

IMPERIAL COLLEGE LONDON

E4.10

C2.1

SC4

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING  
EXAMINATIONS 2006

MSc and EEE PART IV: MEng and ACGI

Corrected Copy

**PROBABILITY AND STOCHASTIC PROCESSES**

Friday, 12 May 10:00 am

Time allowed: 3:00 hours

**There are SIX questions on this paper.**

**Answer FOUR questions.**

*All questions carry equal marks*

**Any special instructions for invigilators and information for candidates are on page 1.**

Examiners responsible      First Marker(s) :      K.K. Leung

Second Marker(s) :      R.B. Vinter

Special Instructions for Invigilator: **None**

Information for Students: **Complementary Normal Distribution**

$$Q(x) = 1 - \Phi(x) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

$x$	$Q(x)$	$x$	$Q(x)$
0	5.00e-01	2.7	3.47e-03
0.1	4.60e-01	2.8	2.56e-03
0.2	4.21e-01	2.9	1.87e-03
0.3	3.82e-01	3.0	1.35e-03
0.4	3.45e-01	3.1	9.68e-04
0.5	3.09e-01	3.2	6.87e-04
0.6	2.74e-01	3.3	4.83e-04
0.7	2.42e-01	3.4	3.37e-04
0.8	2.12e-01	3.5	2.33e-04
0.9	1.84e-01	3.6	1.59e-04
1.0	1.59e-01	3.7	1.08e-04
1.1	1.36e-01	3.8	7.24e-05
1.2	1.15e-01	3.9	4.81e-05
1.3	9.68e-02	4.0	3.17e-05
1.3	8.08e-02	4.5	3.40e-06
1.5	6.68e-02	5.0	2.87e-07
1.6	5.48e-02	5.5	1.90e-08
1.7	4.46e-02	6.0	9.87e-10
1.8	3.59e-02	6.5	4.02e-11
1.9	2.87e-02	7.0	1.28e-12
2.0	2.28e-02	7.5	3.19e-14
2.1	1.79e-02	8.0	6.22e-16
2.2	1.39e-02	8.5	9.48e-19
2.3	1.07e-02	9.0	1.13e-19
2.4	8.20e-03	9.5	1.05e-21
2.5	6.21e-03	10.0	7.62e-24
2.6	4.66e-03		

1. a. At a party,  $N$  men throw their hats in the center of a room. The hats are mixed up, and each man takes turn to randomly select one and then returns it to the center. Assume that all hats are different. Find the expected number of men that select their own hats.

[5]

b. Consider  $K$  independent, scalar random variables  $X_1, X_2, \dots, X_K$ , each of which is exponentially distributed with parameter  $\lambda$ . That is,  $f_X(x) = \lambda e^{-\lambda x}$  if  $x \geq 0$ , and 0 otherwise. Let  $Y$  be the minimum of  $X_1, X_2, \dots$ , and  $X_K$ . Find the probability density function (pdf)  $f_Y(y)$  for  $Y$ .

[8]

c. Given two independent nonnegative random variables  $X$  and  $Y$ , we define a new variable  $Z \equiv X + Y$ . Let the Laplace transforms of the probability density functions (pdf's) for  $X, Y$  and  $Z$  be denoted by  $X^*(s), Y^*(s)$  and  $Z^*(s)$ , respectively. Show that

$$Z^*(s) = X^*(s)Y^*(s).$$

[12]

2. a. A positive integer valued random variable  $X$  has a geometric distribution. That is, for  $n=1, 2, 3, \dots$

$$p_n = P[X = n] = p(1-p)^{n-1}$$

where  $0 < p < 1$ . Determine the probability generation function (i.e., the z-transform) for  $X$ . Hence, or otherwise, obtain the mean and variance of  $X$ .

[10]

- b. Show that the geometric distribution given in part a is equivalent to the conditional-probability property defined below, using the following approach:

- i. Prove that if the random variable  $X$  has the geometric distribution, then  $X$  has the following conditional probability property:

$$P[X > m + n | X > m] = P[X > n]$$

for any two positive integers  $m$  and  $n$ .

[5]

- ii. Prove that if a positive integer valued random value  $X$  satisfies

$$P[X > m + n | X > m] = P[X > n]$$

for any two positive integers  $m$  and  $n$ , then  $X$  must have the geometric distribution. [10]

(Hint: Express the conditional probability in terms of  $a_k = P[X > k]$ .)

