Version SJG/3

EGT2
ENGINEERING TRIPOS PART IIA

Wednesday 25 April 2018 9.30 to 11.10

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.*

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.

1 (a) Let $f(i)$ be a probability mass function (pmf), i.e. $f(i) \geq 0$ and $\sum_{-\infty}^{\infty} f(i) = 1$. Let

$$X_k = X_{k-1} + W_k, \quad \text{for } k = 1, 2, \ldots$$

where $X_0 = i_0$ and $W_1, W_2, \ldots$ are independent and identically distributed random variables and each W_k has pmf f.

(i) Find $p\left(X_{k+1} = j \mid X_k = i\right)$; [10%]

(ii) Find $p(X_1 = i_1, \ldots, X_n = i_n)$; [10%]

(iii) Hence show that $X_1, X_2, \ldots$ is a Markov chain. [20%]

(b) Let $f(-1) = 1/2$, $f(1) = 1/2$ (thus $f(i) = 0$ for all other values of i) and let

$$Y_n = \left(\sum_{j=1}^{n} W_j \right) / \sqrt{n}$$

(c) By computing the characteristic function of Y_n, which is $E\left\{\exp\left(iY_n t\right)\right\}$, show that Y_n tends to a Gaussian random variable as $n \to \infty$. (Hint: you may use the fact that $\cos(t/\sqrt{n})^n \to \exp\left(-t^2/2\right)$ as n tends to infinity.)

(d) A gambler, with initial wealth R, wagers one pound for each bet and the probability of winning the bet is 0.5. Their wealth increases by 1 if the bet is won; otherwise it decreases by 1.

(i) The gambler is allowed to make n successive bets, potentially going into debt. Find an approximation for $\Pr(X_n > 0)$. [25%]

(ii) In a change of the rules, the gambler is allowed to make n successive bets but must stop as soon as their wealth is zero. Give the Markov chain that describes the change in wealth of the gambler and comment on how well the answer in (c)(i) approximates the probability that the gambler's wealth is positive. [10%]

2 Consider the following autoregressive process

$$X_n + a_1 X_{n-1} + a_2 X_{n-2} = \sigma W_n$$

where $\{W_n\}$ is a zero-mean white noise process with variance 1 and σ a positive constant.

(a) Find the power spectrum of $\{X_n\}$. [10%]

(b) Let $\{V_n\}$ be the moving average process

$$V_n = b_0 E_n + b_1 E_{n-1}$$

where $\{E_n\}$ is a zero-mean white noise process with variance 1. Find the power spectrum of $\{V_n\}$. [10%]

(c) Let Y_n be the noisy measurement of X_n given by

$$Y_n = X_n + V_n.$$

Assume the noise sequences $\{W_n\}$ and $\{E_n\}$ are independent. Find the power spectrum of $\{Y_n\}$. [15%]

(d) Based on measurements of $\{V_n\}$, the power spectrum of $\{V_n\}$ is estimated to be

$$\widehat{S}_V(\omega) = 2 + 2\cos\omega$$

Show that valid estimates of b_0 and b_1 are $b_0 = b_1 = 1$ and $b_0 = b_1 = -1$. [25%]

(e) Show that

$$\begin{bmatrix} R_{XX}[0] & R_{XX}[1] \\ R_{XX}[1] & R_{XX}[0] \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = -\begin{bmatrix} R_{XX}[1] \\ R_{XX}[2] \end{bmatrix},$$

$$R_{XX}[0] + a_1 R_{XX}[1] + a_2 R_{XX}[2] = \sigma^2.$$

[30%]

(f) Based on measurements of Y_n as in (c), the following estimates are made for its autocorrelation function:

$$\widehat{R}_{YY}[0] = 4.74, \qquad \widehat{R}_{YY}[1] = 0.54, \qquad \widehat{R}_{YY}[2] = 1.41$$

Use these values to estimate the power spectrum of $\{X_n\}$. [10%]

3 (a) Define the term white noise. What is the form of the autocorrelation function for white noise, and what is its power spectrum? [15%]

(b) A pulse waveform $s_n = n/N$, $n = 0, 1, ..., N$, is buried in noise at a sample time n_0, i.e. the noisy signal is:

$$x_n = \begin{cases} s_{n-n_0} + v_n, & n - n_0 = 0, 1, ..., N \\ v_n, & \text{otherwise,} \end{cases}$$

where v_n is white, zero-mean noise with variance σ_v^2.

A Finite Impulse Response (FIR) smoothing filter with N coefficients $[1, 1, ..., 1]$ is applied to the noisy waveform x_n.

(i) Show that the output y_n of this filter has variance $N\sigma_v^2$ when only the noise v_n is input to the filter (i.e. considering s_{n-n_0} to be zero in the above equation for x_n). [15%]

(ii) Sketch the output of the filter when $\sigma_v = 0$, i.e. no noise is present, but $s_n = n/N$, $n = 0, ..., N$, as in (b) above. [20%]

(iii) What is the maximum expected signal-to-noise ratio (SNR) at the output of the filter, when applied to noisy pulse data x_n, i.e. with $\sigma_v > 0$, and at what value of n does this occur? [20%]

(iv) Now design the optimal filter for detection of the location n_0 (no derivation is required) and compare its performance (in terms of expected SNR) with that of the FIR smoothing filter in (b) above, as N becomes large. [30%]

You may use the result:

$$\sum_{n=1}^{N} n^2 = \frac{N(N+1)(2N+1)}{6}$$

4 (a) A symmetric uniform probability density function (pdf) is defined as

$$f_Y(y|a) = \begin{cases} 1/(2a), & -a \leq y \leq a \\ 0, & \text{otherwise} \end{cases}$$

A sequence of independent measurements $y_1, y_2, y_3, ..., y_n$ is made from a noise source with this pdf, but the scaling of the noise, a, is unknown.

(i) Determine the likelihood function for a when $n = 1$ and when $n = 2$. Sketch these as a function of a, marking on the maximum likelihood estimator in each case. [15%]

(ii) Determine now the maximum likelihood (ML) estimator for a in the general case of n measurements $y_1, y_2, y_3, ..., y_n$, $n \geq 1$. Is this estimate likely to be unbiased for finite n? Explain also whether the estimate will be unbiased as $n \to \infty$. [20%]

(b) A communications network is monitored. It is desired to find the average rate of symbols, λ. The time τ between each symbol is independently and randomly distributed as an exponential random variable with mean $1/\lambda$:

$$f(\tau|\lambda) = \lambda \exp(-\lambda\tau), \ \tau \geq 0$$

(i) The times of arrival of n successive symbols are now measured as $t_1, t_2, t_3, ..., t_n$, where t_0, the first symbol's arrival time, is zero. Show that the likelihood function for λ is:

$$\lambda^n \exp(-\lambda t_n)$$

and determine the ML estimate of λ. [20%]

(ii) Prior information about the network traffic states that λ is distributed in the following way:

$$f(\lambda) = \lambda b^2 \exp(-\lambda b), \ \lambda \geq 0$$

Determine the Bayesian posterior density for λ (including its normalising constant) and the Maximum *a posteriori* (MAP) estimator, given $t_0 = 0$ and measurements $t_1, ..., t_n$ as above. [20%]

(iii) Sketch the prior density, likelihood function and posterior density for λ, marking the MAP and ML estimators clearly on the sketch and commenting on their relationship to the prior. Use the following parameter values: $b = 1$, $n = 4$, $t_4 = 5$. [25%]

In Part (b) you may use the following result, that applies for any integer n:

$$\int_0^\infty x^n \exp(-x)dx = n!$$

END OF PAPER