EGT2
ENGINEERING TRIPOS PART IIA

Wednesday 5 May 2021   9.00 to 10.40

Module 3F3

STATISTICAL SIGNAL PROCESSING

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 and at the top of each answer sheet.*

STATIONERY REQUIREMENTS
Write on single-sided paper.

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed.
You are allowed access to the electronic version of the Engineering Data Books.

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

The time taken for scanning/uploading answers is 15 minutes.

Your script is to be uploaded as a single consolidated pdf containing all answers.

1 (a) Let $\{\ldots, V_{-1}, V_0, V_1, \ldots\}$ be a zero mean *white noise* sequence, that is, $\mathbf{E}\{V_i^2\} = \sigma_v^2$ and $\mathbf{E}\{V_i V_j\} = 0$ for $i \neq j$. Find the autocorrelation function of X_n where

$$X_n = \sum_{j=0}^{t} V_{n-j}.$$

[15%]

(b) Let U be a uniform random variable in the interval $[0, 1]$ and let $X_n = \sin(2\pi U n)$ for $n = 0, 1, \ldots$ Show that $\{X_0, X_1, \ldots\}$ is wide sense stationary. [10%]

(c) Let $\{W_0, W_1, \ldots\}$ be a zero mean wide sense stationary sequence and

$$X_n = b_1 + b_2 n + W_n.$$

(i) Find $\mathbf{E}\{X_n\}$ and $\mathbf{E}\{X_n X_m\}$ and hence determine whether X_n is wide sense stationary. [15%]

(ii) Let $Y_n = X_n - X_{n-1}$ and show that $\{Y_1, Y_2, \ldots\}$ is wide sense stationary. [20%]

(iii) Assume further that $\{W_0, W_1, \ldots\}$ are independent and identically distributed and W_n has characteristic function $\phi(t)$. For constants $t_1, \ldots, t_k$, find the characteristic function $\mathbf{E}\{\exp(iY_1 t_1 + \cdots + iY_k t_k)\}$ of the random vector $(Y_1, \ldots, Y_k)$ from part (c)(ii). [30%]

(iv) Figure 1 shows the data set $X_1, \ldots, X_{50}$ with the linear trend $b_1 + b_2 n$ superimposed. Describe a procedure to assess whether W_n is wide sense stationary or independent in this example. [10%]