3. a. If random variables  $X$  and  $Y$  are identically distributed, not necessarily independent, show that

$$\text{cov}(X + Y, X - Y) = 0. \quad [6]$$

- b. Now consider random variables  $X$ ,  $Y$  and  $Z$ . The conditional covariance of  $X$  and  $Y$ , given  $Z$ , is defined by

$$\text{cov}(X, Y | Z) = E[(X - E[X | Z])(Y - E[Y | Z]) | Z].$$

- i. Show that

$$\text{cov}(X, Y | Z) = E[XY | Z] - E[X | Z]E[Y | Z]. \quad [5]$$

- ii. Show that

$$\text{cov}(X, Y) = E[\text{cov}(X, Y | Z)] + \text{cov}(E[X | Z], E[Y | Z]). \quad [9]$$

(Hint: Take the expectation of both sides of the result in part i.)

- iii. Set  $X=Y$  in part ii and obtain a formula for  $\text{var}(X)$  in terms of conditional expectations. [5]

4. a. Suppose that orders  $X_1, X_2, \dots, X_n$  at a restaurant are independent, identically distributed (iid) random variables with mean  $\mu = \$8$  and standard deviation  $\sigma = \$2$ . Let

$$S_n \equiv \sum_{i=1}^n X_i \text{ and } Z_n \equiv \frac{S_n - n\mu}{\sigma\sqrt{n}}.$$

Using the Central Limit Theorem for  $Z_n$  and the complementary normal distribution given at the beginning of this examination paper,

- i. Estimate the probability that the first 100 customers (orders) spend a total of more than \$840. [5]

- ii. Estimate the probability that the first 100 customers spend a total between \$780 and \$820. [6]

- b. Two random variables  $X$  and  $Y$  are said to have a bivariate normal distribution if their joint probability density function (pdf) is given by

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\}$$

- i. Show that the marginal pdf's for both  $X$  and  $Y$  are normal density functions with respective parameters  $\mu_X, \sigma_X, \mu_Y$  and  $\sigma_Y$ . [6]
- ii. Show that the conditional pdf for  $X$ , given that  $Y=y$ , is the normal pdf with parameters

$$\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y) \text{ and } \sigma_X^2(1-\rho^2). [6]$$

- iii. Show that  $X$  and  $Y$  are independent when  $\rho = 0$ . [2]

5. **Speech Predictor:** Let  $X_1, X_2, X_3, \dots$  be a sequence of samples of a speech voltage waveform. Suppose that  $\{X_k\}$  represents a stationary, second order, stochastic process where  $E[X_k] = 0$  and  $E[X_k^2] = \sigma^2 > 0$  for all  $k = 1, 2, 3, \dots$ . The samples  $X_{k-2}$  and  $X_{k-1}$  are used to predict  $X_k$  with the least mean square error. That is, for  $k = 3, 4, 5, \dots$

$$\hat{X}_k = aX_{k-1} + bX_{k-2} \quad (*)$$

is the least squares predictor of  $X_k$  where  $a$  and  $b$  are constants.

- a. Find  $a$  and  $b$  in terms of the variance  $\sigma^2$  and covariances of  $\{X_k\}$ . [11]
- b. Given the stationary property of the process, the covariances of  $\{X_k\}$  depend on the “distance” between the time indices, but not on the specific index values. Assume that for all  $i$  and  $j$

$$\text{COV}(X_i, X_j) = \rho_{|i-j|} \sigma^2$$

where  $\rho_k > 0$  and  $\rho_0 = 1$ . (For example,  $\text{COV}(X_i, X_{i+2}) = \rho_2 \sigma^2$ .) Find  $a$  and  $b$  in terms of  $\rho_1$  and  $\rho_2$ . [4]

- c. In a practical implementation, it turns out that due to insufficient processing time, only an inaccurate, noisy measurement  $\tilde{X}_{k-1}$  can be obtained for predicting  $X_k$  according to (\*). For each  $k=2, 3, \dots$ ,  $\tilde{X}_{k-1}$  is given by

$$\tilde{X}_{k-1} = X_{k-1} + N_{k-1}$$

where  $\{N_k\}$  is an independent sequence of zero mean, uncorrelated random variables with variance  $\sigma_N^2$ . On the other hand, sufficient processing time is available to obtain an accurate value of  $X_{k-2}$  for purposes of prediction in (\*). Find the optimal  $a$  and  $b$  in terms of  $\rho_1, \rho_2, \sigma^2$  and  $\sigma_N^2$  in the presence of the measurement noise.

