Probability and Stochastic Processes

A Friendly Introduction for Electrical and Computer Engineers

Second Edition

Quiz Solutions

Roy D. Yates and David J. Goodman

May 22, 2004

- The MATLAB section quizzes at the end of each chapter use programs available for download as the archive `matcode.zip`. This archive has programs of general purpose programs for solving probability problems as well as specific `.m` files associated with examples or quizzes in the text. Also available is a manual `probmatlab.pdf` describing the general purpose `.m` files in `matcode.zip`.

- We have made a substantial effort to check the solution to every quiz. Nevertheless, there is a nonzero probability (in fact, a probability close to unity) that errors will be found. If you find errors or have suggestions or comments, please send email to

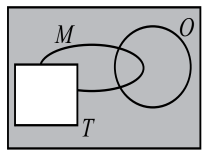
  ryates@winlab.rutgers.edu.

  When errors are found, corrected solutions will be posted at the website.

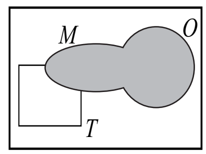
Quiz Solutions – Chapter 1

Quiz 1.1

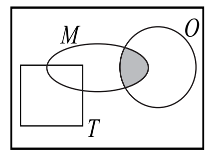
In the Venn diagrams for parts (a)-(g) below, the shaded area represents the indicated set.



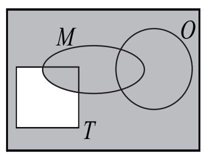
(1) $R = T^c$



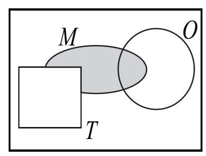
(2) $M \cup O$



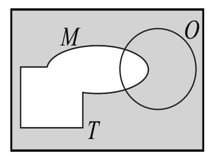
(3) $M \cap O$



(4) $R \cup M$



(4) $R \cap M$



(6) $T^c - M$

Quiz 1.2

(1) $A_1 = \{vvv, vvd, vdv, vdd\}$

(2) $B_1 = \{dvv, dvd, ddv, ddd\}$

(3) $A_2 = \{vvv, vvd, dvv, dvd\}$

(4) $B_2 = \{vdv, vdd, ddv, ddd\}$

(5) $A_3 = \{vvv, ddd\}$

(6) $B_3 = \{vdv, dvd\}$

(7) $A_4 = \{vvv, vvd, vdv, dvv, vdd, dvd, ddv\}$

(8) $B_4 = \{ddd, ddv, dvd, vdd\}$

Recall that A_i and B_i are collectively exhaustive if $A_i \cup B_i = S$. Also, A_i and B_i are mutually exclusive if $A_i \cap B_i = \phi$. Since we have written down each pair A_i and B_i above, we can simply check for these properties.

The pair A_1 and B_1 are mutually exclusive and collectively exhaustive. The pair A_2 and B_2 are mutually exclusive and collectively exhaustive. The pair A_3 and B_3 are mutually exclusive but *not* collectively exhaustive. The pair A_4 and B_4 are not mutually exclusive since dvd belongs to A_4 and B_4. However, A_4 and B_4 are collectively exhaustive.

Quiz 1.3

There are exactly 50 equally likely outcomes: s_{51} through s_{100}. Each of these outcomes has probability 0.02.

(1) $P[\{s_{79}\}] = 0.02$

(2) $P[\{s_{100}\}] = 0.02$

(3) $P[A] = P[\{s_{90}, \ldots, s_{100}\}] = 11 \times 0.02 = 0.22$

(4) $P[F] = P[\{s_{51}, \ldots, s_{59}\}] = 9 \times 0.02 = 0.18$

(5) $P[T \geq 80] = P[\{s_{80}, \ldots, s_{100}\}] = 21 \times 0.02 = 0.42$

(6) $P[T < 90] = P[\{s_{51}, s_{52}, \ldots, s_{89}\}] = 39 \times 0.02 = 0.78$

(7) $P[\text{a } C \text{ grade or better}] = P[\{s_{70}, \ldots, s_{100}\}] = 31 \times 0.02 = 0.62$

(8) $P[\text{student passes}] = P[\{s_{60}, \ldots, s_{100}\}] = 41 \times 0.02 = 0.82$

Quiz 1.4

We can describe this experiment by the event space consisting of the four possible events VB, VL, DB, and DL. We represent these events in the table:

	V	D
L	0.35	?
B	?	?

In a roundabout way, the problem statement tells us how to fill in the table. In particular,

$$P[V] = 0.7 = P[VL] + P[VB] \tag{1}$$
$$P[L] = 0.6 = P[VL] + P[DL] \tag{2}$$

Since $P[VL] = 0.35$, we can conclude that $P[VB] = 0.35$ and that $P[DL] = 0.6 - 0.35 = 0.25$. This allows us to fill in two more table entries:

	V	D
L	0.35	0.25
B	0.35	?

The remaining table entry is filled in by observing that the probabilities must sum to 1. This implies $P[DB] = 0.05$ and the complete table is

	V	D
L	0.35	0.25
B	0.35	0.05

Finding the various probabilities is now straightforward:

(1) $P[DL] = 0.25$

(2) $P[D \cup L] = P[VL] + P[DL] + P[DB] = 0.35 + 0.25 + 0.05 = 0.65.$

(3) $P[VB] = 0.35$

(4) $P[V \cup L] = P[V] + P[L] - P[VL] = 0.7 + 0.6 - 0.35 = 0.95$

(5) $P[V \cup D] = P[S] = 1$

(6) $P[LB] = P[LL^c] = 0$

Quiz 1.5

(1) The probability of exactly two voice calls is

$$P[N_V = 2] = P[\{vvd, vdv, dvv\}] = 0.3 \tag{1}$$

(2) The probability of at least one voice call is

$$P[N_V \geq 1] = P[\{vdd, dvd, ddv, vvd, vdv, dvv, vvv\}] \tag{2}$$
$$= 6(0.1) + 0.2 = 0.8 \tag{3}$$

An easier way to get the same answer is to observe that

$$P[N_V \geq 1] = 1 - P[N_V < 1] = 1 - P[N_V = 0] = 1 - P[\{ddd\}] = 0.8 \tag{4}$$

(3) The conditional probability of two voice calls followed by a data call given that there were two voice calls is

$$P[\{vvd\} | N_V = 2] = \frac{P[\{vvd\}, N_V = 2]}{P[N_V = 2]} = \frac{P[\{vvd\}]}{P[N_V = 2]} = \frac{0.1}{0.3} = \frac{1}{3} \tag{5}$$

(4) The conditional probability of two data calls followed by a voice call given there were two voice calls is

$$P[\{ddv\} | N_V = 2] = \frac{P[\{ddv\}, N_V = 2]}{P[N_V = 2]} = 0 \tag{6}$$

The joint event of the outcome ddv and exactly two voice calls has probability zero since there is only one voice call in the outcome ddv.

(5) The conditional probability of exactly two voice calls given at least one voice call is

$$P[N_V = 2 | N_v \geq 1] = \frac{P[N_V = 2, N_V \geq 1]}{P[N_V \geq 1]} = \frac{P[N_V = 2]}{P[N_V \geq 1]} = \frac{0.3}{0.8} = \frac{3}{8} \tag{7}$$

(6) The conditional probability of at least one voice call given there were exactly two voice calls is

$$P[N_V \geq 1 | N_V = 2] = \frac{P[N_V \geq 1, N_V = 2]}{P[N_V = 2]} = \frac{P[N_V = 2]}{P[N_V = 2]} = 1 \tag{8}$$

Given that there were two voice calls, there must have been at least one voice call.

