Data provided: Background information; table of χ^2 critical values

The University Of Sheffield.

MAS371

SCHOOL OF MATHEMATICS AND STATISTICS **Spring Semester 2019–2020**

Applied Probability **1 hour**

This is an open book exam.
*Answer **all** questions.*
The submission deadline is 10 am (BST), twenty-four hours after it is released. Late submission will not be considered without extenuating circumstances. It is expected that you will be able to complete this exam in approximately one hour and it is recommended that you submit the work within four hours. You will not be penalised for taking longer, however.
Unless it is explicitly stated otherwise, it is intended that calculations are performed by hand (possibly with the aid of a calculator). To gain full marks, you will need to show your working. You will not get full marks if you simply write down output from a computer package.
By uploading your solutions you declare that your submission consists entirely of your own work, that any use of sources or tools other than material provided for this module is cited and acknowledged and that no unfair means have been used.

There are 30 marks available on the paper.

1 The results from a computer simulation of coin tossing are modelled as a discrete time Markov chain with state space $\{H, T\}$ representing heads and tails respectively. The transition matrix is assumed to be

$$\begin{pmatrix} \theta_1 & 1-\theta_1 \\ \theta_2 & 1-\theta_2 \end{pmatrix}.$$

Here θ_1 and θ_2 are unknown values in $(0, 1)$ and represent the probability of a head when the previous result was a head or tail respectively. A sequence of 1000 observations from the simulation gives the following data on transitions.

Numbers of transitions	Head at time $t+1$	Tail at time $t+1$
Head at time t	278	262
Tail at time t	261	198

(a) Give the likelihood of θ_1 and θ_2 given this information and assuming the starting state is known. ***(3 marks)***

(b) Find the maximum likelihood estimates of θ_1 and θ_2 given this information. ***(4 marks)***

(c) Test the null hypothesis that $\theta_1 = \theta_2 = 1/2$. ***(6 marks)***

2 A continuous time Markov chain on state space $\{1, 2, 3\}$ has generator matrix

$$G = \begin{pmatrix} -4 & 3 & 1 \\ 3 & -6 & 3 \\ 1 & 2 & -3 \end{pmatrix}.$$

(a) Find the unique stationary distribution of the chain. ***(4 marks)***

(b) The jump chain is a discrete time Markov chain which gives the sequence of states visited by the continuous time chain. Give the transition matrix of the jump chain. ***(4 marks)***

3 A simple disease model considers two individuals A and B, each of whom may be either infected or uninfected, and has three parameters ϵ, λ and μ, all of which are positive real numbers. During a short time interval of length h, the following things may happen.

- A currently infected individual becomes uninfected with probability $\mu h + o(h)$.
- A currently uninfected individual becomes infected with probability $\epsilon h + o(h)$ if the other individual is uninfected, and with probability $(\epsilon+\lambda)h+o(h)$ if the other individual is infected.
- Multiple changes of infection status happen with probability $o(h)$.

Consider this as a continuous time Markov chain with four states, labelled as follows:

Label	Description
State 1	Both A and B are uninfected
State 2	A is infected, B is uninfected
State 3	A is uninfected, B is infected
State 4	Both A and B are infected

Give the 4×4 generator matrix of the chain. ***(9 marks)***

End of Question Paper

Background material for MAS371 exam

For the purposes of the MAS371 exam, you may assume these results in the somewhat simplified form they are given in this document.

Asymptotic normality of maximum likelihood estimators

Given a vector of unknown parameters $\hat{\boldsymbol{\theta}}$, the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ has a distribution which is asymptotically (in the large sample limit) Normal with mean vector $\boldsymbol{\theta}_0$ and covariance matrix approximately given by $J(\hat{\boldsymbol{\theta}})^{-1}$. Here $\boldsymbol{\theta}_0$ is the true value and $J(\boldsymbol{\theta})$ is the observed information matrix.

The r, s entry of $J(\boldsymbol{\theta})$ is given by

$$J(\boldsymbol{\theta})_{rs} = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\theta_r \partial\theta_s},$$

where $\ell(\boldsymbol{\theta})$ is the log likelihood.

Wilks' Theorem

Let $\boldsymbol{\theta}$ be a vector of unknown parameters with $\boldsymbol{\theta} \in \Theta$, where Θ is a p-dimensional set. Let $\boldsymbol{\theta}_0$ be the true value.

Simple hypothesis

Consider the null hypothesis $H_0 : \boldsymbol{\theta}_0 = \boldsymbol{\theta}^*$, where $\boldsymbol{\theta}^*$ is a specified value. Then, under H_0,

$$W = -2(\ell(\boldsymbol{\theta}^*) - \ell(\hat{\boldsymbol{\theta}})),$$

where $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimator and ℓ is the log likelihood, has an asymptotically (in the large sample limit) χ^2 distribution with p degrees of freedom.

Composite hypothesis

Let Θ_0 be a q-dimensional subset of Θ, and consider the null hypothesis $H_0 : \boldsymbol{\theta}_0 \in \Theta_0$. Then, under H_0,

$$W = -2(\ell(\boldsymbol{\theta}^*) - \ell(\hat{\boldsymbol{\theta}})),$$

where $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimator and ℓ is the log likelihood, has an asymptotically (in the large sample limit) χ^2 distribution with $p - q$ degrees of freedom.

Table of the q-quantile of the χ^2 distribution with ν degrees of freedom

		ν								
		1	2	3	4	5	6	7	8	9
	0.10	0.02	0.21	0.58	1.06	1.61	2.2	2.83	3.49	4.17
	0.50	0.45	1.39	2.37	3.36	4.35	5.35	6.35	7.34	8.34
q	0.90	2.71	4.61	6.25	7.78	9.24	10.64	12.02	13.36	14.68
	0.95	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92
	0.99	6.63	9.21	11.34	13.28	15.09	16.81	18.48	20.09	21.67
		ν								
		10	20	30	40	50	60	70	80	90
	0.10	4.87	12.44	20.6	29.05	37.69	46.46	55.33	64.28	73.29
	0.50	9.34	19.34	29.34	39.34	49.33	59.33	69.33	79.33	89.33
q	0.90	15.99	28.41	40.26	51.81	63.17	74.4	85.53	96.58	107.57
	0.95	18.31	31.41	43.77	55.76	67.5	79.08	90.53	101.88	113.15
	0.99	23.21	37.57	50.89	63.69	76.15	88.38	100.43	112.33	124.12