(Hint: Include the effects of the measurement noise in the results in part a.) [10]

6. a. Consider a two-state continuous-time Markov chain that spends an exponential time with rate  $\lambda$  in state 0 before going to state 1, where it spends an exponential time with rate  $\mu$  before returning to state 0. Let  $\pi_0(t)$  and  $\pi_1(t)$  be the state probability at time  $t$  for state 0 and 1, respectively.

i. Give the differential-difference equations for  $\pi_0(t)$  and  $\pi_1(t)$ . [5]

(Hint:  $\frac{d\pi_j(t)}{dt} = q_{jj}\pi_j(t) + \sum_{k \neq j} q_{kj}\pi_k(t)$  where  $q_{kj}$  is the state transition rate from state  $k$  to  $j$  for  $k \neq j$ , and  $-q_{jj}$  is the state transition rate at which the process departs from state  $j$  when it is in that state.)

ii. Let  $\pi_0(0) = 1$ . Solve these equations for  $\pi_0(t)$  and  $\pi_1(t)$ . [7]

b. Let  $X_1, X_2, \dots$  be independent, identical distributed random variables such that

$P[X_i = j] = \alpha_j, j \geq 0$ . We say that a record occurs at time  $n$

if  $X_n > \max(X_1, \dots, X_{n-1})$ , where  $X_0 = -\infty$ , and if a record does occur at time  $n$ , let

$X_n$  be called the record value. Let  $R_i$  denote the  $i$ th record value.

i. Argue that  $\{R_i, i \geq 1\}$  is a Markov chain. [3]

ii. Compute its transition probabilities. [10]



# Model Answers Probability & Stochastic

Process 2006 E4.10/

C.2.1/SC4

#1

a. Let  $X_i = \begin{cases} 1 & \text{if } i^{\text{th}} \text{ man selects his own hat} \\ 0 & \text{otherwise} \end{cases}$

$$X = \sum_{i=1}^N X_i \Rightarrow E[X] = \sum_{i=1}^N E[X_i]$$

$$\text{For each } i, E[X_i] = \frac{1}{N} \cdot 1 + \frac{N-1}{N} \cdot 0 = \frac{1}{N}$$

$$\Rightarrow E[X] = \sum_{i=1}^N E[X_i] = N \cdot \frac{1}{N} = 1 \quad (\text{The expected number of men selecting their own hats.})$$

b.  $Y = \min(X_1, X_2, \dots, X_k)$

$$P[Y > y] = P[\min(X_1, X_2, \dots, X_k) > y]$$

$$= P[X_1 > y, X_2 > y, \dots, X_k > y]$$

$$= P[X_1 > y] P[X_2 > y] \dots P[X_k > y] \quad \because \text{All } X_i \text{'s are independent.}$$

$$\Rightarrow P[Y > y] = e^{-\lambda y} \cdot e^{-\lambda y} \dots e^{-\lambda y} = e^{-k\lambda y}$$

$$\Rightarrow P[Y \leq y] = 1 - e^{-k\lambda y} \Rightarrow f_Y(y) = k\lambda e^{-k\lambda y}$$

c. Let  $f_X(x)$ ,  $f_Y(y)$  and  $f_Z(z)$  be the pdf for  $X$ ,  $Y$  and  $Z$ , respectively. Similarly, let  $F_X(x)$ ,  $F_Y(y)$  and  $F_Z(z)$  be the CDF for  $X$ ,  $Y$ , and  $Z$ , respectively.

By definition,  $F_Z(z) = P(Z \leq z) = P(X+Y \leq z)$

$$\Rightarrow F_Z(z) = \int_{y=0}^{\infty} P(X+Y \leq z | Y=y) f_Y(y) dy$$

$$= \int_{y=0}^{\infty} P(X \leq z-y | Y=y) f_Y(y) dy$$

$$\Rightarrow F_Z(z) = \int_{y=0}^{\infty} F_X(z-y) f_Y(y) dy$$

$$\Rightarrow \frac{dF_z(z)}{dz} = \int_{y=0}^{\infty} \frac{dF_x(z-y)}{dz} f_Y(y) dy$$

$$\Rightarrow f_z(z) = \int_{y=0}^{\infty} f_x(z-y) f_Y(y) dy \quad \text{--- ①}$$

To show  $Z^*(s) = X^*(s) \cdot Y^*(s)$ , we take the Laplace transform on both sides of ①:

$$\int_{z=0}^{\infty} f_z(z) e^{-sz} dz = \int_{z=0}^{\infty} \int_{y=0}^{\infty} f_x(z-y) f_Y(y) dy \cdot e^{-sz} dz$$

$$\Rightarrow Z^*(s) = \int_{z=0}^{\infty} \int_{y=0}^z f_x(z-y) f_Y(y) e^{-sz} dy dz$$

$$\Rightarrow Z^*(s) = \int_{y=0}^{\infty} \int_{z=y}^{\infty} f_x(z-y) e^{-s(z-y)} dz f_Y(y) e^{-sy} dy$$

$$\Rightarrow Z^*(s) = \int_{y=0}^{\infty} \underbrace{X^*(s)}_{\leftarrow} f_Y(y) e^{-sy} dy$$

$$\Rightarrow Z^*(s) = X^*(s) \cdot Y^*(s) \quad \text{Q.E.D.}$$

#2

a. Let the z-transform be  $P(z)$

$$P(z) = \sum_{n=1}^{\infty} p_n z^n = \sum_{n=1}^{\infty} p(1-p)^{n-1} z^n$$

$$\Rightarrow P(z) = pz \sum_{n=1}^{\infty} [(1-p)z]^{n-1} = \frac{pz}{1-(1-p)z}$$