Quiz 1.6

In this experiment, there are four outcomes with probabilities

$$P[\{vv\}] = (0.8)^2 = 0.64 \qquad P[\{vd\}] = (0.8)(0.2) = 0.16$$
$$P[\{dv\}] = (0.2)(0.8) = 0.16 \qquad P[\{dd\}] = (0.2)^2 = 0.04$$

When checking the independence of any two events A and B, it's wise to avoid intuition and simply check whether $P[AB] = P[A]P[B]$. Using the probabilities of the outcomes, we now can test for the independence of events.

(1) First, we calculate the probability of the joint event:

$$P[N_V = 2, N_V \geq 1] = P[N_V = 2] = P[\{vv\}] = 0.64 \tag{1}$$

Next, we observe that

$$P[N_V \geq 1] = P[\{vd, dv, vv\}] = 0.96 \tag{2}$$

Finally, we make the comparison

$$P[N_V = 2] P[N_V \geq 1] = (0.64)(0.96) \neq P[N_V = 2, N_V \geq 1] \tag{3}$$

which shows the two events are dependent.

(2) The probability of the joint event is

$$P[N_V \geq 1, C_1 = v] = P[\{vd, vv\}] = 0.80 \tag{4}$$

From part (a), $P[N_V \geq 1] = 0.96$. Further, $P[C_1 = v] = 0.8$ so that

$$P[N_V \geq 1] P[C_1 = v] = (0.96)(0.8) = 0.768 \neq P[N_V \geq 1, C_1 = v] \tag{5}$$

Hence, the events are dependent.

(3) The problem statement that the calls were independent implies that the events the second call is a voice call, $\{C_2 = v\}$, and the first call is a data call, $\{C_1 = d\}$ are independent events. Just to be sure, we can do the calculations to check:

$$P[C_1 = d, C_2 = v] = P[\{dv\}] = 0.16 \tag{6}$$

Since $P[C_1 = d]P[C_2 = v] = (0.2)(0.8) = 0.16$, we confirm that the events are independent. Note that this shouldn't be surprising since we used the information that the calls were independent in the problem statement to determine the probabilities of the outcomes.

(4) The probability of the joint event is

$$P[C_2 = v, N_V \text{ is even}] = P[\{vv\}] = 0.64 \tag{7}$$

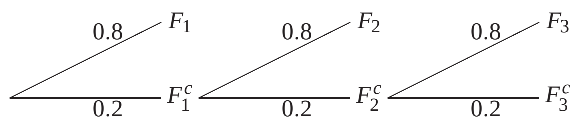
Also, each event has probability

$$P[C_2 = v] = P[\{dv, vv\}] = 0.8, \qquad P[N_V \text{ is even}] = P[\{dd, vv\}] = 0.68 \tag{8}$$

Thus, $P[C_2 = v]P[N_V \text{ is even}] = (0.8)(0.68) = 0.544$. Since $P[C_2 = v, N_V \text{ is even}] \neq 0.544$, the events are dependent.

Quiz 1.7

Let F_i denote the event that that the user is found on page i. The tree for the experiment is



The user is found unless all three paging attempts fail. Thus the probability the user is found is

$$P[F] = 1 - P\left[F_1^c F_2^c F_3^c\right] = 1 - (0.2)^3 = 0.992 \tag{1}$$

Quiz 1.8

(1) We can view choosing each bit in the code word as a subexperiment. Each subexperiment has two possible outcomes: 0 and 1. Thus by the fundamental principle of counting, there are $2 \times 2 \times 2 \times 2 = 2^4 = 16$ possible code words.

(2) An experiment that can yield all possible code words with two zeroes is to choose which 2 bits (out of 4 bits) will be zero. The other two bits then must be ones. There are $\binom{4}{2} = 6$ ways to do this. Hence, there are six code words with exactly two zeroes. For this problem, it is also possible to simply enumerate the six code words:

$$1100, \quad 1010, \quad 1001, \quad 0101, \quad 0110, \quad 0011.$$

(3) When the first bit must be a zero, then the first subexperiment of choosing the first bit has only one outcome. For each of the next three bits, we have two choices. In this case, there are $1 \times 2 \times 2 \times 2 = 8$ ways of choosing a code word.

(4) For the constant ratio code, we can specify a code word by choosing M of the bits to be ones. The other $N - M$ bits will be zeroes. The number of ways of choosing such a code word is $\binom{N}{M}$. For $N = 8$ and $M = 3$, there are $\binom{8}{3} = 56$ code words.

Quiz 1.9

(1) In this problem, k bits received in error is the same as k failures in 100 trials. The failure probability is $\epsilon = 1 - p$ and the success probability is $1 - \epsilon = p$. That is, the probability of k bits in error and $100 - k$ correctly received bits is

$$P\left[S_{k,100-k}\right] = \binom{100}{k} \epsilon^k (1-\epsilon)^{100-k} \tag{1}$$

For $\epsilon = 0.01$,

$$P\left[S_{0,100}\right] = (1-\epsilon)^{100} = (0.99)^{100} = 0.3660 \quad (2)$$

$$P\left[S_{1,99}\right] = 100(0.01)(0.99)^{99} = 0.3700 \quad (3)$$

$$P\left[S_{2,98}\right] = 4950(0.01)^{2}(0.99)^{9}8 = 0.1849 \quad (4)$$

$$P\left[S_{3,97}\right] = 161,700(0.01)^{3}(0.99)^{97} = 0.0610 \quad (5)$$

(2) The probability a packet is decoded correctly is just

$$P[C] = P\left[S_{0,100}\right] + P\left[S_{1,99}\right] + P\left[S_{2,98}\right] + P\left[S_{3,97}\right] = 0.9819 \quad (6)$$

Quiz 1.10

Since the chip works only if all n transistors work, the transistors in the chip are like devices in series. The probability that a chip works is $P[C] = p^n$.

The module works if either 8 chips work or 9 chips work. Let C_k denote the event that exactly k chips work. Since transistor failures are independent of each other, chip failures are also independent. Thus each $P[C_k]$ has the binomial probability

$$P[C_8] = \binom{9}{8}(P[C])^{8}(1-P[C])^{9-8} = 9p^{8n}(1-p^n), \quad (1)$$

$$P[C_9] = (P[C])^{9} = p^{9n}. \quad (2)$$

The probability a memory module works is

$$P[M] = P[C_8] + P[C_9] = p^{8n}(9-8p^n) \quad (3)$$

Quiz 1.11

```
R=rand(1,100);
X=(R<= 0.4) ...
    + (2*(R>0.4).*(R<=0.9)) ...
    + (3*(R>0.9));
Y=hist(X,1:3)
```

For a MATLAB simulation, we first generate a vector `R` of 100 random numbers. Second, we generate vector `X` as a function of `R` to represent the 3 possible outcomes of a flip. That is, `X(i)=1` if flip i was heads, `X(i)=2` if flip i was tails, and `X(i)=3)` is flip i landed on the edge.

To see how this works, we note there are three cases:

- If `R(i) <= 0.4`, then `X(i)=1`.
- If `0.4 < R(i)` and `R(i)<=0.9`, then `X(i)=2`.
- If `0.9 < R(i)`, then `X(i)=3`.

These three cases will have probabilities 0.4, 0.5 and 0.1. Lastly, we use the `hist` function to count how many occurences of each possible value of `X(i)`.

