Version SJG/3

EGT3
ENGINEERING TRIPOS PART IIB

Monday 1 May 2023 9.30 to 11.10

Module 4F7

STATISTICAL SIGNAL ANALYSIS

*Answer not more than **three** questions.*

All questions carry the same number of marks.

*The **approximate** percentage of marks allocated to each part of a question is indicated in the right margin.*

*Write your candidate number **not** your name on the cover sheet.*

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed
Engineering Data Book

10 minutes reading time is allowed for this paper at the start of the exam.

You may not start to read the questions printed on the subsequent pages of this question paper until instructed to do so.

You may not remove any stationery from the Examination Room.

1 Let X be a random variable, with mean $E[X] = x$, and let $Y_1, \ldots, Y_n$ be a further collection of n random variables with $E[Y_i] = m_i$. A linear estimate of X using $Y_1, \ldots, Y_n$ is

$$\hat{X} = h_0 + h_1\ (Y_1 - m_1) + \ldots + h_n\ (Y_n - m_n).$$

(a) The constants $h_0, \ldots, h_n$ that minimise the error $E[(X - \hat{X})^2]$ can be expressed as

$$A \begin{bmatrix} h_0 \\ \vdots \\ h_n \end{bmatrix} = b.$$

Find expressions for the matrix A and vector b in terms of appropriate expectations. [30%]

(b) Assume $E[X] = 0$. A sensor provides n noisy measurements $Y_1, \ldots, Y_n$ of X,

$$Y_i = |X| + W_i$$

for $i = 1, \ldots, n$, where W_i are independent zero mean random variables with variance σ^2. Solve for $h_0, \ldots, h_n$ based on your answer to Part (a). [30%]

(c) (i) Let X be a zero mean Gaussian random variable with variance s^2. Find $E[X|Y_1 = y_1]$, which is the conditional expected value of X given measurement $Y_1 = y_1$. (Hint: you may commence by executing the integral $\int xp(x, y_1)dx$ where $p(x, y_1)$ is the joint probability density function of (X, Y_1).) [20%]

(ii) Determine the estimate $\hat{X}$ under this Gaussian scenario and the observation model of Part (b). [20%]

2 Consider the following state transition probabilities for a Markov model $X_0, X_1, \ldots$ which describes the population count of a prey-predator pair:

$$p\left(X_{k+1} = (c_{k+1}, d_{k+1}) \,|\, X_k = (c_k, d_k)\right)$$
$$= \begin{cases} \theta_1 c_k, & \text{if } (c_{k+1}, d_{k+1}) = (c_k + 1, d_k) \text{ and } c_k < M_c \\ \theta_2 c_k d_k, & \text{if } (c_{k+1}, d_{k+1}) = (c_k - 1, d_k + 1) \text{ and } d_k < M_d \\ \theta_3 d_k, & \text{if } (c_{k+1}, d_{k+1}) = (c_k, d_k - 1) \\ 1 - \theta_1 c_k - \theta_2 c_k d_k - \theta_3 d_k & \text{if } (c_{k+1}, d_{k+1}) = (c_k, d_k) \end{cases}$$

where c_k and d_k are the number of prey and predators at time k (both being non-negative integers). Row 1 corresponds to an increase in prey count, row 2 predation and simultaneous predator population growth, row 3 is predator decline. Here M_c and M_d denote the maximum population count the model permits for the prey and predator respectively. You may assume the initial population counts c_0 and d_0 are independent random variables.

The data received (or the *observations* in the context of a hidden Markov model) are the exact prey count, $\{c_0, c_1, \ldots, c_T\}$. An aim is to estimate the unobserved predator count.

(a) Draw the state transition diagram for X_k to X_{k+1} when $X_k = (0,0)$, $X_k = (0,1), \ldots, X_k = (0, M_d)$. [15%]

(b) Give an expression for $p(d_0, \ldots, d_n | c_0, \ldots, c_n)$ in terms of the transition probabilities $p(X_k | X_{k-1})$ and initial state probability $p(X_0)$. [15%]

(c) Given $p(d_{n-1} | c_0, \ldots, c_{n-1})$, find an expression for $p(d_n | c_0, \ldots, c_n)$. [30%]

(d) Find an expression for

$$\frac{d}{d\theta_i} \log p(c_0, \ldots, c_T),$$

writing your answer as a sum involving the probability mass function $p(d_0, \ldots, d_T | c_0, \ldots, c_T)$. [20%]

(e) Find

$$\frac{d \log p((c_k = 0, d_k) | (c_{k-1} = 0, d_{k-1}))}{d\theta_i}$$

for $i = 1, 2, 3$. [20%]

3 (a) An unknown scalar parameter θ with prior probability density function (pdf) $p(\theta)$ is observed in additive noise

$$Y = \theta + W$$

where $W \in [-\alpha, \alpha]$, for some $\alpha > 0$, is a uniform random variable.

