EGT2
ENGINEERING TRIPOS PART IIA

May 2024

Module 3F3

STATISTICAL SIGNAL PROCESSING - WORKED SOLUTIONS

STATIONERY REQUIREMENTS

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM

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 *Examiners' Comment: Well answered in general, but there were a few recurring errors. E.g., a considerable minority of answers did not display knowledge of how to change variables in a probability density function. Several parts also prompted unnecessary algebra, in particular (c) and (e). Overall, there is a reasonable spread in the answers that displays a range of understanding of material. Several students received full marks.*

(a) Because $\frac{\mathrm{d}w}{\mathrm{d}v} = \frac{1}{v}$,

$$P(w) = \frac{P(v)}{\left|\frac{\mathrm{d}w}{\mathrm{d}v}\right|} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(w+1/4)^2}{2}\right),$$

which is by definition the pdf for a normal distribution with mean $-1/4$ and variance 1.

(b) $X_n = \sum_{i=1}^{n} w_i$ so $E[X_n] = nE[w_i] = -n/4$ and $\mathrm{Var}[X_n] = n\mathrm{Var}[w_i] = n$.

(c) $E[X_n'] = 0$ as required but $\mathrm{Var}[X_n'] = \mathrm{Var}[X_n] = n$. Thus, $\mathrm{Cov}(X_n', X_{n+k}') = n$, which depends on n, so X_n' is not WSS. It is Markov.

(d) $E[Y_n] = E[v_n]Y_{n-1} = e^{n/4}Y_0$. In expectation, our initial investment is growing very rapidly.

(e) X_n is normally distributed with $\mu = -n/4$ and $\sigma^2 = n$. Setting $\alpha = n/8$ in Eq. (3) we get

$$P[X_n > -n/8] \leq \frac{4}{\sqrt{n}} \exp(-n/128)$$

(f) As n becomes large, X_n is practically guaranteed to be less than $-n/8$. In which case, Y_n will be less than $e^{-n/8}$, and so you are virtually guaranteed to lose the money as n becomes large. From this perspective, the investment appears to be a bad idea.

In part (d) it was shown that the expected value of Y_n grows exponentially. The apparent tension occurs because there is an exponentially small chance of making an exponentially large amount of money. In the limit $n \to \infty$, there is technically a possibility of making an arbitrarily large amount of money, but you will actually lose everything with probability 1.

2 *Examiners' Comment: The least popular question, although still attempted by well over half of candidates. It was also the lowest average score. Minimal algebra is required for this question, and several good answers completed it in 2 sides with answers that match the model solution almost verbatim. The weaker answers were confused about the meaning of* $E\left[z^X\right]$*, where* X *is a discrete random variable and* z *is an arbitrary number. Several students failed to write down any equation that looked like an expected value, while others made faulty assumptions such as that* $E\left[z^X\right] = z^{E[X]}$.

(a) $P(X_1 = k, X_2 = j) = \frac{\lambda^{k+j} e^{-2\lambda}}{k!j!}$

(b)

$$\begin{aligned} P(X_1 + X_2 = n) &= \sum_{k=0}^{n} P(X_1 = k, X_2 = n - k) \\ &= \sum_{k=0}^{n} \frac{\lambda^n e^{-2\lambda}}{k!(n-k)!} \\ &= \frac{(2\lambda)^n e^{-2\lambda}}{n!} \end{aligned}$$

(c)

$$G_X(z) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(z\lambda)^k}{k!} = e^{-\lambda} e^{z\lambda} = e^{\lambda(z-1)}$$

(d)

$$E[z^S|k] = E[z^{X_1 + \dots X_k}|k] = \prod_{i=1}^{k} E[z^{X_1}] = G_X(z)^k = e^{k\lambda(z-1)}$$

where the second equality holds by independence.

(e)

$$G_S(z) = E[z^S] = E\big[E[z^S|k]\big] = E\big[G_X(z)^k\big] = G_X(G_X(z))$$

The expected value $E[z^S]$ does not have the same functional form as it would if S were Poisson distributed. Thus, S is not Poisson distributed.

3 *Examiners' Comment:A very popular question, very well answered by most students. Basics of AR models and Wiener filters well known. Most students were accurate with algorithmic derivations of MSE and for the interpretation of noise variances going to zero. Many students did not properly simplify the MSE at the optimal filter solution in part (e).*

A random process is defined as

$$X_n = \alpha X_{n-1} + W_n$$

where $|\alpha| < 1$ and $\{W_n\}$ is zero-mean white noise with variance σ_W^2.

(a) Show that the autocorrelation function for $\{X_n\}$ satisfies, for $k > 0$

$$r_{XX}[k] = \alpha r_{XX}[k-1]$$

[15%]

(b) Show that

$$r_{XX}[0] = \frac{\sigma_W^2}{1-\alpha^2}$$

and hence show that

$$r_{XX}[k] = \frac{\sigma_W^2}{1-\alpha^2}\alpha^{|k|}$$

[15%]

Solution:

Various methods are possible. e.g.