Quiz Solutions – Chapter 2

Quiz 2.1

The sample space, probabilities and corresponding grades for the experiment are

Outcome	$P[\cdot]$	G
BB	0.36	3.0
BC	0.24	2.5
CB	0.24	2.5
CC	0.16	2

Quiz 2.2

(1) To find c, we recall that the PMF must sum to 1. That is,

$$\sum_{n=1}^{3} P_N(n) = c\left(1 + \frac{1}{2} + \frac{1}{3}\right) = 1 \tag{1}$$

This implies $c = 6/11$. Now that we have found c, the remaining parts are straightforward.

(2) $P[N = 1] = P_N(1) = c = 6/11$

(3) $P[N \geq 2] = P_N(2) + P_N(3) = c/2 + c/3 = 5/11$

(4) $P[N > 3] = \sum_{n=4}^{\infty} P_N(n) = 0$

Quiz 2.3

Decoding each transmitted bit is an independent trial where we call a bit error a "success." Each bit is in error, that is, the trial is a success, with probability p. Now we can interpret each experiment in the generic context of independent trials.

(1) The random variable X is the number of trials up to and including the first success. Similar to Example 2.11, X has the geometric PMF

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

(2) If $p = 0.1$, then the probability exactly 10 bits are sent is

$$P[X = 10] = P_X(10) = (0.1)(0.9)^9 = 0.0387 \tag{2}$$

The probability that at least 10 bits are sent is $P[X \geq 10] = \sum_{x=10}^{\infty} P_X(x)$. This sum is not too hard to calculate. However, its even easier to observe that $X \geq 10$ if the first 10 bits are transmitted correctly. That is,

$$P[X \geq 10] = P[\text{first 10 bits are correct}] = (1-p)^{10} \tag{3}$$

For $p = 0.1$, $P[X \geq 10] = 0.9^{10} = 0.3487$.

(3) The random variable Y is the number of successes in 100 independent trials. Just as in Example 2.13, Y has the binomial PMF

$$P_Y(y) = \binom{100}{y} p^y (1-p)^{100-y} \tag{4}$$

If $p = 0.01$, the probability of exactly 2 errors is

$$P[Y = 2] = P_Y(2) = \binom{100}{2} (0.01)^2 (0.99)^{98} = 0.1849 \tag{5}$$

(4) The probability of no more than 2 errors is

$$\begin{aligned} P[Y \leq 2] &= P_Y(0) + P_Y(1) + P_Y(2) && (6) \\ &= (0.99)^{100} + 100(0.01)(0.99)^{99} + \binom{100}{2}(0.01)^2(0.99)^{98} && (7) \\ &= 0.9207 && (8) \end{aligned}$$

(5) Random variable Z is the number of trials up to and including the third success. Thus Z has the Pascal PMF (see Example 2.15)

$$P_Z(z) = \binom{z-1}{2} p^3 (1-p)^{z-3} \tag{9}$$

Note that $P_Z(z) > 0$ for $z = 3, 4, 5, \ldots$.

(6) If $p = 0.25$, the probability that the third error occurs on bit 12 is

$$P_Z(12) = \binom{11}{2} (0.25)^3 (0.75)^9 = 0.0645 \tag{10}$$

Quiz 2.4

Each of these probabilities can be read off the CDF $F_Y(y)$. However, we must keep in mind that when $F_Y(y)$ has a discontinuity at y_0, $F_Y(y)$ takes the upper value $F_Y(y_0^+)$.

(1) $P[Y < 1] = F_Y(1^-) = 0$

(2) $P[Y \leq 1] = F_Y(1) = 0.6$

(3) $P[Y > 2] = 1 - P[Y \leq 2] = 1 - F_Y(2) = 1 - 0.8 = 0.2$

(4) $P[Y \geq 2] = 1 - P[Y < 2] = 1 - F_Y(2^-) = 1 - 0.6 = 0.4$

(5) $P[Y = 1] = P[Y \leq 1] - P[Y < 1] = F_Y(1^+) - F_Y(1^-) = 0.6$

(6) $P[Y = 3] = P[Y \leq 3] - P[Y < 3] = F_Y(3^+) - F_Y(3^-) = 0.8 - 0.8 = 0$

Quiz 2.5

(1) With probability 0.7, a call is a voice call and $C = 25$. Otherwise, with probability 0.3, we have a data call and $C = 40$. This corresponds to the PMF

$$P_C(c) = \begin{cases} 0.7 & c = 25 \\ 0.3 & c = 40 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

(2) The expected value of C is

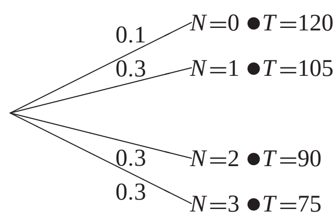
$$E[C] = 25(0.7) + 40(0.3) = 29.5 \text{ cents} \tag{2}$$

Quiz 2.6

(1) As a function of N, the cost T is

$$T = 25N + 40(3 - N) = 120 - 15N \tag{1}$$

(2) To find the PMF of T, we can draw the following tree:



From the tree, we can write down the PMF of T:

$$P_T(t) = \begin{cases} 0.3 & t = 75, 90, 105 \\ 0.1 & t = 120 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

From the PMF $P_T(t)$, the expected value of T is

$$\begin{aligned} E[T] &= 75P_T(75) + 90P_T(90) + 105P_T(105) + 120P_T(120) \qquad (3) \\ &= (75 + 90 + 105)(0.3) + 120(0.1) = 62 \qquad (4) \end{aligned}$$

Quiz 2.7

(1) Using Definition 2.14, the expected number of applications is

$$E[A] = \sum_{a=1}^{4} a P_A(a) = 1(0.4) + 2(0.3) + 3(0.2) + 4(0.1) = 2 \tag{1}$$

(2) The number of memory chips is $M = g(A)$ where

$$g(A) = \begin{cases} 4 & A = 1, 2 \\ 6 & A = 3 \\ 8 & A = 4 \end{cases} \tag{2}$$

(3) By Theorem 2.10, the expected number of memory chips is

$$E[M] = \sum_{a=1}^{4} g(A) P_A(a) = 4(0.4) + 4(0.3) + 6(0.2) + 8(0.1) = 4.8 \tag{3}$$

Since $E[A] = 2$, $g(E[A]) = g(2) = 4$. However, $E[M] = 4.8 \neq g(E[A])$. The two quantities are different because $g(A)$ is not of the form $\alpha A + \beta$.

Quiz 2.8

The PMF $P_N(n)$ allows to calculate each of the desired quantities.

(1) The expected value of N is

$$E[N] = \sum_{n=0}^{2} n P_N(n) = 0(0.1) + 1(0.4) + 2(0.5) = 1.4 \tag{1}$$

(2) The second moment of N is

$$E\left[N^2\right] = \sum_{n=0}^{2} n^2 P_N(n) = 0^2(0.1) + 1^2(0.4) + 2^2(0.5) = 2.4 \tag{2}$$

(3) The variance of N is

$$\text{Var}[N] = E\left[N^2\right] - (E[N])^2 = 2.4 - (1.4)^2 = 0.44 \tag{3}$$

(4) The standard deviation is $\sigma_N = \sqrt{\text{Var}[N]} = \sqrt{0.44} = 0.663$.

