Tutorial 2

- Solutions (handout)
- Extra comments in www.richardsconsulting.co.uk/lectures/Tutorial2.pdf

Common mistakes and questions

- Is ${}_tp_x^{gh}$ in the multi-state Markov model the same thing as ${}_tp_x$?

⇒ No.
In fact, it is precisely the opposite. ${}_tp_x^{gh}$ is the probability of jumping from state g to state h. ${}_tp_x$ is the probability of surviving time t. If you think of the basic lifetime model as two states, Alive (a) and Dead (d), then ${}_tp_x$ is the probability that the life x starts alive and is still alive at time t. Thus, ${}_tp_x$ in Lecture 1 notation is actually ${}_tp_x^{\overline{aa}}$ in Markov notation. Similarly, ${}_tp_x^{\overline{ad}}$ in Markov notation is actually ${}_tq_x$ in actuarial and Lecture 1 notation.

- How do I choose the baseline for a Cox model? Does it make a difference?

⇒ You can choose whatever you like as a baseline. . .and it makes no difference.
Note that some questions will steer you towards a choice, e.g. if one group is labelled 'Control' and the other group is the drug being applied, you ought to choose the Control as the baseline.

- Why the e^β in the Cox partial likelihood? Why not just β?

⇒ Convenience.
We use e^β because we invariably end up working with the log-likelihood. By using e^β, we end up working with β in the log-likelihood.
⇒ R (and other software) uses e^β for the same reason.

- How do I interpret z in the z-test?

⇒ $z \approx N(0, 1)$.
If $z \in [-1.96, 1.96]$, then there is insufficient information to reject the null hypothesis that $\beta_0 = 0$ at the 5% level. If z lies outwith this range, there is only a 5% probability that this happened by chance.

- Where does the 1.96 come from?

⇒ 1.96 is the upper 2.5% point of the $N(0, 1)$ distribution.
-1.96 is the lower 2.5% point.

- What is the meaning behind z^2, S and $-2 \log \Lambda$ all having similar values?

⇒ No particular meaning.
They are all testing the same thing, and all test statistics are approximately distributed as per χ^2_1. We expect the test statistics to have similar values, otherwise they will contradict each other.

Data and estimation in two-state Markov model

- We saw three different methods to obtain $\hat{\mu}$ in the two-state model
- Consider the likelihood equivalent of the confidence interval, the *support interval*

First, expand $\ell(\mu)$ as a Taylor series around $\hat{\mu}$:

$$\begin{aligned}\ell(\mu) &= \ell(\hat{\mu}) + \ell'(\hat{\mu})(\mu - \hat{\mu}) + \ell''(\hat{\mu})\frac{(\mu - \hat{\mu})^2}{2!} + \dots \\ &= \ell(\hat{\mu}) + \ell''(\hat{\mu})\frac{(\mu - \hat{\mu})^2}{2!} + \dots\end{aligned}$$

as $\ell'(\hat{\mu}) = 0$ by definition of $\hat{\mu}$. Re-arranging gives:

$$-2\left(\ell(\mu) - \ell(\hat{\mu})\right) = \ell''(\hat{\mu})(\mu - \hat{\mu})^2 + \dots$$

where we recognise the left-hand side as the test statistic for the liklihood ratio test. We want the values of μ such that $\ell(\mu) = \ell(\hat{\mu}) - 2$ for the 2-support interval, i.e.

$$\begin{aligned}-2\left(\ell(\hat{\mu}) - 2 - \ell(\hat{\mu})\right) &\approx \ell''(\hat{\mu})\left(\mu^2 - 2\mu\hat{\mu} + \hat{\mu}^2\right) \\ \Rightarrow \frac{4}{\ell''(\hat{\mu})} &\approx \mu^2 - 2\hat{\mu}\mu + \hat{\mu}^2 \\ \Rightarrow 0 &\approx \mu^2 - 2\hat{\mu}\mu + \hat{\mu}^2 - \frac{4}{\ell''(\hat{\mu})}\end{aligned}$$

which is a quadratic equation in μ with roots:

$$\frac{2\hat{\mu} \pm \sqrt{4\hat{\mu}^2 - 4\left(\hat{\mu}^2 - \frac{4}{\ell''(\hat{\mu})}\right)}}{2}$$

which simplifies to:

$$\hat{\mu} \pm 2\sqrt{\frac{1}{\ell''(\hat{\mu})}}$$

which we recognise as the standard form for a confidence interval. The term inside the $\sqrt{\ }$ is the variance, so $\mathrm{Var}(\hat{\mu}) \approx \ell''(\hat{\mu})^{-1}$

$$\ell''(\mu) \equiv \frac{\partial^2 \ell}{\partial \mu^2} \equiv -I(\mu)$$

where $I(\mu)$ is called *Fisher's information function*, so $\mathrm{Var}(\hat{\mu}) \approx I(\hat{\mu})^{-1}$

Example 5.1 on page 46 of Red Book

- $\hat{\mu} = 1.0909$
- 2-support interval is (0.26, 2.88)
- $\dfrac{\partial^2 \ell}{\partial \mu^2} = -\dfrac{\sum d_i}{\mu^2}$

  $\Rightarrow I(\mu) = \dfrac{\sum d_i}{\mu^2}$

  $\Rightarrow \mathrm{Var}(\hat{\mu}) \approx \dfrac{\hat{\mu}^2}{\sum d_i} = \dfrac{\left(\frac{\sum d_i}{\sum v_i}\right)^2}{\sum d_i} = \dfrac{\sum d_i}{\left(\sum v_i\right)^2}$

  $\Rightarrow$ Approximate 95% confidence interval is $1.0909 \pm 2\dfrac{\sqrt(3)}{2.75}$, i.e. (-0.17, 2.35)
- Interval from R output (exponential fit) is (0.34, 3.46)
- These intervals are rather different, but only because the sample size is so small.