Let  $q \equiv 1-p$ . Then  $P(z) = \frac{pz}{1-qz}$ .

$$\frac{dP(z)}{dz} = \frac{d}{dz} [pz(1-qz)^{-1}]$$

$$\Rightarrow \frac{dP(z)}{dz} = p(1-qz)^{-1} + pz(-1)(1-qz)^{-2}(-q)$$

$$\Rightarrow \frac{dP(z)}{dz} = p(1-qz)^{-1} + pqz(1-qz)^{-2} \quad \text{--- ①}$$

$$\Rightarrow \frac{d^2P(z)}{dz^2} = p(-1)(1-qz)^{-2}(-q) + pq(1-qz)^{-2} + pqz(-2)(1-qz)^{-3}(-q)$$

$$\Rightarrow \frac{d^2P(z)}{dz^2} = pq(1-qz)^{-2} + pq(1-qz)^{-2} + 2pq^2z(1-qz)^{-3}$$

$$\Rightarrow \frac{d^2P(z)}{dz^2} = 2pq(1-qz)^{-2} + 2pq^2z(1-qz)^{-3} \quad \text{--- ②}$$

Since  $E[X] = \frac{dP(z)}{dz} \Big|_{z=1}$ , from ① we have

$$E(X) = p(1-q)^{-1} + pq(1-q)^{-2}$$

$$\Rightarrow E(X) = p \cdot p^{-1} + pq p^{-2}$$

$$\Rightarrow E(X) = 1 + \frac{q}{p} = 1 + \frac{1-p}{p} = \frac{1}{p}$$

p.3

$$\left. \frac{d^2 P(z)}{dz^2} \right|_{z=1} = E[X(X-1)]$$

From ②, we have

$$E[X(X-1)] = 2pq(1-q)^{-2} + 2pq^2(1-q)^{-3}$$

$$\Rightarrow E[X(X-1)] = 2pq p^{-2} + 2pq^2 p^{-3}$$

$$\Rightarrow E[X(X-1)] = 2q/p + 2q^2/p^2 = 2 \frac{q}{p} \left(1 + \frac{q}{p}\right)$$

$$\Rightarrow E[X(X-1)] = 2 \cdot \frac{1-p}{p} \cdot \left(1 + \frac{1-p}{p}\right) = 2 \left(\frac{1-p}{p^2}\right)$$

Since  $E(X) = \frac{1}{p}$ , subs. this into the above yields

$$E(X^2) = 2 \cdot \frac{1-p}{p^2} + \frac{1}{p} = \frac{2}{p^2} - \frac{1}{p}$$

Variance of  $X$ :

$$\sigma_X^2 = E(X^2) - [E(X)]^2$$

$$= \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2}$$

$$\Rightarrow \sigma_X^2 = \frac{1-p}{p^2}$$

b. i) Consider  $m, n > 0$

$$\begin{aligned} P(X > m+n \mid X > m) &= \frac{P(X > m+n \text{ and } X > m)}{P(X > m)} \\ &= \frac{P(X > m+n)}{P(X > m)} \quad \because n > 0 \end{aligned} \quad \text{--- (3)}$$

Define  $a_k \triangleq P(X > k)$ .

Given that  $X$  has a geometric distribution,

$$a_k = P(X > k) = \sum_{n=k+1}^{\infty} p(1-p)^{n-1} = p \cdot \frac{(1-p)^k}{1-(1-p)}$$

$$\Rightarrow a_k = (1-p)^k$$

Subs.  $a_k$  into (3), we have

$$P(X > m+n \mid X > m) = \frac{P(X > m+n)}{P(X > m)} = \frac{a_{m+n}}{a_m} \quad \text{--- (4)}$$

$$\Rightarrow P(X > m+n \mid X > m) = \frac{(1-p)^{m+n}}{(1-p)^m} = (1-p)^n = a_n$$

$$\Rightarrow P(X > m+n \mid X > m) = P(X > n) \quad \text{Q.E.D.}$$

ii) Since the derivation for (4) does not involve the distribution for  $X$ , the equation remains valid in general.