Quiz 2.9

(1) From the problem statement, we learn that the conditional PMF of N given the event I is

$$P_{N|I}(n) = \begin{cases} 0.02 & n = 1, 2, \ldots, 50 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

(2) Also from the problem statement, the conditional PMF of N given the event T is

$$P_{N|T}(n) = \begin{cases} 0.2 & n = 1, 2, 3, 4, 5 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

(3) The problem statement tells us that $P[T] = 1 - P[I] = 3/4$. From Theorem 1.10 (the law of total probability), we find the PMF of N is

$$P_N(n) = P_{N|T}(n)\, P[T] + P_{N|I}(n)\, P[I] \tag{3}$$

$$= \begin{cases} 0.2(0.75) + 0.02(0.25) & n = 1, 2, 3, 4, 5 \\ 0(0.75) + 0.02(0.25) & n = 6, 7, \ldots, 50 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$= \begin{cases} 0.155 & n = 1, 2, 3, 4, 5 \\ 0.005 & n = 6, 7, \ldots, 50 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

(4) First we find

$$P[N \leq 10] = \sum_{n=1}^{10} P_N(n) = (0.155)(5) + (0.005)(5) = 0.80 \tag{6}$$

By Theorem 2.17, the conditional PMF of N given $N \leq 10$ is

$$P_{N|N\leq 10}(n) = \begin{cases} \frac{P_N(n)}{P[N\leq 10]} & n \leq 10 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

$$= \begin{cases} 0.155/0.8 & n = 1, 2, 3, 4, 5 \\ 0.005/0.8 & n = 6, 7, 8, 9, 10 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

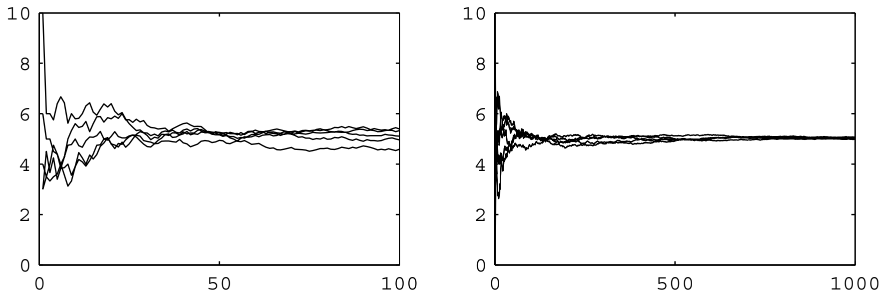
$$= \begin{cases} 0.19375 & n = 1, 2, 3, 4, 5 \\ 0.00625 & n = 6, 7, 8, 9, 10 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

(5) Once we have the conditional PMF, calculating conditional expectations is easy.

$$E[N|N \leq 10] = \sum_n n P_{N|N\leq 10}(n) \tag{10}$$

$$= \sum_{n=1}^{5} n(0.19375) + \sum_{n=6}^{10} n(0.00625) \tag{11}$$

$$= 3.15625 \tag{12}$$



(a) `samplemean(100)` **(b)** `samplemean(1000)`

Figure 1: Two examples of the output of `samplemean(k)`

(6) To find the conditional variance, we first find the conditional second moment

$$E\left[N^2|N \leq 10\right] = \sum_n n^2 P_{N|N\leq 10}(n) \tag{13}$$

$$= \sum_{n=1}^{5} n^2(0.19375) + \sum_{n=6}^{10} n^2(0.00625) \tag{14}$$

$$= 55(0.19375) + 330(0.00625) = 12.71875 \tag{15}$$

The conditional variance is

$$\text{Var}[N|N \leq 10] = E\left[N^2|N \leq 10\right] - (E\left[N|N \leq 10\right])^2 \tag{16}$$

$$= 12.71875 - (3.15625)^2 = 2.75684 \tag{17}$$

Quiz 2.10

The function `samplemean(k)` generates and plots five m_n sequences for $n = 1, 2, \ldots, k$. The ith column `M(:,i)` of `M` holds a sequence $m_1, m_2, \ldots, m_k$.

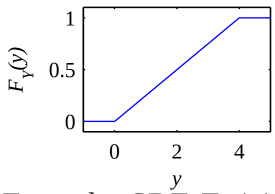
```
function M=samplemean(k);
K=(1:k)';
M=zeros(k,5);
for i=1:5,
    X=duniformrv(0,10,k);
    M(:,i)=cumsum(X)./K;
end;
plot(K,M);
```

Examples of the function calls (a) `samplemean(100)` and (b) `samplemean(1000)` are shown in Figure 1. Each time `samplemean(k)` is called produces a random output. What is observed in these figures is that for small n, m_n is fairly random but as n gets large, m_n gets close to $E[X] = 5$. Although each sequence $m_1, m_2, \ldots$ that we generate is random, the sequences always converges to $E[X]$. This random convergence is analyzed in Chapter 7.

Quiz Solutions – Chapter 3

Quiz 3.1

The CDF of Y is



$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y/4 & 0 \le y \le 4 \\ 1 & y > 4 \end{cases} \tag{1}$$

From the CDF $F_Y(y)$, we can calculate the probabilities:

(1) $P[Y \le -1] = F_Y(-1) = 0$

(2) $P[Y \le 1] = F_Y(1) = 1/4$

(3) $P[2 < Y \le 3] = F_Y(3) - F_Y(2) = 3/4 - 2/4 = 1/4$

(4) $P[Y > 1.5] = 1 - P[Y \le 1.5] = 1 - F_Y(1.5) = 1 - (1.5)/4 = 5/8$

Quiz 3.2

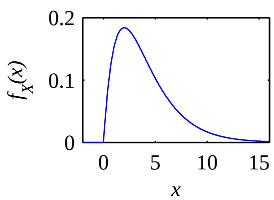
(1) First we will find the constant c and then we will sketch the PDF. To find c, we use the fact that $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$. We will evaluate this integral using integration by parts:

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^{\infty} cxe^{-x/2}\,dx \tag{1}$$

$$= \underbrace{-2cxe^{-x/2}\Big|_0^{\infty}}_{=0} + \int_0^{\infty} 2ce^{-x/2}\,dx \tag{2}$$

$$= -4ce^{-x/2}\Big|_0^{\infty} = 4c \tag{3}$$

Thus $c = 1/4$ and X has the Erlang $(n = 2, \lambda = 1/2)$ PDF



$$f_X(x) = \begin{cases} (x/4)e^{-x/2} & x \ge 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

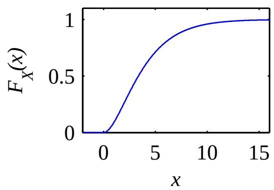
(2) To find the CDF $F_X(x)$, we first note X is a nonnegative random variable so that $F_X(x) = 0$ for all $x < 0$. For $x \geq 0$,

$$F_X(x) = \int_0^x f_X(y)\, dy = \int_0^x \frac{y}{4} e^{-y/2}\, dy \tag{5}$$

$$= -\frac{y}{2} e^{-y/2} \Big|_0^x - \int_0^x -\frac{1}{2} e^{-y/2}\, dy \tag{6}$$

$$= 1 - \frac{x}{2} e^{-x/2} - e^{-x/2} \tag{7}$$

The complete expression for the CDF is



$$F_X(x) = \begin{cases} 1 - \left(\frac{x}{2} + 1\right) e^{-x/2} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

(3) From the CDF $F_X(x)$,

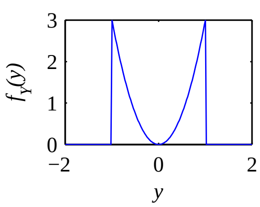
$$P[0 \leq X \leq 4] = F_X(4) - F_X(0) = 1 - 3e^{-2}. \tag{9}$$

(4) Similarly,

$$P[-2 \leq X \leq 2] = F_X(2) - F_X(-2) = 1 - 3e^{-1}. \tag{10}$$

Quiz 3.3

The PDF of Y is



$$f_Y(y) = \begin{cases} 3y^2/2 & -1 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

(1) The expected value of Y is

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\, dy = \int_{-1}^{1} (3/2) y^3\, dy = (3/8) y^4 \Big|_{-1}^{1} = 0. \tag{2}$$

Note that the above calculation wasn't really necessary because $E[Y] = 0$ whenever the PDF $f_Y(y)$ is an even function (i.e., $f_Y(y) = f_Y(-y)$).

