(80-60) * x1,
                        mean = coef(weibull.adix.2)[c("dose", "clinic", "prison", "dose:prison")],
                        cov=vcov(weibull.adix.2)[c("dose", "clinic", "prison", "dose:prison"),
                                                 c("dose", "clinic", "prison", "dose:prison")],
                        ses=TRUE)

#point estimate
exp(AFT.point.3)

#confidence interval
exp(c(AFT.point.3 - 1.96*AFT.se.3,
       AFT.point.3 + 1.96*AFT.se.3))
```

```
#using wald test
SE.int.term <- deltamethod(g = ~exp((80-60) * x1),
                            mean=coef(weibull.adix.2)["dose:prison"],
                            cov=vcov(weibull.adix.2)["dose:prison", "dose:prison"],
                            ses=TRUE)

wald.test.stat <- (exp((80-60) * coef(weibull.adix.2)["dose:prison"])-1) / SE.int.term

wald.test.stat

#wald p-value
2 * pnorm(-abs(wald.test.stat))


library(haven)
link = "https://github.com/dmccoomes/Survival/raw/master/Homework%204/diabetesST11.dta"
diab <- read_dta(url(link))

#create study time in years variable
diab$time_year <- diab$time/12
#create age at blindness
diab$age_blind <- diab$age + diab$time_year
#subset data to include only untreated eyes
diab.untreat <- diab[which(diab$treat == 0),]



#creating survival function that includes left truncation
surv.diab.lt <- with(diab.untreat, Surv(age, age_blind, status))
survfit.diab.lt <- survfit(surv.diab.lt ~ 1, data=diab.untreat, conf.type = "log-log")

#creating survival function that does NOT include left truncation
surv.diab <- with(diab.untreat, Surv(age_blind, status))
survfit.diab <- survfit(surv.diab ~ 1, data=diab.untreat, conf.type = "log-log")

#fitting a curve that is time since diagnosis
surv.diab.3 <- with(diab.untreat, Surv(time, status))
survfit.diab.3 <- survfit(surv.diab.3 ~ 1, data=diab.untreat, conf.type = "log-log")


summary(survfit.diab.lt)$table


plot(survfit.diab.lt,
     conf.int=TRUE,
     main="Kaplan-Meier survival estimate accounting for left truncation \n(includes 95% CI)",
     ylab="Survival probability", xlab="Age",
     col=c("blue", "light blue", "light blue"),
     lty=c("solid", "dashed", "dashed"),
     lwd=c(2, 1, 1),
     caption ="caption")
```

```
summary(survfit.diab)$table


plot(survfit.diab,
     conf.int=TRUE,
     main="Kaplan-Meier survival estimate not accounting for left truncation \n(includes 95% CI)",
     ylab="Survival probability", xlab="Age",
     col=c("blue", "light blue", "light blue"),
     lty=c("solid", "dashed", "dashed"),
     lwd=c(2, 1, 1),
     caption ="caption")



plot(survfit.diab.3,
     conf.int=TRUE,
     main="Kaplan-Meier survival estimate not accounting for left truncation \n(includes 95% CI)",
     ylab="Survival probability", xlab="Time since diagnosis (months)",
     col=c("blue", "light blue", "light blue"),
     lty=c("solid", "dashed", "dashed"),
     lwd=c(2, 1, 1),
     caption ="caption")
```