MATH3085/6143 Survival Models – Worksheet 3

1. In a Cox regression model with smoking and age as covariates, the following estimates of the coefficients associated with smoking and age were obtained:

Covariates	Coefficient estimates
Age (years)	0.005
Cigarette Smoker	0.05

Show that according to the fitted model, a smoker aged 55 years has the same risk of death as a non-smoker aged 65 years.

2. A study was undertaken into the length of spells of unemployment among young people in a certain city. A sample of young people was monitored from the time they started to claim unemployment benefit until either they resumed work, or they moved away from the city. None of the members of the sample died during the study. The study investigated the impact of age, sex and educational qualifications on the hazard of returning to work using the following covariates:

   A a young person's age when he/she started claiming benefit (measured in exact years since his/her 16^{th} birthday).
   F a dummy variable taking value 1 if the person was male and 0 if the person was female.
   E a dummy variable taking the value 1 if the person had passed a school leaving examination in mathematics and 0 otherwise.

   with the associated parameters β_A, β_F and β_E. The investigators decided to use a Cox proportional hazards regression model for the study.

   i) Write down the equation of the model that was estimated, defining the terms you use (other than those defined above).

   ii) List the characteristics of the young person to whom the baseline hazard applies.

   iii) The results showed:

   - The hazard of resuming work for males who started claiming benefit aged 17 years exact and who had passed the mathematics examination was 1.5 times the hazard for males who started claiming benefit aged 16 years exact but who had not passed the mathematics examination.
   - Females who had passed the mathematics examination were twice as likely to take up a new job as were males of the same age who had failed the mathematics examination.
   - Females who started claiming benefit aged 20 years exact and who had passed the mathematics examination were twice as likely to resume work as were males who started claiming benefit aged 16 years exact and who had also passed the mathematics examination.

   Calculate the estimated values of the parameters β_A, β_F and β_E.

   [Based on Question 11 in the Institute and Faculty of Actuaries CT4 examination in September 2009.]

3. A farmer is concerned that he is losing a lot of his birds to a predator, so he decides to build a new enclosure using taller fencing. This fencing is expensive and he cannot afford to build a large enough area for all his birds. He therefore decides to put half his birds in the new enclosure and leave the others in the existing enclosure. He is convinced that the new enclosure is an improvement, but has asked an actuarial student to determine whether the new enclosure will result in an increase in the life expectancy of his birds. The student has fitted a Cox proportional hazards model to data on the duration until a bird is killed by a predator and calculated the following figures relating to the regression parameters:

	Bird			Enclosure		Sex	
	Chicken	Duck	Goose	New	Old	Male	Female
Parameter estimate	0	−0.210	0.075	0.125	0	0.2	0
Variance	0	0.002	0.004	0.0015	0	0.0026	0

i) State the features of the bird to which the baseline hazard applies.

ii) For each regression parameter, define the associated covariate and calculate the 95% confidence interval based on the standard error.

iii) Comment on the farmer's belief that the new enclosure will result in an increase in his birds' life expectancy.

iv) Calculate, using this model, the probability that a female duck in the new enclosure has been killed by a predator at the end of six months, given that the probability that a male goose in the old enclosure has been killed at the end of the same period is 0.1 (all other possible causes of death can be ignored).

[Based on Question 9 in the Institute and Faculty of Actuaries CT4 examination in April 2010.]

4. Data are available from a small portfolio of employment protection policies, and show, for two classes of policyholders, the time in months until a claim is made; the + indicates that the observation was censored.

Group A	2	5+	6	7+	7+	7+	8	10	14
Group B	3+	4	7+	7+	7+	7+	9	11+	15+

Suppose that the Cox proportional hazards model holds for these data and that Group A is taken as the reference category, i.e., in the standard notation, $h_B(t) = h_A(t)\exp(\beta)$.

i) Show that the partial log-likelihood is

$$\ell(\beta) = 2\beta - 6\log(1 + e^{\beta}) - \log(2 + 3e^{\beta}) - \log(3024)$$

ii) Deduce the value of $\hat{\beta}$, the maximum partial likelihood estimate of β, together with its standard error.

iii) Test the hypothesis $H_0 : \beta = 0$ against the alternative $H_1 : \beta \neq 0$ and report your conclusions.

5. In a preliminary study of a new medical treatment, 3 patients were given the new treatment and 3 an existing treatment. For all patients the time (in days) to recurrence of symptoms is recorded (with a `+` indicating a right-censored observation) below. After 10 weeks, the trial was terminated for analysis.

```
New treatment:  53  70+  70+      Existing treatment:  18  67  70+
```

(a) A Cox proportional hazards model is proposed, where the hazard functions, h_N in the new treatment group and h_E in the existing treatment group, are related by

$$h_N(t) = \exp(\beta)\, h_E(t).$$

Write down an expression for the partial likelihood $L(\beta)$ for these data.

(b) Is the censoring at 70 days informative or non-informative? Explain your reasoning.

6. An engineering company uses a Weibull accelerated failure model to investigate whether a new type of sealant has the potential to reduce failure rates of the components they manufacture. Failure times $t_1, \ldots, t_n$ are recorded along with the single explanatory variable x_i $(i = 1, \ldots, n)$ which was recorded as 1 if the new sealant was used and 0 for the old sealant.

(a) Write down carefully the model which has been fitted here.

(b) Describe one way in which this model differs from a Cox proportional hazards model?

(c) The regression parameter β for this explanatory variable was estimated to be 0.14 with standard error 0.09. Calculate a 95% confidence interval for β. What conclusion can be reached about the effectiveness of the new sealant? Note that $z_{0.975} = 1.96$.