(2) The second moment of Y is

$$E\left[Y^2\right] = \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy = \int_{-1}^{1} (3/2) y^4\, dy = (3/10) y^5 \Big|_{-1}^{1} = 3/5. \tag{3}$$

(3) The variance of Y is

$$\text{Var}[Y] = E\left[Y^2\right] - (E[Y])^2 = 3/5. \tag{4}$$

(4) The standard deviation of Y is $\sigma_Y = \sqrt{\text{Var}[Y]} = \sqrt{3/5}$.

Quiz 3.4

(1) When X is an exponential (λ) random variable, $E[X] = 1/\lambda$ and $\text{Var}[X] = 1/\lambda^2$. Since $E[X] = 3$ and $\text{Var}[X] = 9$, we must have $\lambda = 1/3$. The PDF of X is

$$f_X(x) = \begin{cases} (1/3)e^{-x/3} & x \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

(2) We know X is a uniform (a, b) random variable. To find a and b, we apply Theorem 3.6 to write

$$E[X] = \frac{a+b}{2} = 3 \qquad \text{Var}[X] = \frac{(b-a)^2}{12} = 9. \tag{2}$$

This implies

$$a + b = 6, \qquad b - a = \pm 6\sqrt{3}. \tag{3}$$

The only valid solution with $a < b$ is

$$a = 3 - 3\sqrt{3}, \qquad b = 3 + 3\sqrt{3}. \tag{4}$$

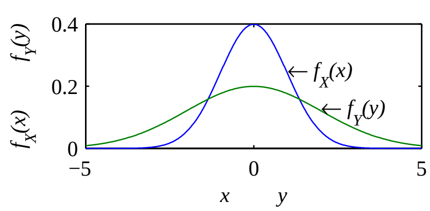
The complete expression for the PDF of X is

$$f_X(x) = \begin{cases} 1/(6\sqrt{3}) & 3 - 3\sqrt{3} \leq x < 3 + 3\sqrt{3}, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Quiz 3.5

Each of the requested probabilities can be calculated using $\Phi(z)$ function and Table 3.1 or $Q(z)$ and Table 3.2. We start with the sketches.

(1) The PDFs of X and Y are shown below. The fact that Y has twice the standard deviation of X is reflected in the greater spread of $f_Y(y)$. However, it is important to remember that as the standard deviation increases, the peak value of the Gaussian PDF goes down.



(2) Since X is Gaussian (0, 1),

$$P[-1 < X \leq 1] = F_X(1) - F_X(-1) \tag{1}$$
$$= \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = 0.6826. \tag{2}$$

(3) Since Y is Gaussian (0, 2),

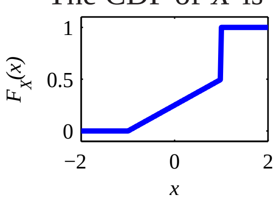
$$P[-1 < Y \leq 1] = F_Y(1) - F_Y(-1) \tag{3}$$
$$= \Phi\left(\frac{1}{\sigma_Y}\right) - \Phi\left(\frac{-1}{\sigma_Y}\right) = 2\Phi\left(\frac{1}{2}\right) - 1 = 0.383. \tag{4}$$

(4) Again, since X is Gaussian (0, 1), $P[X > 3.5] = Q(3.5) = 2.33 \times 10^{-4}$.

(5) Since Y is Gaussian (0, 2), $P[Y > 3.5] = Q(\frac{3.5}{2}) = Q(1.75) = 1 - \Phi(1.75) = 0.0401$.

Quiz 3.6

The CDF of X is



$$F_X(x) = \begin{cases} 0 & x < -1, \\ (x+1)/4 & -1 \leq x < 1, \\ 1 & x \geq 1. \end{cases} \tag{1}$$

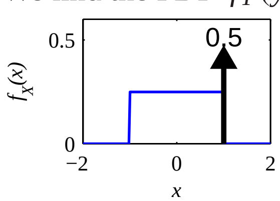
The following probabilities can be read directly from the CDF:

(1) $P[X \leq 1] = F_X(1) = 1$.

(2) $P[X < 1] = F_X(1^-) = 1/2$.

(3) $P[X = 1] = F_X(1^+) - F_X(1^-) = 1 - 1/2 = 1/2$.

(4) We find the PDF $f_Y(y)$ by taking the derivative of $F_Y(y)$. The resulting PDF is



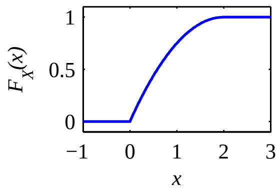
$$f_X(x) = \begin{cases} 1/4 & -1 \leq x < 1, \\ (1/2)\delta(x-1) & x = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Quiz 3.7

(1) Since X is always nonnegative, $F_X(x) = 0$ for $x < 0$. Also, $F_X(x) = 1$ for $x \geq 2$ since its always true that $x \leq 2$. Lastly, for $0 \leq x \leq 2$,

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy = \int_0^x (1 - y/2)\, dy = x - x^2/4. \tag{1}$$

The complete CDF of X is



$$F_X(x) = \begin{cases} 0 & x < 0, \\ x - x^2/4 & 0 \leq x \leq 2, \\ 1 & x > 2. \end{cases} \tag{2}$$

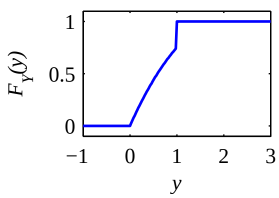
(2) The probability that $Y = 1$ is

$$P[Y = 1] = P[X \geq 1] = 1 - F_X(1) = 1 - 3/4 = 1/4. \tag{3}$$

(3) Since X is nonnegative, Y is also nonnegative. Thus $F_Y(y) = 0$ for $y < 0$. Also, because $Y \leq 1$, $F_Y(y) = 1$ for all $y \geq 1$. Finally, for $0 < y < 1$,

$$F_Y(y) = P[Y \leq y] = P[X \leq y] = F_X(y). \tag{4}$$

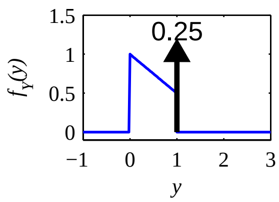
Using the CDF $F_X(x)$, the complete expression for the CDF of Y is



$$F_Y(y) = \begin{cases} 0 & y < 0, \\ y - y^2/4 & 0 \leq y < 1, \\ 1 & y \geq 1. \end{cases} \tag{5}$$

As expected, we see that the jump in $F_Y(y)$ at $y = 1$ is exactly equal to $P[Y = 1]$.

(4) By taking the derivative of $F_Y(y)$, we obtain the PDF $f_Y(y)$. Note that when $y < 0$ or $y > 1$, the PDF is zero.



$$f_Y(y) = \begin{cases} 1 - y/2 + (1/4)\delta(y-1) & 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Quiz 3.8

(1) $P[Y \leq 6] = \int_{-\infty}^{6} f_Y(y)\, dy = \int_0^6 (1/10)\, dy = 0.6$.