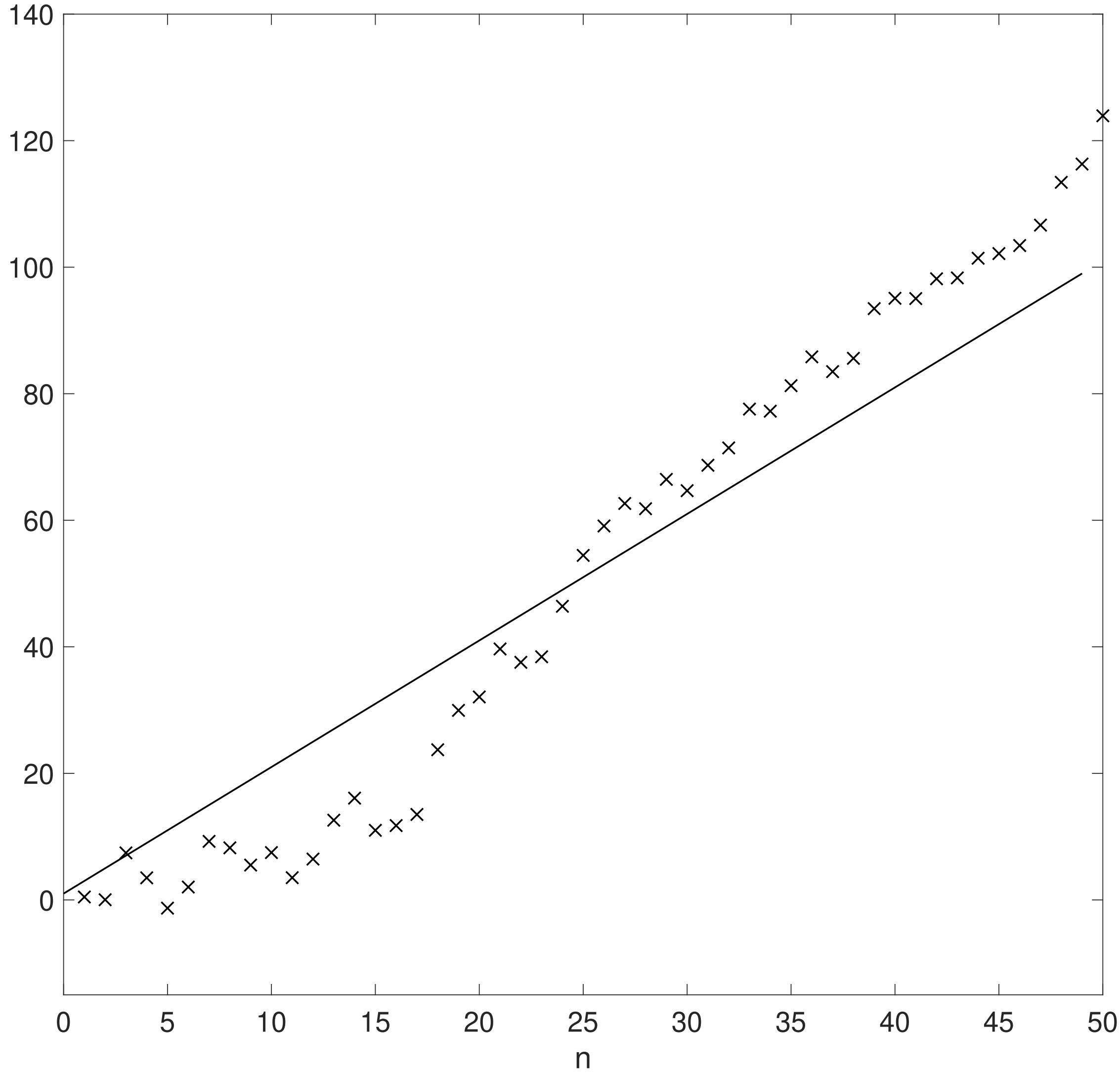

Fig. 1

(TURN OVER

2 (a) In a betting game, a player bets one pound at a time. Each time a bet is made, the player wins a pound with probability α (where $0 < \alpha < 1$) independently of the outcomes of previous bets. Let X_n be the wealth of the player after the outcome of the n-th bet. Explain why X_n is a Markov process. Draw its state transition diagram and give $p_{i,j} = \Pr(X_{n+1} = j | X_n = i)$ for all i, j. [15%]

(b) The player's strategy is to carry on betting until they have attained wealth M or they have lost everything.

(i) The player has an initial wealth of m pounds and let E_m denote the event that the player reaches their target wealth M when starting from m. List two possible outcomes in the event E_m.

(ii) Let $q_m = \Pr(E_m)$, for $m = 0, 1, \ldots, M$. Write q_m in terms of q_{m-1} and q_{m+1}.

[15%]

(c) Show that

$$q_m - q_{m-1} = \left(\frac{1-\alpha}{\alpha}\right)^{m-1} (q_1 - q_0).$$

[15%]

(d) Solve for q_1. [25%]

(e) The input U of an amplifier is $\mathcal{N}(0, 1)$ and its output is $X = ZU + Y$ where Z is the amplifier's random gain with probability density

$$f(z) = \frac{3}{7}z^2, \qquad 1 \leq z \leq 2.$$

The amplifier has a non-negative random bias Y with conditional probability density $p_{Y|Z}(y|z) = z \exp(-zy)$ for $y \geq 0$. Assume U is independent of Z and Y.

(i) Find the variance of the random variables X, Y and Z. [15%]

(ii) Find the covariance matrix of the random vector $[X, Y, Z]'$. [15%]

3 A random communications signal $\{X_n\}$ is observed through a noisy, reverberant channel as follows:

$$Y_n = 0.8X_n - 0.6X_{n-1} + V_n$$

where $\{V_n\}$ is a zero-mean white noise process independent of $\{X_n\}$ and having variance $\sigma_V^2 = 0.1$. $\{X_n\}$ is an iid random binary process taking values:

$$X_n = \begin{cases} +1, & \text{with probability } p \\ -1, & \text{with probability } 1-p \end{cases}$$

(a) Determine the mean of $\{X_n\}$ and show that the autocorrelation function for $\{X_n\}$ is

$$R_{XX}[n, n+m] = (2p-1)^2 + \delta_m(1-(2p-1)^2)$$

where

$$\delta_m = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases}$$

Hence state whether the process is wide-sense stationary. Determine also whether the process is mean ergodic. [20%]

(b) With $p = 0.5$, i.e. equally probable symbols, determine the cross-correlation function between $\{Y_n\}$ and $\{X_n\}$ and the autocorrelation function for the observed process $\{Y_n\}$. [35%]

(c) An FIR filter with coefficients α and β is applied to $\{Y_n\}$ in order to estimate the binary signal as follows:

$$\hat{X}_n = \alpha Y_n + \beta Y_{n-1}$$

Determine the optimal value of the coefficients in order to estimate X_n with minimum mean-squared error, assuming symbols are equally probable ($p = 0.5$). [25%]

(d) What is the mean-squared error for the optimum filter? Would this be adequate for determining whether $+1$ or -1 has been transmitted, assuming that the estimation error $\hat{X}_n - X_n$ is approximately zero mean and Gaussian? [20%]

4 Some experimental climate data is believed to be normally distributed around the long term trend. However, some extreme values are occasionally observed, which leads a scientist to model the data as normal, but with a random *precision* parameter U_t. The model for the measurements at time t is as follows:

$$p(Y_t|U_t) = \mathcal{N}(Y_t|0, 1/U_t)$$

where $\mathcal{N}(X|0,V) = \frac{1}{\sqrt{2\pi V}} \exp(-\frac{1}{2V}X^2)$.

(a) From a single data point Y_t, determine the maximum likelihood estimate for the precision parameter U_t and sketch the form of the likelihood as a function of U_t. [20%]

(b) Comment on whether this is likely to be a useful estimate of the precision in practice. [10%]

(c) As an alternative procedure, an exponential prior distribution is assigned to U_t, taking the form

$$p(U_t) = \lambda \exp(-\lambda U_t), \quad U_t \geq 0$$

Determine the posterior probability density function for U_t using this prior and determine the maximum *a posteriori* estimator for U_t. Sketch the posterior density and compare its behaviour with that of the likelihood function in part (a), illustrating what happens as Y_t^2 becomes small, intermediate and large relative to λ. [30%]

(d) Using the same prior and likelihood functions as in parts (a) and (c), determine the *marginal* probability of the data point Y_t, $p(Y_t)$. Sketch this density function and comment on its behaviour for modelling data with extreme values, compared with the normal distribution. [40%]

You may require the following definition of the gamma function, which is a well known tabulated function:

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x}\, dx, \quad z > 0$$

END OF PAPER