THE STATE UNIVERSITY OF NEW JERSEY

RUTGERS

College of Engineering
Department of Electrical and Computer Engineering

332:541 **Stochastic Signals and Systems** **Fall 2007**

Quiz III

There are 4 questions. You have three hours to answer them. Show all work. Answers given without work will receive no credit. GOOD LUCK!

1. *(35 points)* Let k be a parameter which can take on values $k = 0, 1, \cdots$. We form a random variable R as

$$R = \sum_{\ell=1}^{k^2} G_\ell$$

where the G_ℓ are i.i.d. zero mean unit variance Gaussian random variables. If $k = 0$ we define $R = 0$

(a) *(5 points)* Given k, what is the probability density for R?

SOLUTION: *Given k we have a sum of k Gaussians so R is a zero mean Gaussian with variance k^2.*

$$f_{R|k}(r|k) = \sqrt{\frac{1}{2\pi}}\frac{1}{k}e^{-\frac{r^2}{2k^2}}$$

Notice that this probability model comports with the definition of R since at $k = 0$, R has an impulse distribution at $R = 0$.

(b) *(10 points)* What is the ML estimate for k, $\hat{k}(R)$? What is $E[\hat{k}(R)|k]$? Is this estimate biased or unbiased?

SOLUTION: *We maximize $f_{R|k}(r|k)$ in k. We can take the log to make life easier*

$$\frac{d \ln f_{R|k}(r|k)}{dk} = -\frac{1}{k} + \frac{r^2}{k^3} = \frac{1}{k}(\frac{r^2}{k^2} - 1)$$

which is strictly monotone decreasing in k. Since the value at $k = 0$ is $+\infty$, $f_{R|k}(r|k)$ initially increases and then decreases. So when we find the value at which the derivative is zero, we'll have found the function maximum. Thus, accounting for the fact that k must be non-negative,

$$\hat{k}(R) = |R|$$

$E[\hat{k}(R)|k] = \sqrt{\frac{2}{\pi}}k \neq k$ *so the estimate is biased.*

(c) *(10 points)* What is the ML estimate for k^2, $\hat{k}^2(R)$? What is $E[\hat{k}^2(R)|k]$? Is this estimate biased or unbiased?

SOLUTION: *Letting $N = k^2$ we have*

$$\frac{d \ln f_{R|\sqrt{n}}(r|\sqrt{n})}{dn} = -\frac{1}{2n} + \frac{r^2}{2n^2} = \frac{1}{2n}(\frac{r^2}{n} - 1)$$

so that

$$\hat{k}^2(R) = R^2$$

This estimator has $E[\hat{k}^2(R)|k] = k^2$ *so the estimate is unbiased.*

(d) *(10 points)* Suppose we have a series of identically composed, but independent measurements R_m where $m = 1, 2, \cdots, M$. What is the ML estimate for k^2 based on these M measurements? Is the estimate biased or unbiased? Is the estimate consistent?

SOLUTION: *The* R_m *are independent so*

$$f_{\mathbf{R}|k}(\mathbf{r}|\sqrt{n}) = \left(\frac{1}{2\pi n}\right)^{M/2} e^{-\frac{1}{2n}\sum_{m=1}^{M} r_m^2}$$

Defining $\rho^2 = \sum_{m=1}^{M} r_m^2$ *we have,*

$$\frac{d \ln f_{\mathbf{R}|k}(\mathbf{r}|\sqrt{n})}{dn} = -\frac{M}{2n} + \frac{\rho^2}{2n^2} = \frac{1}{n}(\frac{\rho^2}{n} - M)$$

so that

$$\hat{k}^2(\mathbf{r}) = \frac{1}{M}\sum_{m=1}^{M} r_m^2$$

The mean of the estimator is

$$E\left[\frac{1}{M}\sum_{m=1}^{M} r_m^2\right] = k^2$$

so the estimator is unbiased. We also see that $\hat{k}^2(\mathbf{r})$ *is the sample mean for* R^2 *and must therefore converge to the mean of* R^2 *with probability 1. Since the mean of* $E[R^2] = k^2$*, the estimator is consistent.*

2. *(35 points)* Suppose Y and X are zero mean jointly Gaussian random variables and you wish to estimate Y from X. If we define $\mathbf{Z}$ as

$$\mathbf{Z} = \begin{bmatrix} X \\ Y \end{bmatrix}$$

then the joint distribution is

$$f_{XY}(x, y) = f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\mathbf{z}^\top \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix} \mathbf{z}}$$

where $-1 < \rho = E[XY] < 1$

(a) *(10 points)* What is the *Linear* MMSE estimate of Y given x?

SOLUTION: $\hat{Y}(x) = \frac{\rho}{\sigma_x^2} x = \rho x$

(b) *(10 points)* Please derive the maximum likelihood estimate of Y given x.

