A.4 Markov models; Model testing; Proportional-hazards regression

1. Public health officials often compare the effects of different changes in population health by considering the resulting change in life expectancy. But what about the personal and societal costs of illness and disability? This has led to the notion of "Disability Adjusted Life Years" (DALYs) or "Quality Adjusted Life Years" (QALYs), in which years are weighted by some measure of the "quality of life". Suppose we have a Markov representation of lifetimes, in which there are m non-absorbing "alive" states, and 1 absorbing "dead" state. The "quality" of a year of life in state i is w_i. Individuals begin in state i with probability p_i.

   (a) Show that the expected total value of a life is given by $p^T(-Q_*^{-1})w$, where T means "transpose", and Q_* is the submatrix of Q corresponding to the non-absorbing states.

   (b) Suppose the model is the simple Healthy-Sick-Dead model described in section 9.5.2, with $\sigma = 0.1$, $\rho = 0$, $\delta = 0.01$, and $\gamma = 0.2$. What is the life expectancy of someone initially healthy?

   (c) Suppose we judge a year of being sick to be worth half of a year of health. What is the expected number of quality-adjusted years in the lifetime of someone initially healthy? What about the number remaining to someone who has just gotten sick?

   (d) How many QALYs would be saved by a cure that now raises ρ to 0.2? Suppose the alternative were a treatment that would lower γ by some amount. Could you achieve the same QALY savings by such a treatment? How low would γ have to go?

   (e) What fraction of individuals are sick right before they die?

   (f) Suppose after 30 years the survivors move into "old age", in which the rate of becoming sick σ rises to 0.5. What is the effect on QALYs?

2. (a) Supppose that we have a random sample which includes right-censored data (censoring assumed non-informative). We wish to decide whether or not a Weibull distribution is appropriate. Using an estimator of the survival function how might we graphically investigate the appropriateness of the model? Given that the model appears to be appropriate how would you test whether or not the special case of an exponential model is valid? Suppose that the Weibull model does not appear to be appropriate what graph would you use to consider a log-logistic model?

   (b) Now suppose that there are two groups to be considered (eg smokers v. non-smokers). What graphs would be appropriate for consideration of a proportional hazards model, accelerated life model respectively?

   (c) Gehan (1965) studied 42 leukaemia patients. Some were treated with the drug *6-mercaptopurine* and the rest are controls. The trial was designed as matched pairs, but both members of a pair observed until both came out of remission or the study ended. (The data are included under the name `gehan` in the `R` package `MASS`. The description attached to these data there says that in each pair both were withdrawn from the trial when either came out of remission. If you have a look at the data, you can see that this is clearly not true.) The observed times to recurrence (in months) were:

$$\begin{aligned}\text{Controls: } & 1, 1, 2, 2, 3, 4, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23\\ \text{Treatment: } & 6+, 6, 6, 6, 7, 9+, 10+, 10, 11+, 13, 16, 17+, 19+, 20+, 22, 23, 25+,\\ & 32+, 32+, 34+, 35+\end{aligned}$$

   Here + indicates censored times. Investigate these data in respect of both a) and b).

3. (a) Describe the proportional hazards model, explaining what is meant by the partial likelihood and how this can be used to estimate resgression coefficients. How might standard errors be generated?

(b) Drug addicts are treated at two clinics (clinic 0 and clinic 1) on a drug replacement therapy. The response variables are the time to relapse (to re-taking drugs) and the status relapse =1 and censored =0. There are three explanatory variables, clinic (0 or 1), previous stay in prison (no=0, yes=1) and the prescribed amount of the replacement dose. The following results are obtained using a proportional hazards model, $h(t,x) = e^{\beta x} h_0(t)$.

Variable	Coeff	St Err	p-value
clinic	-1.009	0.215	0.000
prison	0.327	0.167	0.051
dose	-0.035	0.006	0.000

What is the estimated hazard ratio for a subject from clinic 1 who has not been in prison as compared to a subject from clinic 0 who has been in prison, given that they are each assigned the same dose?

(c) Find a 95% confidence interval for the hazard ratio comparing those who have been in prison to those who have not, given that clinic and dose are the same.