(2) From Definition 3.15, the conditional PDF of Y given $Y \leq 6$ is

$$f_{Y|Y\leq 6}(y) = \begin{cases} \frac{f_Y(y)}{P[Y\leq 6]} & y \leq 6, \\ 0 & \text{otherwise,} \end{cases} = \begin{cases} 1/6 & 0 \leq y \leq 6, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

(3) The probability $Y > 8$ is

$$P[Y > 8] = \int_8^{10} \frac{1}{10}\, dy = 0.2\,. \quad (2)$$

(4) From Definition 3.15, the conditional PDF of Y given $Y > 8$ is

$$f_{Y|Y>8}(y) = \begin{cases} \frac{f_Y(y)}{P[Y>8]} & y > 8, \\ 0 & \text{otherwise,} \end{cases} = \begin{cases} 1/2 & 8 < y \leq 10, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

(5) From the conditional PDF $f_{Y|Y\leq 6}(y)$, we can calculate the conditional expectation

$$E[Y|Y \leq 6] = \int_{-\infty}^{\infty} y f_{Y|Y\leq 6}(y)\; dy = \int_0^6 \frac{y}{6}\, dy = 3. \quad (4)$$

(6) From the conditional PDF $f_{Y|Y>8}(y)$, we can calculate the conditional expectation

$$E[Y|Y > 8] = \int_{-\infty}^{\infty} y f_{Y|Y>8}(y)\; dy = \int_8^{10} \frac{y}{2}\, dy = 9. \quad (5)$$

Quiz 3.9

A natural way to produce random variables with PDF $f_{T|T>2}(t)$ is to generate samples of T with PDF $f_T(t)$ and then to discard those samples which fail to satisfy the condition $T > 2$. Here is a MATLAB function that uses this method:

```
function t=t2rv(m)
i=0;lambda=1/3;
t=zeros(m,1);
while (i<m),
    x=exponentialrv(lambda,1);
    if (x>2)
        t(i+1)=x;
        i=i+1;
    end
end
```

A second method exploits the fact that if T is an exponential (λ) random variable, then $T' = T + 2$ has PDF $f_{T'}(t) = f_{T|T>2}(t)$. In this case the command

```
t=2.0+exponentialrv(1/3,m)
```

generates the vector `t`.

Quiz Solutions – Chapter 4

Quiz 4.1

Each value of the joint CDF can be found by considering the corresponding probability.

(1) $F_{X,Y}(-\infty, 2) = P[X \leq -\infty, Y \leq 2] \leq P[X \leq -\infty] = 0$ since X cannot take on the value $-\infty$.

(2) $F_{X,Y}(\infty, \infty) = P[X \leq \infty, Y \leq \infty] = 1$. This result is given in Theorem 4.1.

(3) $F_{X,Y}(\infty, y) = P[X \leq \infty, Y \leq y] = P[Y \leq y] = F_Y(y)$.

(4) $F_{X,Y}(\infty, -\infty) = P[X \leq \infty, Y \leq -\infty] = 0$ since Y cannot take on the value $-\infty$.

Quiz 4.2

From the joint PMF of Q and G given in the table, we can calculate the requested probabilities by summing the PMF over those values of Q and G that correspond to the event.

(1) The probability that $Q = 0$ is

$$P[Q = 0] = P_{Q,G}(0,0) + P_{Q,G}(0,1) + P_{Q,G}(0,2) + P_{Q,G}(0,3) \tag{1}$$

$$= 0.06 + 0.18 + 0.24 + 0.12 = 0.6 \tag{2}$$

(2) The probability that $Q = G$ is

$$P[Q = G] = P_{Q,G}(0,0) + P_{Q,G}(1,1) = 0.18 \tag{3}$$

(3) The probability that $G > 1$ is

$$P[G > 1] = \sum_{g=2}^{3} \sum_{q=0}^{1} P_{Q,G}(q, g) \tag{4}$$

$$= 0.24 + 0.16 + 0.12 + 0.08 = 0.6 \tag{5}$$

(4) The probability that $G > Q$ is

$$P[G > Q] = \sum_{q=0}^{1} \sum_{g=q+1}^{3} P_{Q,G}(q, g) \tag{6}$$

$$= 0.18 + 0.24 + 0.12 + 0.16 + 0.08 = 0.78 \tag{7}$$

Quiz 4.3

By Theorem 4.3, the marginal PMF of H is

$$P_H(h) = \sum_{b=0,2,4} P_{H,B}(h,b) \tag{1}$$

For each value of h, this corresponds to calculating the row sum across the table of the joint PMF. Similarly, the marginal PMF of B is

$$P_B(b) = \sum_{h=-1}^{1} P_{H,B}(h,b) \tag{2}$$

For each value of b, this corresponds to the column sum down the table of the joint PMF. The easiest way to calculate these marginal PMFs is to simply sum each row and column:

$P_{H,B}(h,b)$	$b=0$	$b=2$	$b=4$	$P_H(h)$
$h=-1$	0	0.4	0.2	0.6
$h=0$	0.1	0	0.1	0.2
$h=1$	0.1	0.1	0	0.2
$P_B(b)$	0.2	0.5	0.3	

(3)

Quiz 4.4

To find the constant c, we apply $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$. Specifically,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\;dx\,dy = \int_0^2\int_0^1 cxy\,dx\,dy \tag{1}$$

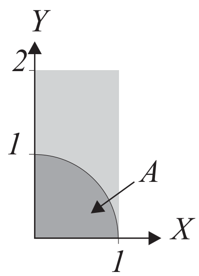
$$= c\int_0^2 y\left(x^2/2\Big|_0^1\right)\,dy \tag{2}$$

$$= (c/2)\int_0^2 y\,dy = (c/4)y^2\Big|_0^2 = c \tag{3}$$

Thus $c = 1$. To calculate $P[A]$, we write

$$P[A] = \iint_A f_{X,Y}(x,y)\;dx\,dy \tag{4}$$

To integrate over A, we convert to polar coordinates using the substitutions $x = r\cos\theta$, $y = r\sin\theta$ and $dx\,dy = r\,dr\,d\theta$, yielding



$$P[A] = \int_0^{\pi/2}\int_0^1 r^2\sin\theta\cos\theta\, r\,dr\,d\theta \tag{5}$$

$$= \left(\int_0^1 r^3\,dr\right)\left(\int_0^{\pi/2}\sin\theta\cos\theta\,d\theta\right) \tag{6}$$

$$= \left(r^4/4\Big|_0^1\right)\left(\frac{\sin^2\theta}{2}\Big|_0^{\pi/2}\right) = 1/8 \tag{7}$$

Quiz 4.5

By Theorem 4.8, the marginal PDF of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy \tag{1}$$

For $x < 0$ or $x > 1$, $f_X(x) = 0$. For $0 \le x \le 1$,

$$f_X(x) = \frac{6}{5}\int_0^1 (x + y^2)\, dy = \frac{6}{5}\left(xy + y^3/3\right)\Big|_{y=0}^{y=1} = \frac{6}{5}(x + 1/3) = \frac{6x+2}{5} \tag{2}$$

The complete expression for the PDf of X is

$$f_X(x) = \begin{cases} (6x+2)/5 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

By the same method we obtain the marginal PDF for Y. For $0 \le y \le 1$,

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy \tag{4}$$

$$= \frac{6}{5}\int_0^1 (x + y^2)\, dx = \frac{6}{5}\left(x^2/2 + xy^2\right)\Big|_{x=0}^{x=1} = \frac{6}{5}(1/2 + y^2) = \frac{3+6y^2}{5} \tag{5}$$

Since $f_Y(y) = 0$ for $y < 0$ or $y > 1$, the complete expression for the PDF of Y is

$$f_Y(y) = \begin{cases} (3+6y^2)/5 & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Quiz 4.6