SOLUTION: *First we need the marginals for X and Y – both are going to be Gaussian since they're jointly Gaussian. The inverse of the covariance matrix is*

$$\mathbf{K}^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$$

and we recognize then verify that

$$\mathbf{K} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

so that $\sigma_x^2 = \sigma_y^2 = 1$.

Another approach would be to integrate the joint PDF which we expand out from the definition

$$f_X(x) = \int_{-\infty}^{\infty} \frac{1}{2\pi}\frac{1}{\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)} dy$$

We rearrange the exponent to obtain

$$f_X(x) = \frac{1}{2\pi}\frac{1}{\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(\rho^2x^2-2\rho xy+y^2+(1-\rho^2)x^2)} dy$$

and thence

$$f_X(x) = \frac{1}{2\pi}\frac{1}{\sqrt{1-\rho^2}} e^{-\frac{x^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(y-\rho x)^2} dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

a zero mean unit variance Gaussian. Symmetry dictates the same result for y.

We need to maximize $f_{X|Y}(x|y)$ in y for the ML estimate.

$$f_{X|Y}(x|y) = f_{\mathbf{Z}}(\mathbf{z})/f_Y(y) = \sqrt{\frac{1}{2\pi}}\frac{1}{\sqrt{1-\rho^2}} e^{-\frac{1}{2}\left((x^2-2\rho xy+y^2)-(1-\rho^2)y^2\right)}$$

The only thing that matters in the maximization is the exponent. We minimize the exponent in y and find $\hat{Y}_{ml}(x) = x/\rho$.

Some of you did the MAP estimate which is the maximum of $f_{\mathbf{Z}}(\mathbf{z})$ with respect to y. I gave full credit if you did this correctly. Once again the exponent is all that matters and $x^2 - 2\rho xy + y^2$ is minimzed in y when $y = \rho x$, so $\hat{Y}_{map}(x) = \rho x$.

(c) *(15 points)* Please derive the MMSE estimate of Y given x.

SOLUTION: *To obtain the expected value of Y given x we need*

$$f_{Y|X}(y|x) = f_{\mathbf{Z}}(\mathbf{z})/f_X(x) = \sqrt{\frac{1}{2\pi}}\frac{1}{\sqrt{1-\rho^2}} e^{-\frac{1}{2}\left((x^2-2\rho xy+y^2)-(1-\rho^2)x^2\right)}$$

Rearranging we have

$$f_{Y|X}(y|x) = \sqrt{\frac{1}{2\pi}}\frac{1}{1-\rho^2} e^{-\frac{1}{2(1-\rho^2)}(\rho x-y)^2}$$

so the conditional mean of Y given x is $\hat{Y}(x) = \rho x$.

3. *(30 points)*

A discrete time linear system has a random input $u(t)$ with

$$p_{U(t)}(u(t)) = \begin{cases} \alpha & u(t) = 1 \\ 1-\alpha & u(t) = -1 \end{cases}$$

where $0 < \alpha < 1$. The $u(t)$ at different points in time are all mutually independent. The difference equation which describes the system is

$$x(t+1) = x(t) + u(t)$$

Assume $x(0) = 0$.

(a) *(10 points)* What is the probability that $x(n) = 0$, for n an integer greater than zero?

SOLUTION: *This is a random walk markov chain in disguise. You can't get back to state zero in an odd number of steps so if n is odd, Prob$(x(n) = 0) = 0$.*

For n even, the number of forward steps has to be equal the number of backward steps in order to return to zero. There are therefore $\binom{n}{n/2}$ ways to return to zero, each having probability $\alpha^{\frac{n}{2}}(1-\alpha)^{\frac{n}{2}}$. Thus

$$\textit{Prob}(x(n) = 0) = \begin{cases} 0 & n \textit{ odd} \\ \binom{n}{\frac{n}{2}}(\alpha(1-\alpha))^{\frac{n}{2}} & n \textit{ even} \end{cases}$$

(b) *(10 points)* Please derive the Linear minimum mean square estimate for $x(T)$ based on the values of $x(t)$ for $t = 1, 2, ..., T-1$. Assume $x(0) = 0$ and $\alpha = 1/2$.

HINT: Don't just try to turn a crank.

SOLUTION: *This is a markov chain. Where you go depends only on where you are, so estimating $x(T)$ depends only on $x(T-1)$. Given $x(T-1)$ there are only two possibilities for $x(T)$, $x(T-1)+1$ and $x(T-1)-1$, both equally likely. The conditional mean is the MMSE estimate and is easily found to be*

$$E[x(T)|x(T-1)] = \frac{1}{2}(x(T-1)+1) + \frac{1}{2}(x(T-1)-1) = x(T-1)$$

This also happens to be a linear estimate.

It's also guaranteed to be WRONG because we know $x(T) \neq x(T-1)$. However, it's the estimate which minimizes the mean square error.

(c) *(10 points)* Please derive the minimum mean square estimate for $x(T)$ based on the values of $x(t)$ for $t = 1, 2, ..., T-1$. Assume $x(0) = 0$ and $\alpha = 1/2$.