If  $P(X > m+n \mid X > m) = P(X > n)$ , then we now have

$$P(X > m+n \mid X > m) = \frac{a_{m+n}}{a_m} = P(X > n) = a_n$$

That is,  $a_{m+n} = a_m a_n$  for  $m, n$  positive integers

Let  $n=1$ ,  $a_{m+1} = a_m a_1$

Consider  $m=1, 2, 3, \dots$ . We then have

$$a_{m+1} = a_1^{m+1} \quad \text{--- (5)}$$

by definition,  $a_1 = P(X > 1) = 1 - p(X=1) \triangleq 1 - p$

Therefore, from (5),  $a_m = a_1^m = (1-p)^m$  for  $m=1, 2, 3, \dots$

Since  $P(X=n) = P(X \geq n) - P(X > n)$

$$= a_{n-1} - a_n$$

$$= (1-p)^{n-1} - (1-p)^n$$

$$\Rightarrow P(X=n) = p(1-p)^{n-1} \quad \text{for } n=1, 2, 3, \dots$$

A geometric distribution ! Q.E.D.

#3

$$\begin{aligned} a. \quad \text{Cov}(X+Y, X-Y) &= \text{Cov}(X, X) + \text{Cov}(X, -Y) + \text{Cov}(Y, X) \\ &\quad + \text{Cov}(Y, -Y) \\ &= \text{Var}(X) - \text{Cov}(X, Y) + \text{Cov}(Y, X) - \text{Var}(Y) \\ &= \text{Var}(X) - \text{Var}(Y) \\ &= 0 \end{aligned}$$

$\because X$  and  $Y$  are identically distributed.

$$\begin{aligned} b. \quad i) \quad \text{Cov}(X, Y|Z) &= E\{XY - E[X|Z]Y - XE[Y|Z] + E[X|Z]E[Y|Z] | Z\} \\ &= E[XY|Z] - E[X|Z]E[Y|Z] - E[X|Z]E[Y|Z] \\ &\quad + E[X|Z] \cdot E[Y|Z] \quad \because E[\cdot] \text{ is a constant} \\ &= E[XY|Z] - E[X|Z]E[Y|Z] \end{aligned}$$

ii) Take the expectation on both sides of result in part i:

$$\begin{aligned} E[\text{Cov}(X, Y|Z)] &= E\{E[XY|Z] - E[X|Z]E[Y|Z]\} \\ \Rightarrow E[\text{Cov}(X, Y|Z)] &= E[XY] - E\{E[X|Z]E[Y|Z]\} \quad \text{--- (1)} \end{aligned}$$

Consider  $\text{Cov}(E[X|Z], E[Y|Z])$ .

$$\begin{aligned} \text{Cov}(E[X|Z], E[Y|Z]) &= E\{E[X|Z]E[Y|Z]\} \\ &\quad - E\{E[X|Z]\} \cdot E\{E[Y|Z]\} \\ \Rightarrow \text{Cov}(E[X|Z], E[Y|Z]) &= E\{E[X|Z]E[Y|Z]\} - E[X]E[Y] \end{aligned}$$

--- (2)

Combine ① and ② :

$$\begin{aligned} E[\text{cov}(X, Y|Z)] + \text{cov}(E[X|Z], E[Y|Z]) \\ = E[XY] - E(X)E(Y) \\ = \text{cov}(X, Y) \end{aligned} \quad \text{Q.E.D.}$$

iii)  $\text{cov}(X, Y|Z) = \text{var}(X|Z)$  when  $Y=X$ .

The LHS of result in part ii:  
 $\text{cov}(X, Y) = \text{var}(X)$

The RHS :  $E[\text{cov}(X, X|Z)] + \text{cov}(E[X|Z], E[X|Z])$   
 $= E[\text{var}(X|Z)] + \text{var}(E[X|Z])$   
 $\Rightarrow \text{var}(X) = E[\text{var}(X|Z)] + \text{var}(E[X|Z])$



#4

Q.  $n=100$   $S_{100} = X_1 + X_2 + \dots + X_{100}$

$$E[S_{100}] = n\mu = 800 \quad \sigma_{S_{100}}^2 = n\sigma^2 = 400.$$

The normalized form of  $S_{100}$  is

$$Z_{100} = \frac{S_{100} - 800}{20}$$

i)  $P(S_{100} > 840) = P\left(Z_{100} > \frac{840 - 800}{20}\right)$

$$= P(Z_{100} > 2)$$

$$\doteq Q(2) = 2.28 \times 10^{-2} \quad (\text{from the CDF table})$$

ii)  $P(780 \leq S_{100} \leq 820)$

$$= P(-1 \leq Z_{100} \leq 1)$$

$$\doteq 1 - 2Q(1)$$

$$= 1 - 2 \times 1.59 \times 10^{-1}$$

$$= 0.682$$

$$b. i) f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

$$= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left\{ \frac{-1}{2(1-\rho^2)} \left[ \left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho \frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} \right] \right\} dy$$