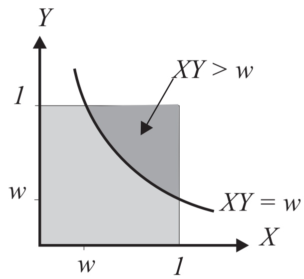
(A) The time required for the transfer is $T = L/B$. For each pair of values of L and B, we can calculate the time T needed for the transfer. We can write these down on the table for the joint PMF of L and B as follows:

$P_{L,B}(l, b)$	$b = 14,400$	$b = 21,600$	$b = 28,800$
$l = 518,400$	0.20 (T=36)	0.10 (T=24)	0.05 (T=18)
$l = 2,592,000$	0.05 (T=180)	0.10 (T=120)	0.20 (T=90)
$l = 7,776,000$	0.00 (T=540)	0.10 (T=360)	0.20 (T=270)

From the table, writing down the PMF of T is straightforward.

$$P_T(t) = \begin{cases} 0.05 & t = 18 \\ 0.1 & t = 24 \\ 0.2 & t = 36, 90 \\ 0.1 & t = 120 \\ 0.05 & t = 180 \\ 0.2 & t = 270 \\ 0.1 & t = 360 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

(B) First, we observe that since $0 \leq X \leq 1$ and $0 \leq Y \leq 1$, $W = XY$ satisfies $0 \leq W \leq 1$. Thus $f_W(0) = 0$ and $f_W(1) = 1$. For $0 < w < 1$, we calculate the CDF $F_W(w) = P[W \leq w]$. As shown below, integrating over the region $W \leq w$ is fairly complex. The calculus is simpler if we integrate over the region $XY > w$. Specifically,



$$F_W(w) = 1 - P[XY > w] \tag{2}$$
$$= 1 - \int_w^1 \int_{w/x}^1 dy\,dx \tag{3}$$
$$= 1 - \int_w^1 (1 - w/x)\,dx \tag{4}$$
$$= 1 - \left(x - w \ln x|_{x=w}^{x=1}\right) \tag{5}$$
$$= 1 - (1 - w + w \ln w) = w - w \ln w \tag{6}$$

The complete expression for the CDF is

$$F_W(w) = \begin{cases} 0 & w < 0 \\ w - w \ln w & 0 \leq w \leq 1 \\ 1 & w > 1 \end{cases} \tag{7}$$

By taking the derivative of the CDF, we find the PDF is

$$f_W(w) = \frac{dF_W(w)}{dw} = \begin{cases} 0 & w < 0 \\ -\ln w & 0 \leq w \leq 1 \\ 0 & w > 1 \end{cases} \tag{8}$$

Quiz 4.7

(A) It is helpful to first make a table that includes the marginal PMFs.

$P_{L,T}(l,t)$	$t = 40$	$t = 60$	$P_L(l)$
$l = 1$	0.15	0.1	0.25
$l = 2$	0.3	0.2	0.5
$l = 3$	0.15	0.1	0.25
$P_T(t)$	0.6	0.4	

(1) The expected value of L is

$$E[L] = 1(0.25) + 2(0.5) + 3(0.25) = 2. \tag{1}$$

Since the second moment of L is

$$E\left[L^2\right] = 1^2(0.25) + 2^2(0.5) + 3^2(0.25) = 4.5, \tag{2}$$

the variance of L is

$$\text{Var}[L] = E\left[L^2\right] - (E[L])^2 = 0.5. \tag{3}$$

(2) The expected value of T is

$$E[T] = 40(0.6) + 60(0.4) = 48. \tag{4}$$

The second moment of T is

$$E\left[T^2\right] = 40^2(0.6) + 60^2(0.4) = 2400. \tag{5}$$

Thus

$$\text{Var}[T] = E\left[T^2\right] - (E[T])^2 = 2400 - 48^2 = 96. \tag{6}$$

(3) The correlation is

$$\begin{aligned} E[LT] &= \sum_{t=40,60} \sum_{l=1}^{3} lt P_{LT}(lt) \quad &(7) \\ &= 1(40)(0.15) + 2(40)(0.3) + 3(40)(0.15) \quad &(8) \\ &\quad + 1(60)(0.1) + 2(60)(0.2) + 3(60)(0.1) \quad &(9) \\ &= 96 \quad &(10) \end{aligned}$$

(4) From Theorem 4.16(a), the covariance of L and T is

$$\text{Cov}[L, T] = E[LT] - E[L]E[T] = 96 - 2(48) = 0 \tag{11}$$

(5) Since $\text{Cov}[L, T] = 0$, the correlation coefficient is $\rho_{L,T} = 0$.

(B) As in the discrete case, the calculations become easier if we first calculate the marginal PDFs $f_X(x)$ and $f_Y(y)$. For $0 \le x \le 1$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy = \int_0^2 xy\, dy = \frac{1}{2}xy^2\Big|_{y=0}^{y=2} = 2x \tag{12}$$

Similarly, for $0 \le y \le 2$,

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx = \int_0^2 xy\, dx = \frac{1}{2}x^2y\Big|_{x=0}^{x=1} = \frac{y}{2} \tag{13}$$

The complete expressions for the marginal PDFs are

$$f_X(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \qquad f_Y(y) = \begin{cases} y/2 & 0 \le y \le 2 \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

From the marginal PDFs, it is straightforward to calculate the various expectations.

(1) The first and second moments of X are

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^1 2x^2\, dx = \frac{2}{3} \tag{15}$$

$$E\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \int_0^1 2x^3\, dx = \frac{1}{2} \tag{16}$$

$$\tag{17}$$

The variance of X is $\text{Var}[X] = E[X^2] - (E[X])^2 = 1/18$.

(2) The first and second moments of Y are

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\, dy = \int_0^2 \frac{1}{2}y^2\, dy = \frac{4}{3} \tag{18}$$

$$E\left[Y^2\right] = \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy = \int_0^2 \frac{1}{2}y^3\, dy = 2 \tag{19}$$

The variance of Y is $\text{Var}[Y] = E[Y^2] - (E[Y])^2 = 2 - 16/9 = 2/9$.

(3) The correlation of X and Y is

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x, y)\, dx, dy \tag{20}$$

$$= \int_0^1\int_0^2 x^2y^2\, dx, dy = \left.\frac{x^3}{3}\right|_0^1 \left.\frac{y^3}{3}\right|_0^2 = \frac{8}{9} \tag{21}$$

(4) The covariance of X and Y is

$$\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = \frac{8}{9} - \left(\frac{2}{3}\right)\left(\frac{4}{3}\right) = 0. \tag{22}$$

(5) Since $\text{Cov}[X, Y] = 0$, the correlation coefficient is $\rho_{X,Y} = 0$.