(i) Using N independent samples from a pdf $q(\theta)$, give an importance sampling estimate for the integral $\int h(\theta)p(\theta|Y = y)\mathrm{d}\theta$ for the case when $p(y)$ can be evaluated. Write an expression for the variance of the importance sampling estimate. [20%]

(ii) Using $q(\theta) = p(\theta)$, find an expression for the variance of the importance sampling estimator as a function of the noise interval parameter α. [10%]

(b) Consider the following measurement model

$$Y_k = \theta + W_k, \qquad k = 1, 2, \ldots$$

where $W_k \in [-a_k, a_k]$ is a uniform random variable for some $a_k > 0$. Assume $W_1, W_2, \ldots$ are independent random variables.

(i) Determine the conditional pdf $\pi_k(\theta) = p(\theta|y_{1:k})$. Using $\pi_k(\theta)$, find $\pi_{k+1}(\theta)$. [10%]

(ii) Let $\theta_k^1, \theta_k^2, \ldots, \theta_k^N$ be N samples of θ and $w_k^1, w_k^2, \ldots, w_k^N$ be N random variables, known as *weights*, such that

$$E\left\{h(\theta_k^i)w_k^i\right\} = \int h(\theta)p(\theta, y_{1:k})\mathrm{d}\theta \qquad (1)$$

where $h(\theta)$ is any real-valued function. Find the function h such that

$$\frac{\mathrm{d}}{\mathrm{d}a_1}\log p(y_{1:k}) = \int h(\theta)p(\theta|y_{1:k})\mathrm{d}\theta$$

and find the importance sampling estimate of this gradient. [20%]

(iii) Let $J_1, \ldots, J_N$ be independent and identically distributed random variables with probability mass function

$$\Pr(J = j) = \frac{w_k^j}{\sum_{i=1}^N w_k^i}.$$

Using sequential importance sampling with resampling and the weighted samples of equation (1), find an estimate of

$$\int h(\theta)\pi_{k+1}(\theta)\mathrm{d}\theta.$$

[20%]

(c) Let $y_1 = y, y_2 = y, \ldots, y_T = y$ where y is the single data point from Part (a).

(i) Find $a_1, \ldots, a_T$ so that $\pi_T(\theta) = p(\theta | Y = y)$. [10%]

(ii) Discuss whether there is any potential benefit of a sequential importance sampling estimate of $p(\theta | Y = y)$ found using $y_1 = y, y_2 = y, \ldots, y_T = y$ compared to the estimate in Part (a). [10%]

4 The concentration X (in parts per-million) of a particular pollutant in the atmosphere has been found by a modeller to have mean $E[X] = m$ and variance σ_0^2.

You are given further observations of X,

$$Y_n = X + V_n \quad \text{for } n = 1, 2, \ldots$$

where $V_1, V_2, \ldots$ are independent and identically distributed zero-mean Gaussian random variables with variance $E[V_n^2] = \sigma_v^2$.

(a) Let $\widehat{X}_{n-1}$ be an estimate of X using $\{Y_1, \ldots, Y_{n-1}\}$. Upon receiving Y_n, the estimate of X is updated to

$$\widehat{X}_n = G_n \left(Y_n - \widehat{X}_{n-1} \right) + \widehat{X}_{n-1}$$

where G_n is a gain.

(i) Find $\widehat{X}_0$ which results in the estimate being unbiased for all n. [10%]

(ii) Find the value of the gain G_n that minimises the mean square error $E\left[(\widehat{X}_n - X)^2 \right]$. Carefully explain the steps in your derivation. [40%]

(b) The modeller further proposes the Gaussian density as an appropriate model, that is X is $\mathcal{N}(m, \sigma_0^2)$.

(i) Find the joint probability density function $p(x, y_1)$ of (X, Y_1) and then find $p(x|y_1)$. [20%]

(ii) Find the conditional probability density function of X given $Y_1 = y_1, \ldots, Y_n = y_n$, justifying the steps of your solution. [20%]

(iii) Give an unbiased estimate of X and verify its unbiasedness. [10%]

END OF PAPER