

# Problem Set 4

## Applied Stats II

Due: April 4, 2022

### Instructions

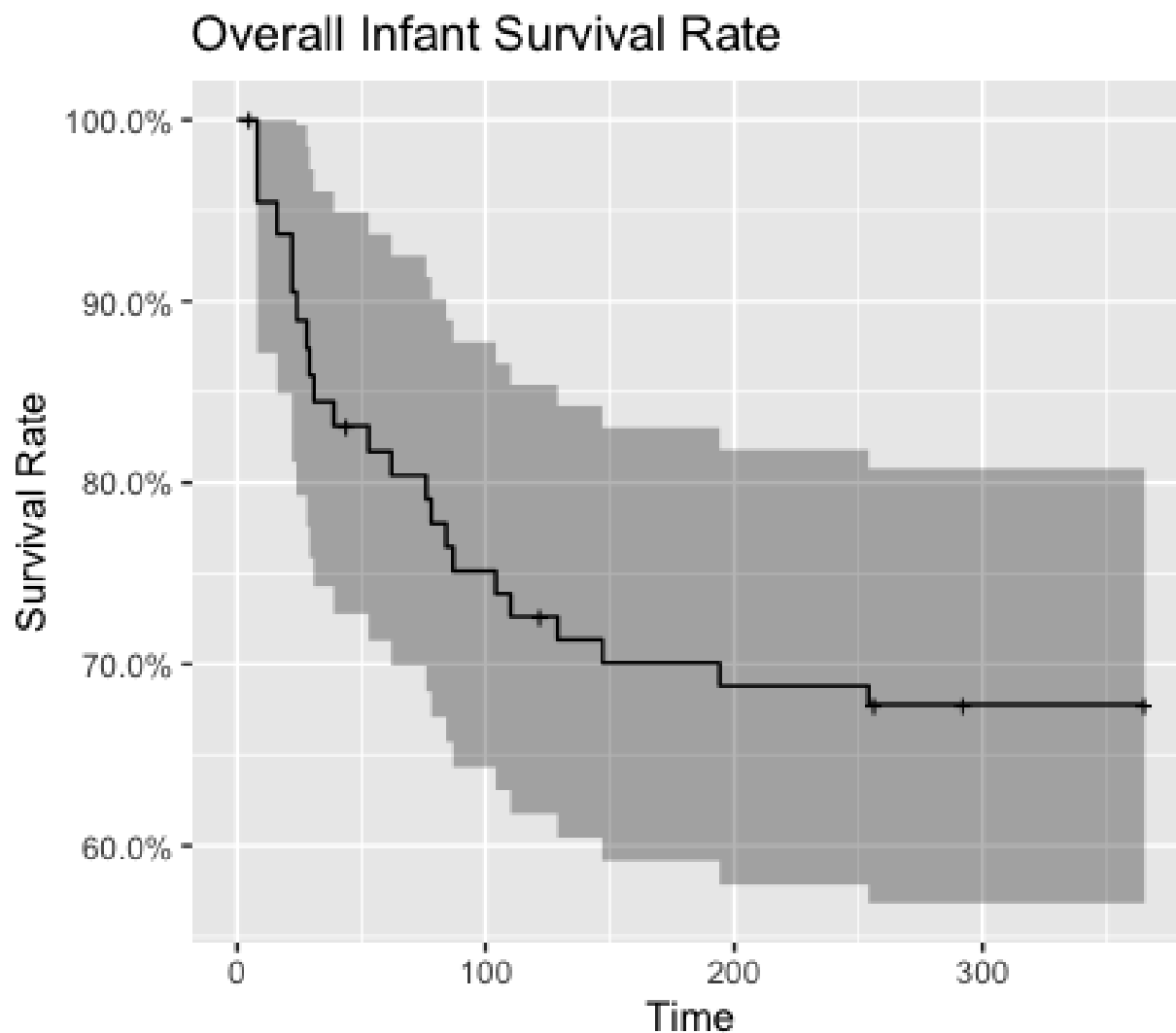
- Please show your work! You may lose points by simply writing in the answer. If the problem requires you to execute commands in R, please include the code you used to get your answers. Please also include the .R file that contains your code. If you are not sure if work needs to be shown for a particular problem, please ask.
- Your homework should be submitted electronically on GitHub in .pdf form.
- This problem set is due before class on Monday April 4, 2022. No late assignments will be accepted.
- Total available points for this homework is 80.

### 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.

### Question 1 - My Answer

```
1
2 # Load infants dataset
3 data(infants)
4
5 # using Tutorial 10 as reference, where 'child' was used
6 infant_survivor <- with(infants, Surv(enter, exit, event))
7
8 # using survit to create the survival model
9
10 ## and need to run a Cox Proportional Hazard regression on the data
11
```



```

12 km <- survfit(infant_survivor ~ 1, data = infants)
13
14
15 # summary(km, times = seq(0, 15, 1))
16 # event = 0, exit = 365, exit = 5
17
18 summary(km, times = seq(0, 365, 5))
19 autoplot(km, main="Overall Infant Survival Rate",
20           xlab = "Time",
21           ylab = "Survival Rate")

```

```

1
2
3 # Run a Cox Proportional Hazard Regression
4 # referencing tutorial 10
5 cox <- coxph(Surv(enter, exit, event) ~ sex + age, data = infants)

```

```

6 summary(cox)
7 drop1(cox, test = "Chisq")
8 # stargazer(cox, type = "text")
9 stargazer(cox)
10
11 ## For the Survival Rate in the first 100 days there is an increase, which
    then decrease after
12 ## 100 days but the overall infant survival rate is shown to be quite low.

```

Table 1: Cox Proportional Hazard Regression

<i>Dependent variable:</i>	
	enter
sexboy	−0.485 (0.442)
age	−0.040 (0.045)
Observations	105
R <sup>2</sup>	0.019
Max. Possible R <sup>2</sup>	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)
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	