$X_{n+1} = \alpha X_n + W_{n+1}$

$X_{n+2} = \alpha(\alpha(X_n + W_{n+1}))+W_{n+2}$

$X_{n+k} = \alpha^k X_n + \alpha^{k-1} W_{n+1} + \alpha^{k-2} W_{n+2} + \ldots + W_{n+k}$

So,

$$r_{XX}[k] = E[X_n X_{n+k}] = \alpha^k r_{XX}[0] \quad (1)$$

since $E[X_n W_{n+i}] = 0$ for $i > 0$.

But,

$$r_{XX}[0] = E[X_n(\alpha X_{n-1} + W_n)] = r_X X[1] + E[X_n W_n] = \alpha r_{XX}[0] + \sigma_W^2$$

since

$$E[X_n W_n] = \sigma_W^2,$$

and hence

$$r_{XX}[0] = \frac{\sigma_W^2}{1-\alpha^2}$$

and substituting in (**??**) leads to

$$r_{XX}[k] = \frac{\sigma_W^2}{1-\alpha^2}\alpha^{|k|}$$

noting that $r_{XX}[k] = r_{XX}[-k]$.

(c) Noisy observations of $\{X_n\}$ are made in which

$$Y_n = X_n + V_n$$

where $\{V_n\}$ is zero-mean white noise with variance σ_V^2, independent of $\{X_n\}$. Determine the autocorrelation function of $\{Y_n\}$ and the cross-correlation function between $\{X_n\}$ and $\{Y_n\}$, as a function of α, σ_W^2 and σ_V^2. [20%]

Solution:

$$\begin{aligned} r_{YY}[k] &= E[(X_n + V_n)(X_{n+k} + V_{n+k})] \\ &= r_{XX}[k] + \sigma_V^2 \delta_k \end{aligned}$$

since all $E[XV] = E[XV]$ terms are zero (independent and zero mean V).

$$\begin{aligned} r_{XY}[k] &= E[(X_n)(X_{n+k} + V_{n+k})] \\ &= r_{XX}[k] \end{aligned}$$

which should be expanded out in terms of the form of r_{XX} in part (a) and (b).

(d) It is desired to predict X_{n+p} where $p > 0$ is a lookahead of p samples, using just the current observed data point Y_n, i.e.

$$\hat{X}_{n+p} = hY_n$$

where h is a coefficient to be determined. Derive the value of h which minimises the mean-squared error of the prediction. [25%]

Solution:

Expected error is $\epsilon = E[(hY_n - X_{n+p})^2]$:

Minimum is achieved when:

$$\frac{d\epsilon}{dh} = 0$$

i.e.

$$E[2(hY_n - X_{n+p})Y_n] = 2hr_{YY}[0] - 2r_{YX}[p] = 0$$

which is satisfied when $h = r_{YX}[p]/r_{YY}[0]$, and using result from above:

$$h = r_{XX}[p]/(r_{XX}[0] + \sigma_V^2) = \frac{\sigma_W^2}{1-\alpha^2}\alpha^{|p|}/(\frac{\sigma_W^2}{1-\alpha^2} + \sigma_V^2) = \alpha^{|p|}/(1 + (1-\alpha^2)\sigma_V^2/\sigma_W^2)$$

(e) Derive the mean-squared error for this optimal predictor, and show that as σ_V becomes small the error tends towards:

$$\sigma_W^2 \frac{1-\alpha^{2p}}{1-\alpha^2}.$$

Explain in terms of the model for $\{X_n\}$ why this error tends to zero as σ_W^2 becomes small.

Solution:

At the optimum,

$$\epsilon = E[(hY_n - X_{n+p})^2] = h^2 r_{YY}[0] + r_{XX}[0] - 2hr_{YX}[p] = r_{XX}[0] - hr_{YX}[p]$$

since $2hr_{YY}[0] - 2r_{YX}[p] = 0$ at the optimum solution.

Substituting in, we get:

$$\epsilon = r_{XX}[0](1 - \alpha^{2|p|}/(1 + (1-\alpha^2)\sigma_V^2/\sigma_W^2) = \frac{\sigma^{W2}}{1-\alpha^2}(1 - \alpha^{2|p|}\sigma_W^2/(\sigma_W^2 + (1-\alpha^2)\sigma_V^2)$$

Then with small σ_V^2 the term $\sigma_W^2/(\sigma_W^2 + (1-\alpha^2)\sigma_V^2) \to 1$ and we get

$$\epsilon = \frac{\sigma^{W2}}{1-\alpha^2}(1 - \alpha^{2|p|}$$

as required. We expect this to tend to zero as σ_W^2 becomes small since then the AR model for $\{X_n\}$ is predictable with zero error.

