M3S3 - STATISTICAL THEORY II
ASSESSED COURSEWORK 2

Official Deadline: Friday 24th March

Please hand in to Room 523
(or via the Central Office if necessary)

1. Suppose that $X_1, \ldots, X_n$ are an i.i.d. sample from a Weibull distribution with parameters $\boldsymbol{\theta} = (\alpha, \lambda)^\mathsf{T}$, with pdf

$$f_{X|\boldsymbol{\theta}}(x|\boldsymbol{\theta}) = \frac{\alpha}{\lambda^\alpha} x^{\alpha-1} \exp\left\{-\left(\frac{x}{\lambda}\right)^\alpha\right\} \qquad x > 0.$$

Find the form of the profile likelihood for α.

[5 MARKS]

2. In a comparison of the rates of deaths on two roads (denoted A and B), a commonly used model assumes that deaths occur according to a Poisson process. That is, the number of deaths in one year on road A is a random variable $X_A \sim Poisson(\lambda_A)$, with a similar random variable, $X_B \sim Poisson(\lambda_B)$ for road B. It can be assumed that numbers of deaths on the two roads are independent random variables, and the Poisson Process assumptions imply independence across successive years.

The following count data were recorded for the numbers of deaths per year on the two roads:

Year	1	2	3	4	5	6
Road A	3	0	2	1	5	1
Road B	1	2	0	1		

(i) Find the profile likelihood for parameter of interest $\theta = \lambda_A/\lambda_B$.

[3 MARKS]

(ii) By focussing on θ or otherwise, assess the evidence in these data that the rates of deaths in the two roads are statistically significantly different. Comment on any asymptotic approximations you make. Statistical Tables can be found at the course website

`stats.ma.ic.ac.uk/~das01/M3S3/`

[4 MARKS]

3. In the statistical analysis of 2×2 tables, a model that assumes independent Binomial sampling, conditional on row totals n_1 and n_2, within the two rows of the table is typically used. Under this model, the entries in the first column, X_1 and X_2 are distributed as follows

$$X_1|n_1, \pi_1 \sim Binomial(n_1, \pi_1) \qquad X_2|n_2, \pi_2 \sim Binomial(n_2, \pi_2)$$

where $n_1, n_2 > 0$ and $0 < \pi_1, \pi_2 < 1$.

(i) Find the ML estimator, $\widehat{\theta}_n$, for the *odds-ratio*

$$\theta = \frac{\pi_1(1-\pi_2)}{\pi_2(1-\pi_1)}$$

[4 MARKS]

(ii) Find an asymptotic approximation to the distribution of

$$\log \widehat{\theta}_n$$

where $n = n_1 + n_2$.

Hint: use the Delta Method.

[4 MARKS]