STAT3032 SURVIVAL MODELS

TUTORIAL WEEK SEVEN

Question One

Use the first two terms in a Taylor series expansion of $\exp(-x)$ to demonstrate that the Nelson-Aalen estimator and the Kaplan-Meier estimator are roughly equal.

Question Two

The table below contains information on the survival times after a particular operation. There is also information on which hospital the operation was performed at.

Observation Time	Hospital
5	A
9	A
4	B
8	B
6	B
15*	B
13	B
1*	A
2	A

(a) In R create three vectors containing the above information (time, censor/death indicator and hospital indicator). The hospital indicator will take the value 1 for one of the hospitals and 0 for the other.

(b) Use the coxph command in R to fit a Cox regression model to the above data. Comment on the significance of hospital in determining survival probabilities for the above data.

Question Three

The effect of 3 risk factors (rf) on a disease was of interest. The variable z_i represents the presence or absence of factor i, so

$$z_i = \begin{cases} 1 \text{ if individual has rf } i \\ 0 \text{ if individual doesn't have rf } i \end{cases}$$

The data which was observed is

Survival	z_1	z_2	z_3
5	0	1	0
6	0	1	1
4*	1	1	0

1	1	0	1
2	1	0	0
8*	0	0	1
9	0	0	1
5	1	0	1

(a) Write out the partial likelihood for this data.

(b) Use R to fit and test the model in (a). Are any of the risk factors statistically significant in their impact on survival?

(c) Describe the estimated impact of risk factor 1 on both the hazard functions and survival probabilities.

Question Three

The output from a Cox proportional-hazards regression analysis of a Recidivism dataset (Rossi, Berk and Lenihan 1980) is provided below. The purpose of the analysis was to investigate whether certain covariates were related to survival time (in this context survival time being the time until first-arrest upon release from prison). The covariates included in the fitted model are:

- *fin*: a categorical variable taking the value 1 if financial aid was received and 0 otherwise.
- *age*: age in years.
- *race*: a categorical variable taking the value 1 for blacks and 0 otherwise.
- *mar*: a categorical variable taking the value 1 if married and 0 otherwise.
- *prio*: the number of prior convictions

```
> summary(fit)
Call:
coxph(formula = Surv(week, arrest) ~ fin + age +
I(age^2) + race +
    mar + prio, data = Rossi)

  n= 432, number of events= 114

            coef    exp(coef)  se(coef)  z      Pr(>|z|)
fin      -0.373246  0.688496  0.191009 -1.954  0.05069
age      -0.276401  0.758509  0.136125 -2.031  0.04231
age^2     0.003907  1.003915  0.002409  1.622  0.10485
race      0.344845  1.411770  0.308424  1.118  0.26353
mar      -0.417402  0.658756  0.378661 -1.102  0.27033
prio      0.099941  1.105106  0.027367  3.652  0.00026
---

Concordance= 0.642  (se = 0.027 )
Rsquare= 0.078   (max possible= 0.956 )
Likelihood ratio test= 34.99  on 6 df,   p=4.336e-06
Wald test            = 35.54  on 6 df,   p=3.389e-06
```

```
Score (logrank) test = 37.09  on 6 df,   p=1.693e-06
```

Note in the questions that follow, β_1 refers to the parameter corresponding to the variable *fin*, β_2 to the parameter corresponding to the variable *age*, and so on.

Using the R-output above answer the following questions:

a) Estimate the ratio of the hazards of time until first-arrest for individual A and individual B. *Individual A*: receives financial assistance, aged 23, is black, has never been married and has 2 prior convictions. *Individual B*: receives no financial assistance, aged 26, is black, married and has 0 prior convictions.

b) Provide a 95% confidence interval for the multiplicative increase in the hazard ratio for an increase in the number of prior convictions of two, everything else held constant.

c) Provide a standard error for the estimate of $\exp(\beta_1)$ obtained from the fitted model.

d) Is marital status related to time until first-arrest? You must provide statistically sound reasons for your answer.

e) Does the positive coefficient estimate obtained for the covariate representing the number of prior convictions seem reasonable? You must provide justification for your answer.

$Var(\beta)$

$Var(e^{\beta}) = var(\beta)\cdot \cancel{e^{2\beta}} (e^{\beta})^2$

$\cancel{Var(e^{\beta})} = Var(\beta)\cdot e^{2\beta}$

$= \cancel{Var(\beta)}$