4. (a) Sketch the shape of the hazard function in the following cases, paying attention to any changes of shape due to changes in value of κ where appropriate.

i. Weibull: $S(t) = e^{-\rho t^\kappa}$.

ii. Log-logistic: $S(t) = \frac{1}{1+(\rho t)^\kappa}$.

(b) Suppose that it is thought that an accelerated life model is valid and that the hazard function has a maximum at a non-zero time point. Which parametric models might be appropriate?

(c) Suppose that $t_1, \ldots, t_n$ are observations from a lifetime distribution with respective vectors of covariates $x_1, \ldots, x_n$. It is thought that an appropriate distribution for lifetime y is Weibull with parameters ρ, κ, where the link is $\log \rho = \beta' x$. In the case that there is no censoring write down the likelihood and, using maximum likelihood, give equations from which the vector of estimated regression coefficients β (and also the estimate for κ) could be found. What would be the asymptotic distribution of the vector of estimators? How would the likelihood differ if some of the observations t_i were right censored (assuming independent censoring)?

5. Coronary Heart Disease (CHD) remains the leading cause of death in many countries. The evidence is substantial that males are at higher risk than females, but the role of genetic factors versus the gender factor is still under investigation. A study was performed to assess the gender risk of death from CHD, controlling for genetic factors. A dataset consisting of non-identical twins was assembled. The age at which each person died of CHD was recorded. Individuals who either had not died or had died from other causes had censored survival times (age). A randomly selected subsample from the data is as follows. (* indicates a censored observation.)

Age male twin	Age female twin
50	63*
49*	52
56*	70*
68	75
74*	72
69*	69*
70*	70*
67	70
74*	74*
81*	81*
61	58
75*	73*

(a) Write down the times of events and list the associated risk sets.

(b) Suppose the censoring mechanism is independent of death times due to CHD, and that the mortality rates for male and female twins satisfy the PH assumption, and let β be the regression coefficient for the binary covariate that codes gender as 0 or 1 for male or female respectively. Write down the partial-likelihood function. Using a computer or programmable calculator, compute and plot the partial-likelihood for a range of values of β. What is the Cox-regression estimate for β? What does this mean?

(c) Estimate the survival function for male twins.

(d) Suppose now only that the censoring mechanism is independent of death times due to CHD, perform the log-rank test for equivalence of hazard amongst these two groups. Contrast the test statistic and associated p-value with the results from the Fleming–Harrington test using a weight $W(t_i) = \hat{S}(t_{i-1})$.

(e) Do you think the assumption of a non-informative censoring mechanism is appropriate? Give reasons.

6. *(Based on Exercise 11.1 of* **[KM03]***.)* The dataset `larynx` in the package `KMsurv` includes times of death (or censoring by the end of the study) of 90 males diagnosed with cancer of the larynx between 1970 and 1978 at a single hospital. One important covariate is the stage of the cancer, coded as 1,2,3,4.

(a) Why would it probably not be a good idea to fit the Cox model with relative risk $e^{\beta \cdot \text{stage}}$? What should be done instead?

(b) Explain how you would use a martingale residual plot to show that `stage` does not enter as a linear covariate.

(c) Which residual plot would you use to test whether the proportional-hazards assumption holds for `age` or `stage`, or whether the proportional effect of one of these covariates changes over time.

(d) Explain how you would use a Cox–Snell residual plot to test whether the Cox model is appropriate to these data. Describe the calculations you would perform, the plot that you would create, and describe the visual features you would be looking for to evaluate the goodness of fit.

(e) **[optional]** Carry out these computations in `R`.

i. One way of making `R` treat the `stage` variable appropriately is to replace it in the model definition by `factor(stage)`. Show that this produces the same result as defining

ii. Try adding year of diagnosis or age at diagnosis as a linear covariate (in the exponent of the relative risk). Is either statistically significant?

iii. Use a martingale residual plot to show that `stage` does not enter as a linear covariate.

iv. Use a residual plot to test whether one or the other of these covariates might more appropriately enter the model in a different functional form — for example, as a step function.

v. Use a Cox–Snell residual plot to test whether the Cox model is appropriate to these data.