Problem Set 4

Doireanna Craven

Due: April 4, 2022

Question 1

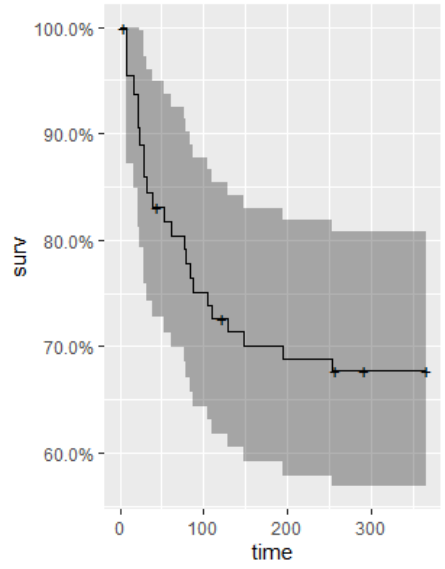
We're interested in modeling the historical causes of infant mortality. We have data from 5641 first-born in seven Swedish parishes 1820-1895. Using the "infants" dataset in the `eha` library, fit a Cox Proportional Hazard model using mother's age and infant's gender as covariates. Present and interpret the output.
R's `Surv` function returns a hazard function, which is the failure rate at time t, conditional on a person having survived to that time.

```
infants_surv <- with(infants, Surv(enter, exit, event))
```

`survfit` creates survival curves from the survival model.

```
km_infants <- survfit(infants_surv ~ 1, data = infants)
# 'survfit' creates survival curves from the survival model
summary(km_infants, times = seq(0, 365, 5))
autoplot(km_infants)
```

Let's look at the model.



Now for the Cox Proportional Hazard regression.

```
km_infants <- survfit(infants_surv ~ 1, data = infants)
# 'survfit' creates survival curves from the survival model
summary(km_infants, times = seq(0, 365, 5))
autoplot(km_infants)
```

	Dependent variable:
	infants_surv
sexboy	-0.485
	(0.442)
age	-0.040
	(0.045)
Observations	105
R^2	0.019
Max. Possible R^2	0.800
Log Likelihood	-83.626
Wald Test	2.000 (df = 2)
LR Test	1.992 (df = 2)
Score (Logrank) Test	2.034 (df = 2)
Note:	*p<0.1; **p<0.05; ***p<0.01

- The expected log of the hazard for male infants decreases by 0.49 compared to female infants, holding mother's age constant.
- The expected log of the hazard for infants decreases by 0.04 for every additional year of mother's age, holding infant's sex constant.

Exponentiate parameter estimates to obtain hazard ratios:

- The hazard for male infants is 0.62 times the hazard for female babies, i.e., female infant deaths are 38% lower than male infant deaths, holding mother's age constant at its mean.
- The hazard for infants is 4% lower for every extra year of mother's age, holding sex constant at its mean.