$$\Rightarrow f_x(x) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{ \frac{-1}{2(1-\rho^2)} \left(\frac{x-\mu_x}{\sigma_x}\right)^2 \right\} \int_{-\infty}^{\infty} \exp\left\{ \frac{-1}{2(1-\rho^2)} \left[ \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho \frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \rho^2 \left(\frac{x-\mu_x}{\sigma_x}\right)^2 - \rho^2 \left(\frac{x-\mu_x}{\sigma_x}\right)^2 \right] \right\} dy$$

$$\Rightarrow f_x(x) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{ \left(\frac{x-\mu_x}{\sigma_x}\right)^2 \left[ \frac{-1}{2(1-\rho^2)} + \frac{\rho^2}{2(1-\rho^2)} \right] \right\} \int_{-\infty}^{\infty} \exp\left\{ \frac{-1}{2(1-\rho^2)} \left[ \frac{y-\mu_y}{\sigma_y} - \rho \frac{x-\mu_x}{\sigma_x} \right]^2 \right\} dy$$

$$\Rightarrow f_x(x) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{ \frac{-1}{2} \left(\frac{x-\mu_x}{\sigma_x}\right)^2 \right\}$$

$$\int_{-\infty}^{\infty} \exp\left\{ \frac{-1}{2} \left( \frac{y-\mu_y - \rho\sigma_y(x-\mu_x)/\sigma_x}{\sigma_y\sqrt{1-\rho^2}} \right)^2 \right\} dy$$

$$\Rightarrow f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left\{ \frac{-1}{2} \left(\frac{x-\mu_x}{\sigma_x}\right)^2 \right\}$$

$$\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}\sigma_y} \int_{-\infty}^{\infty} \exp\left\{ \frac{-1}{2} \left( \frac{y-\mu_y - \rho\sigma_y(x-\mu_x)/\sigma_x}{\sigma_y\sqrt{1-\rho^2}} \right)^2 \right\} dy$$

$$\Rightarrow f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left\{ \frac{-1}{2} \left(\frac{x-\mu_x}{\sigma_x}\right)^2 \right\} = 1 \quad \begin{array}{l} \because \text{Integrating} \\ \text{normal distribution} \\ \text{density} = 1 \end{array} \quad p.10$$

Similarly,

$$f_Y(y) = \frac{1}{\sqrt{2\pi} \sigma_Y} \cdot \exp \left\{ \frac{-1}{2} \left( \frac{y - \mu_Y}{\sigma_Y} \right)^2 \right\}$$

$$\text{(ii) } f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

$$= \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[ \left( \frac{x - \mu_X}{\sigma_X} \right)^2 + \left( \frac{y - \mu_Y}{\sigma_Y} \right)^2 - 2\rho \frac{(x - \mu_X)(y - \mu_Y)}{\sigma_X \sigma_Y} \right] \right\} \cdot \sqrt{2\pi} \sigma_Y \exp \left\{ \frac{1}{2} \left( \frac{y - \mu_Y}{\sigma_Y} \right)^2 \right\}$$

$$\Rightarrow f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi} \sigma_X \sqrt{1-\rho^2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[ \left( \frac{x - \mu_X}{\sigma_X} \right)^2 - 2\rho \frac{(x - \mu_X)(y - \mu_Y)}{\sigma_X \sigma_Y} + \rho^2 \left( \frac{y - \mu_Y}{\sigma_Y} \right)^2 - \rho^2 \left( \frac{y - \mu_Y}{\sigma_Y} \right)^2 + \left( \frac{y - \mu_Y}{\sigma_Y} \right)^2 \right] \right\} \cdot \exp \left\{ \frac{1}{2} \left( \frac{y - \mu_Y}{\sigma_Y} \right)^2 \right\}$$

$$\Rightarrow f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi} \sigma_X \sqrt{1-\rho^2}} \cdot \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[ \left( \frac{x - \mu_X}{\sigma_X} - \rho \cdot \frac{y - \mu_Y}{\sigma_Y} \right)^2 \right] + \underbrace{(-\rho^2 + 1 - (1-\rho^2))}_{0} \left( \frac{y - \mu_Y}{\sigma_Y} \right)^2 \right\}$$

$$\Rightarrow f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi} \sigma_X \sqrt{1-\rho^2}} \cdot \exp \left\{ \frac{-1}{2} \left( \frac{x - \mu_X - \rho \frac{\sigma_X}{\sigma_Y} (y - \mu_Y)}{\sigma_X \sqrt{1-\rho^2}} \right)^2 \right\}$$