4 *Examiners' comment: Another very popular question and well answered. In part e) most students had good intuition about the posterior behaviour as N becomes large, but most did not comment that the posterior density becomes more concentrated about the MAP/ML solution.*

The exponential random variable X has probability density function

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

(a) Determine from first principles the mean μ and variance σ^2 of the distribution. [15%]

Solution:

Integrate by parts:

$$\mu = E[X] = \int_0^\infty x\lambda e^{-\lambda x} dx = [-xe^{-\lambda x}]_0^\infty + \int_0^\infty e^{-\lambda x} dx = 0 - \frac{1}{\lambda}[e^{-\lambda x}]_0^\infty = \frac{1}{\lambda}$$

$$E[X^2] = \int_0^\infty x^2 \lambda e^{-\lambda x} dx = [-x^2 e^{-\lambda x}]_0^\infty + \int_0^\infty 2xe^{-\lambda x} dx = \frac{2}{\lambda^2}$$

and hence

$$var[X] = E[X^2] - E[X]^2 = \frac{1}{\lambda^2}$$

(b) Given N independent samples drawn from the exponential distribution, determine the maximum likelihood ML estimate for λ and for μ. [30%]

Solution:

Samples $x_1, ..., x_N$.

Likelihood function is

$$p(x_1, ..., x_N|\lambda) = \prod_{i=1}^{N} \lambda \exp(-\lambda x_i) = \lambda^N \exp(-\lambda \sum_i x_i)$$

Now ML solution for λ is found by:

$$dp(x_1, ..., x_N|\lambda)/d\lambda = N\lambda^{N-1} \exp(...) - \lambda^N \sum_i x_i \exp(...) = 0$$

and hence

$$\hat{\lambda} = N/\sum_i x^i$$

Now $\mu = 1/\lambda$ which is a monotonic function, so the ML solution for μ is just $\hat{\mu} = (1/N)\sum_i x^i$. This needs to be justified in the answer. If not then a full calculus solution would be required, as I expect many students will do...

(c) Is the estimate for μ unbiased? [15%]

Solution:

Check its expected value:

$$E[\hat{\mu}] = E[(1/N)\sum_i x^i] = (1/N)\sum_i E[x^i] = (1/N)\sum_i \mu = \mu$$

hence it is unbiased.

(d) What is the variance of the ML estimate for μ? Explain whether the ML estimate is therefore *consistent*. [15%]

Solution:

A quick solution would require the result $var(k\sum Z_i) = k^2 \sum var(Z_i)$:

$$var[\hat{\mu}] = (1/N)^2 \sum_i var(x^i) = (1/N)^2 \sum_i var[x^i] = \frac{1}{N}\frac{1}{\lambda^2}$$

and hence it is consistent since $var[\hat{\mu}] \to 0$ as $N \to \infty$ and the estimator is unbiased.

A less advanced answer expected from most undergraduates would compute:

$$E[\hat{\mu}^2] = (1/N^2)E[(\sum_i x^i)^2] = (1/N^2)(\sum_i \sum_j E[x^i x^j]) = (1/N^2)(N^2\mu^2 + N\sigma^2)$$

since $E[x^i x^j] = \mu^2 + \sigma^2$ when $i = j$ and just μ^2 when $i \neq j$.

Thus $var[\hat{\mu}] = E[\hat{\mu}^2] - E[\hat{\mu}]^2 = \frac{1}{N}\frac{1}{\lambda^2}$, as before.

(e) A Bayesian prior is to be assigned to λ,

$$p(\lambda) \propto \lambda^{\alpha-1} e^{-\lambda\beta}$$

where $\alpha > 0$ and $\beta > 0$. Determine the MAP estimator for λ, and comment on its behaviour as N becomes large. Sketch the posterior probability density function for λ; how would you expect this density to change as N increases? [25%]

Solution:

Posterior distribution:

$$p(\lambda|x_1,...,x_N) \propto p(\lambda)p(x_1,...,x_N|\lambda) = \lambda^{N+\alpha-1}\exp(-\lambda(\beta + \sum_i x_i))$$

The MAP estimator is the maximiser of this:

$$dp(\lambda|x_1,...,x_N)/d\lambda = (N+\alpha-1)\lambda^{N+\alpha-2}\exp(...) - (\beta + \sum_i x_i)\lambda^{N+\alpha-1}\exp(...) = 0$$

which is solved for

$$\lambda^{MAP} = (N+\alpha-1)/(\beta + \sum_i x_i)$$

As N becomes large $(N+\alpha-2)/(\beta+\sum_i x_i) \to (N)/(\sum_i x_i)$, i.e. tends to the ML solution.

Sketch has a unique maximum at λ^{MAP}, decaying to zero either side. As N increases the peak around the maximum becomes narrower and narrower and the peak tends to the ML estimate.

END OF PAPER