The University Of Sheffield.

SCHOOL OF MATHEMATICS AND STATISTICS **Spring Semester 2018–2019**

Epidemiology and Time Series **2 hours**

Attempt all the questions. The allocation of marks is shown in brackets.

RESTRICTED OPEN BOOK EXAMINATION

Candidates may bring to the examination lecture notes and associated lecture material (but no textbooks) plus a calculator that conforms to University regulations.

There are 60 marks available on the paper.

Please leave this exam paper on your desk
Do not remove it from the hall

Registration number from U-Card (9 digits)
to be completed by student

Blank

1 (i) Consider that 200 observations of a time series $\{y_t\}$ gave values of the sample partial autocorrelation function (PACF) and sample autocorrelation function (ACF) tabulated below:

Lag h	1	2	3	4
ACF (r_h)	0.4	0.3	0.1	0.1
PACF ($a_h^{(h)}$)	$\star$	$\star\star$	0.01	0.04

(a) Find the values of $\star$ and $\star\star$. *(3 marks)*

(b) Test whether $\{y_t\}$ is consistent with moving average models: MA(1) and MA(2). *(4 marks)*

(c) Test whether $\{y_t\}$ is consistent with autoregressive models: AR(1) and AR(2). *(3 marks)*

(ii) Consider the time series $\{y_t\}$, defined as

$$y_t = (-1)^t x_t,$$

where x_t is a time series generated by the autoregressive model:

$$x_t = \frac{1}{2}x_{t-1} - \frac{1}{3}x_{t-2} + \epsilon_t, \tag{1}$$

with ϵ_t being a white noise process with variance 1.

(a) Show that the time series $\{x_t\}$ is causal. *(4 marks)*

(b) Show that the time series $\{y_t\}$ is stationary. *(6 marks)*

2 (i) Consider the non-stationary time series $\{y_t\}$ so that the time series

$$x_t = (1 - B^4)y_t$$

is stationary.

Suppose that x_t is modelled with the autoregressive model

$$x_t = 0.7x_{t-1} + \epsilon_t,$$

where ϵ_t is white noise with variance 2.

8 observations of y_t are recorded in the table below

t	1	2	3	4	5	6	7	8
y_t	12	8	14	18	11	9	16	20

(a) Based on the data above, calculate the 1-step and 5-step ahead forecasts of the observations y_9 and y_{13}, respectively. ***(4 marks)***

(b) Calculate the 1-step and 5-step ahead forecast variances of the observations y_9 and y_{13}. ***(7 marks)***

(c) Provide a 95% (5-step ahead) prediction interval for the observation y_{13}. ***(1 mark)***

(ii) Consider that a time series $\{y_t\}$ is generated by the moving average model (MA):

$$y_t = \epsilon_t + \beta_1\epsilon_{t-1} + \beta_2\epsilon_{t-2},$$

where ϵ_t is a white noise sequence with variance σ^2 and the MA parameters β_1 and β_2 are assumed known.

Suppose this model is fitted to data $y_{1:n} = (y_1, \dots, y_n)$, for some given value of n. Based on these data, show that a 95% (2-step ahead) prediction interval for the observation y_{n+2} is given by

$$\beta_2\epsilon_n \pm 2\sigma\sqrt{1 + \beta_1^2}.$$

(8 marks)

3 It is well known that an economy's growth as measured by gross domestic product (GDP) is related to unemployment rate. Arthur Okun studied how much GDP is likely to fall, if unemployment increased by a certain level. Let y_t denote the GDP growth of the UK economy at time t and let x_t denote the unemployment rate at time t. A simple regression between the two can reveal the likely decrease of the GDP growth, for an increase of the unemployment rate. However, a more elaborate analysis, considers the following dynamic regression model:

$$y_t = \alpha + \gamma_t x_t + \epsilon_t,$$

where α is a static intercept, ϵ_t is a Gaussian white noise with variance 1 and γ_t is a time-varying slope, which follows the autoregressive model

$$\gamma_t = 0.3\gamma_{t-1} + \nu_t,$$

with ν_t a Gaussian white noise with variance 2. It is further assumed that ϵ_t and ν_s are independent for any t, s and that γ_0 is independent of ν_t, for any t.

(i) Define the state vector

$$\beta_t = \begin{bmatrix} \alpha \\ \gamma_t \end{bmatrix}.$$

Write the above model in state space form,

$$y_t = L_t^T \beta_t + \kappa_t,$$
$$\beta_t = F\beta_{t-1} + \zeta_t,$$

and determine the design vector L_t and the evolution matrix F. Write down κ_t and ζ_t and obtain their distributions. ***(4 marks)***

(ii) The above state space model is fitted to data of length n. The posterior distribution of β_n, given information $y_{1:n} = \{y_1, \dots, y_n\}$ is

$$\beta_n \mid y_{1:n} \sim N\left\{ \begin{bmatrix} -1.5 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix} \right\}.$$

(a) If $x_{n+1} = 0.5$ and $y_{n+1} = 1$, then obtain the posterior distribution $p(\beta_{n+1} \mid y_{1:n+1})$ of β_{n+1}, given information $y_{1:n+1}$. ***(14 marks)***

(b) Provide a 95% credible interval for γ_{n+1}, given $y_{1:n+1}$. ***(2 marks)***

End of Question Paper