a normal density with

$$\text{mean} = \mu_X + \rho \frac{\sigma_X}{\sigma_Y} (y - \mu_Y)$$

$$\text{variance} = \frac{\sigma_X^2}{1-\rho^2} \quad \text{p.11}$$

iii) When  $\rho=0$

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi\sigma_x\sigma_y} \cdot \exp\left\{-\frac{1}{2} \left[ \left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 \right]\right\} \\ &= \underbrace{\frac{1}{\sqrt{2\pi}\sigma_x} \exp\left\{-\frac{1}{2} \left(\frac{x-\mu_x}{\sigma_x}\right)^2\right\}}_{f_x(x)} \cdot \underbrace{\frac{1}{\sqrt{2\pi}\sigma_y} \exp\left\{-\frac{1}{2} \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}}_{f_y(y)} \end{aligned}$$

$$\Rightarrow f(x, y) = f_x(x) f_y(y)$$

Therefore,  $X$  and  $Y$  are independent.

#5

a. Minimize  $E[(X_k - aX_{k-1} - bX_{k-2})^2]$   
 a, b

Take the derivative of the mean square error w.r.t.  $a$  and set the result to zero. We get

$$E[(X_k - aX_{k-1} - bX_{k-2})X_{k-1}] = 0$$

$$\Rightarrow E[X_k X_{k-1}] - aE[X_{k-1} X_{k-1}] - bE[X_{k-2} X_{k-1}] = 0$$

Similarly, differentiation w.r.t.  $b$  yields

$$E[(X_k - aX_{k-1} - bX_{k-2})X_{k-2}] = 0$$

$$\Rightarrow E[X_k X_{k-2}] - aE[X_{k-1} X_{k-2}] - bE[X_{k-2}^2] = 0 \quad \text{--- (2)}$$

Solving (1) and (2) yields

$$a = \frac{\text{var}(X_{k-2}) \text{cov}(X_k, X_{k-1}) - \text{cov}(X_k, X_{k-2}) \text{cov}(X_{k-2}, X_{k-1})}{\text{var}(X_{k-1}) \text{var}(X_{k-2}) - \text{cov}(X_{k-1}, X_{k-2})^2}$$

$$b = \frac{\text{var}(X_{k-1}) \text{cov}(X_k, X_{k-2}) - \text{cov}(X_k, X_{k-1}) \text{cov}(X_{k-1}, X_{k-2})}{\text{var}(X_{k-2}) \cdot \text{var}(X_{k-1}) - \text{cov}(X_{k-2}, X_{k-1})^2}$$

As  $\text{var}(X_k) = \sigma^2$  for  $\forall k$ ,

$$a = \frac{\sigma^2 \text{cov}(X_k, X_{k-1}) - \text{cov}(X_k, X_{k-2}) \text{cov}(X_{k-1}, X_{k-2})}{\sigma^4 - \text{cov}(X_{k-1}, X_{k-2})^2}$$

$$b = \frac{\sigma^2 \text{cov}(X_k, X_{k-2}) - \text{cov}(X_k, X_{k-1}) \text{cov}(X_{k-1}, X_{k-2})}{\sigma^4 - \text{cov}(X_{k-1}, X_{k-2})^2}$$

b. Subs.  $cov(X_i, X_j) = \rho_{|i-j|} \sigma^2$  into the expressions for a and b in part b:

$$a = \frac{\sigma^2 \rho_1 \sigma^2 - \rho_2 \sigma^2 \rho_1 \sigma^2}{\sigma^4 - \rho_1^2 \sigma^4} = \frac{\rho_1 (1 - \rho_2)}{1 - \rho_1^2}$$

$$b = \frac{\sigma^2 \rho_2 \sigma^2 - \rho_1 \sigma^2 \rho_1 \sigma^2}{\sigma^4 - \rho_1^2 \sigma^4} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

c. minimize  $E \left[ (X_k - a \tilde{X}_{k-1} - b X_{k-2})^2 \right]$   
 a, b

Using the results in part a:

$$a = \frac{\text{var}(X_{k-2}) \text{cov}(X_k, \tilde{X}_{k-1}) - \text{cov}(X_k, X_{k-2}) \text{cov}(\tilde{X}_{k-1}, X_{k-2})}{\text{var}(\tilde{X}_{k-1}) \text{var}(X_{k-2}) - \text{cov}(\tilde{X}_{k-1}, X_{k-2})^2}$$

Let us consider  $\text{var}(\tilde{X}_{k-1})$ .