SOLUTION: *We solved this already in the previous part. The MMSE estimate also happens to be a linear estimate. How cool is that?*

4. *(35 points)* In between solving the world's research problems, Rutgera Univera, the world famous Rutgers University graduate student shops for food at Infinite Food Emporium, a supermarket with infinite floorspace. The only thing finite about Infinite Food Emporium is the number of checkout lines, N. The lines are identical, idependent and can be as long as needed. The time a customer spends with a checkout clerk (cashier) is an exponential random variable with mean $1/\mu$.

(a) *(5 points)* Suppose people (including Rutgera) arrive to the checkout lines as a Poisson process with rate λ and then choose one of the N lines randomly. What value of service rate μ guarantees that the service system is stable (that queue lengths do not tend toward infinity)?

SOLUTION: *The random selection gives the arrival rate to each queue as λ/N so we must have $\mu > \lambda/N$ for stability.*

(b) *(5 points)* What is the steady state distribution of each line (number of customers in the line, including the one being served), assuming that the conditions established in the previous part on finite waiting time are satisfied?

SOLUTION: *We have a set of independent $M/M/1/\infty$ queues each with arrival rate λ/N and service rate μ. We established in class (and it's easy to establish by cutting the corresponding markov chain between states) that the steady state distribution is geometric:*

$$f_K(k) = \left(\frac{\lambda}{N\mu}\right)^k \left(1 - \frac{\lambda}{N\mu}\right)$$

(c) *(5 points)* Assume the store has been open a long time by the time Rutgera reaches the checkout line. What is Rutgera's mean waiting time (the time spent in the line before she begins service with the checkout clerk)?

SOLUTION: *Random incidence of the Poisson process suggests that Rutgera arrives at a random point in time which means she finds whatever queue she's chosen at random in steady state. The probability of k customers in the queue ahead of her is $f_K(k)$ as given in the previous part. As the $k+1^{st}$ customer, she'll have to wait for the k folks ahead of her to be serviced. The mean amount of time each spends in service is $1/\mu$, so her mean waiting time is k/μ given there are already k customers present. Therefore, her overall mean waiting time is*

$$E[K]/\mu = \frac{1}{\mu} \frac{\frac{\lambda}{N\mu}}{\left(1 - \frac{\lambda}{N\mu}\right)} = \frac{\frac{\lambda}{N\mu}}{\mu - \frac{\lambda}{N}}$$

By the way, many of you MISapplied Little Theorem $\bar{N} = \lambda \bar{T}$ by thinking for $\bar{N}$ the average number of customers in the steady state that $\bar{T}$ was the average WAITING time WAITING. It is NOT. $\bar{T}$ is the average time IN THE SYSTEM.

(d) *(5 points)* Rutgera is the last customer to reach the checkout lanes. After she joins a queue, she notices that all the lines have exactly C customers. Assume there are no further arrivals to the checkout lanes. What is the probability that Rutgera's lane finishes first?

HINT: Remember that these are exponential servers.

SOLUTION: *Owing to the exponential property of the servers, when Rutgera joins the system and there are C customers in each queue, the states of the N independent queues are identical. The symmetry of the problem dictates that any of the queues is equally likely to be the fastest, so the probability that Rutgera's queue is fastest to empty is $1/N$.*

(e) *(15 points)* Suppose Rutgera, instead of choosing a line at random, chooses the line with the least number of customers in it. What is her mean waiting time?

SOLUTION: *This is a min-of-N problem. The probability that all of a set of N iid random variables X is greater than z is $(1 - F_X(z))^N$, so that the CDF on Z the min-of-N over the N $\{X_i\}$ is $1 - (1 - F_X(z))^N$*

The CDF of the state occupancy distribution is

$$F_K(k) = \sum_{m=0}^{k} f_K(m) = \left(1 - \frac{\lambda}{N\mu}\right) \frac{1 - \left(\frac{\lambda}{N\mu}\right)^{k+1}}{1 - \frac{\lambda}{N\mu}} = 1 - \left(\frac{\lambda}{N\mu}\right)^{k+1}$$

so that our min-of-N CDF is

$$F_Z(z) = 1 - \left(\frac{\lambda}{N\mu}\right)^{N(k+1)}$$

Which is immediately recognized as the CDF of a geometric distribution with parameter $\left(\frac{\lambda}{N\mu}\right)^N$. Thus, the probability distribution on the minimum queue size is

$$f_Z(z) = (1 - \left(\frac{\lambda}{N\mu}\right)^N) \left(\frac{\lambda}{N\mu}\right)^{Nz}$$

with expected value

$$E[Z] = \frac{\left(\frac{\lambda}{N\mu}\right)^N}{1 - \left(\frac{\lambda}{N\mu}\right)^N}$$

so that the mean time Rutgera spends in line before service is $\frac{\left(\frac{\lambda}{N\mu}\right)^N}{\mu - \mu\left(\frac{\lambda}{N\mu}\right)^N}$.