$$\text{var}(\tilde{X}_{k-1}) = \text{var}(X_{k-1} + N_{k-1})$$

$$= \text{var}(X_{k-1}) + \text{var}(N_{k-1})$$

∵  $X_{k-1}$  &  $N_{k-1}$  are independent

$$\Rightarrow \text{var}(\tilde{X}_{k-1}) = \sigma^2 + \sigma_N^2$$

Consider  $\text{cov}(X_k, \tilde{X}_{k-1})$ :

$$\text{cov}(X_k, \tilde{X}_{k-1}) = \text{cov}(X_k, X_{k-1} + N_{k-1})$$

$$= \text{cov}(X_k, X_{k-1}) + \underbrace{\text{cov}(X_k, N_{k-1})}_0$$

$$= \text{cov}(X_k, X_{k-1})$$

Similarly,  $\text{cov}(X_{k-2}, \tilde{X}_{k-1}) = \text{cov}(X_{k-2}, X_{k-1})$

Subs. these into the expressions for a and b :

$$a = \frac{\sigma^2 \rho_1 \sigma^2 - \rho_2 \sigma^2 \rho_1 \sigma^2}{(\sigma^2 + \sigma_N^2) \sigma^2 - \rho_1^2 \sigma^4}$$

$$\Rightarrow a = \frac{\sigma^2 \rho_1 (1 - \rho_2)}{\sigma^2 + \sigma_N^2 - \rho_1^2 \sigma^2}$$

$$b = \frac{(\sigma^2 + \sigma_N^2) \rho_2 \sigma^2 - \rho_1 \sigma^2 \rho_1 \sigma^2}{\sigma^2 (\sigma^2 + \sigma_N^2) - \rho_1^2 \sigma^4}$$

$$\Rightarrow b = \frac{(\sigma^2 + \sigma_N^2) \rho_2 - \rho_1^2 \sigma^2}{\sigma^2 + \sigma_N^2 - \rho_1^2 \sigma^2}$$

#6

a. i) The differential-difference equations for the Markov chain are:

$$\frac{d\pi_0(t)}{dt} = -\lambda\pi_0(t) + \mu\pi_1(t) \quad \text{--- (1)}$$

$$\frac{d\pi_1(t)}{dt} = \lambda\pi_0(t) - \mu\pi_1(t) \quad \text{--- (2)}$$

ii) Since  $\pi_0(t) + \pi_1(t) = 1$  for  $\forall t$ , subs. this into (1) yields

$$\frac{d\pi_0(t)}{dt} = -\lambda\pi_0(t) + \mu(1 - \pi_0(t))$$

$$\Rightarrow \frac{d\pi_0(t)}{dt} + (\lambda + \mu)\pi_0(t) = \mu$$

$$\Rightarrow e^{(\lambda + \mu)t} \left[ \frac{d\pi_0(t)}{dt} + (\lambda + \mu)\pi_0(t) \right] = \mu e^{(\lambda + \mu)t}$$

$$\Rightarrow \frac{d}{dt} \left[ e^{(\lambda + \mu)t} \pi_0(t) \right] = \mu e^{(\lambda + \mu)t}$$

$$\Rightarrow e^{(\lambda + \mu)t} \pi_0(t) = \frac{\mu}{\lambda + \mu} e^{(\lambda + \mu)t} + c$$

$$\Rightarrow \pi_0(t) = \frac{\mu}{\lambda + \mu} + c e^{-(\lambda + \mu)t}$$

where  $c$  is a constant.

$$\text{Given } \pi_0(0) = 1, \quad \pi_0(0) = \frac{\mu}{\lambda + \mu} + c = 1$$

$$\Rightarrow c = \frac{\lambda}{\lambda + \mu}$$

$$\text{Thus, } \pi_0(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t}$$

$$\pi_1(t) = 1 - \pi_0(t) = \frac{\lambda}{\lambda + \mu} \left[ 1 - e^{-(\lambda + \mu)t} \right]$$



b. i) Consider that the current record value is  $R_i$

for some  $i \geq 1$ . The value of the next record value  $R_{i+1}$  must be larger than  $R_i$ .

So given  $R_i$ , the value of  $R_{i+1}$  is completely characterized by the  $X_i$ 's, which are i.i.d.

Therefore,  $\{R_i, i \geq 1\}$  is a Markov chain.

ii) Without loss of generality, let  $X_k$  be the element in the sequence  $\{X_j, j \geq 1\}$  that corresponds to  $R_{i+1}$ . We have

$$\begin{aligned} P(R_{i+1} = n / R_i = m) &= P(X_k = n / X_k > m) \\ &= \frac{\alpha_n}{\sum_{j=m+1}^{\infty} \alpha_j} \end{aligned}$$

where  $n > m$ . So, considering the cases with  $n \leq m$  leads us to

$$P(R_{i+1} = n / R_i = m) = \begin{cases} 0 & \text{for } n \leq m \\ \frac{\alpha_n}{\sum_{j=m+1}^{\infty} \alpha_j} & \text{for } n > m. \end{cases}$$

all  
for  $m = 0, 1, 2, \dots$

(Last page)