Quiz 4.8

(A) Since the event $V > 80$ occurs only for the pairs $(L, T) = (2, 60)$, $(L, T) = (3, 40)$ and $(L, T) = (3, 60)$,

$$P[A] = P[V > 80] = P_{L,T}(2, 60) + P_{L,T}(3, 40) + P_{L,T}(3, 60) = 0.45 \tag{1}$$

By Definition 4.9,

$$P_{L,T|A}(l, t) = \begin{cases} \frac{P_{L,T}(l,t)}{P[A]} & lt > 80 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

We can represent this conditional PMF in the following table:

$P_{L,T\|A}(l,t)$	$t=40$	$t=60$
$l=1$	0	0
$l=2$	0	4/9
$l=3$	1/3	2/9

The conditional expectation of V can be found from the conditional PMF.

$$E[V|A] = \sum_l \sum_t lt P_{L,T|A}(l,t) \tag{3}$$

$$= (2\cdot 60)\frac{4}{9} + (3\cdot 40)\frac{1}{3} + (3\cdot 60)\frac{2}{9} = 133\,\frac{1}{3} \tag{4}$$

For the conditional variance $\mathrm{Var}[V|A]$, we first find the conditional second moment

$$E\left[V^2|A\right] = \sum_l \sum_t (lt)^2 P_{L,T|A}(l,t) \tag{5}$$

$$= (2\cdot 60)^2\frac{4}{9} + (3\cdot 40)^2\frac{1}{3} + (3\cdot 60)^2\frac{2}{9} = 18,400 \tag{6}$$

It follows that

$$\mathrm{Var}[V|A] = E\left[V^2|A\right] - (E[V|A])^2 = 622\,\frac{2}{9} \tag{7}$$

(B) For continuous random variables X and Y, we first calculate the probability of the conditioning event.

$$P[B] = \iint_B f_{X,Y}(x,y)\,dx\,dy = \int_{40}^{60}\int_{80/y}^{3} \frac{xy}{4000}\,dx\,dy \tag{8}$$

$$= \int_{40}^{60} \frac{y}{4000}\left(\left.\frac{x^2}{2}\right|_{80/y}^{3}\right) dy \tag{9}$$

$$= \int_{40}^{60} \frac{y}{4000}\left(\frac{9}{2} - \frac{3200}{y^2}\right) dy \tag{10}$$

$$= \frac{9}{8} - \frac{4}{5}\ln\frac{3}{2} \approx 0.801 \tag{11}$$

The conditional PDF of X and Y is

$$f_{X,Y|B}(x,y) = \begin{cases} f_{X,Y}(x,y)/P[B] & (x,y)\in B \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

$$= \begin{cases} Kxy & 40 \le y \le 60,\ 80/y \le x \le 3 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

where $K = (4000P[B])^{-1}$. The conditional expectation of W given event B is

$$E[W|B] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y|B}(x,y)\,dx\,dy \tag{14}$$

$$= \int_{40}^{60}\int_{80/y}^{3} Kx^2y^2\,dx\,dy \tag{15}$$

$$= (K/3)\int_{40}^{60} y^2x^3\Big|_{x=80/y}^{x=3}\,dy \tag{16}$$

$$= (K/3)\int_{40}^{60}\left(27y^2 - 80^3/y\right)\,dy \tag{17}$$

$$= (K/3)\left(9y^3 - 80^3\ln y\right)\Big|_{40}^{60} \approx 120.78 \tag{18}$$

The conditional second moment of K given B is

$$E\left[W^2|B\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(xy)^2 f_{X,Y|B}(x,y)\,dx\,dy \tag{19}$$

$$= \int_{40}^{60}\int_{80/y}^{3} Kx^3y^3\,dx\,dy \tag{20}$$

$$= (K/4)\int_{40}^{60} y^3x^4\Big|_{x=80/y}^{x=3}\,dy \tag{21}$$

$$= (K/4)\int_{40}^{60}\left(81y^3 - 80^4/y\right)\,dy \tag{22}$$

$$= (K/4)\left((81/4)y^4 - 80^4\ln y\right)\Big|_{40}^{60} \approx 16,116.10 \tag{23}$$

It follows that the conditional variance of W given B is

$$\text{Var}[W|B] = E\left[W^2|B\right] - (E[W|B])^2 \approx 1528.30 \tag{24}$$

Quiz 4.9

(A) (1) The joint PMF of A and B can be found from the marginal and conditional PMFs via $P_{A,B}(a,b) = P_{B|A}(b|a)P_A(a)$. Incorporating the information from the given conditional PMFs can be confusing, however. Consequently, we can note that A has range $S_A = \{0, 2\}$ and B has range $S_B = \{0, 1\}$. A table of the joint PMF will include all four possible combinations of A and B. The general form of the table is

$P_{A,B}(a,b)$	$b=0$	$b=1$
$a=0$	$P_{B\|A}(0\|0)P_A(0)$	$P_{B\|A}(1\|0)P_A(0)$
$a=2$	$P_{B\|A}(0\|2)P_A(2)$	$P_{B\|A}(1\|2)P_A(2)$

Substituting values from $P_{B|A}(b|a)$ and $P_A(a)$, we have

$P_{A,B}(a,b)$	$b=0$	$b=1$
$a=0$	$(0.8)(0.4)$	$(0.2)(0.4)$
$a=2$	$(0.5)(0.6)$	$(0.5)(0.6)$

or

$P_{A,B}(a,b)$	$b=0$	$b=1$
$a=0$	0.32	0.08
$a=2$	0.3	0.3

(2) Given the conditional PMF $P_{B|A}(b|2)$, it is easy to calculate the conditional expectation

$$E\left[B|A=2\right] = \sum_{b=0}^{1} b P_{B|A}\left(b|2\right) = (0)(0.5) + (1)(0.5) = 0.5 \qquad (1)$$

(3) From the joint PMF $P_{A,B}(a,b)$, we can calculate the the conditional PMF

$$P_{A|B}\left(a|0\right) = \frac{P_{A,B}\left(a,0\right)}{P_B\left(0\right)} = \begin{cases} 0.32/0.62 & a=0 \\ 0.3/0.62 & a=2 \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

$$= \begin{cases} 16/31 & a=0 \\ 15/31 & a=2 \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

(4) We can calculate the conditional variance $\text{Var}[A|B=0]$ using the conditional PMF $P_{A|B}(a|0)$. First we calculate the conditional expected value

$$E\left[A|B=0\right] = \sum_a a P_{A|B}\left(a|0\right) = 0(16/31) + 2(15/31) = 30/31 \qquad (4)$$

The conditional second moment is

$$E\left[A^2|B=0\right] = \sum_a a^2 P_{A|B}\left(a|0\right) = 0^2(16/31) + 2^2(15/31) = 60/31 \qquad (5)$$

The conditional variance is then

$$\text{Var}[A|B=0] = E\left[A^2|B=0\right] - \left(E\left[A|B=0\right]\right)^2 = \frac{960}{961} \qquad (6)$$

(B) (1) The joint PDF of X and Y is

$$f_{X,Y}\left(x,y\right) = f_{Y|X}\left(y|x\right) f_X\left(x\right) = \begin{cases} 6y & 0 \le y \le x, 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \qquad (7)$$

(2) From the given conditional PDF $f_{Y|X}(y|x)$,

$$f_{Y|X}\left(y|1/2\right) = \begin{cases} 8y & 0 \le y \le 1/2 \\ 0 & \text{otherwise} \end{cases} \qquad (8)$$

(3) The conditional PDF of Y given $X = 1/2$ is $f_{X|Y}(x|1/2) = f_{X,Y}(x, 1/2)/f_Y(1/2)$. To find $f_Y(1/2)$, we integrate the joint PDF.

$$f_Y(1/2) = \int_{-\infty}^{\infty} f_{X,1/2}()\, dx = \int_{1/2}^{1} 6(1/2)\, dx = 3/2 \tag{9}$$

Thus, for $1/2 \le x \le 1$,

$$f_{X|Y}(x|1/2) = \frac{f_{X,Y}(x, 1/2)}{f_Y(1/2)} = \frac{6(1/2)}{3/2} = 2 \tag{10}$$

(4) From the pervious part, we see that given $Y = 1/2$, the conditional PDF of X is uniform $(1/2, 1)$. Thus, by the definition of the uniform (a, b) PDF,

$$\text{Var}[X|Y = 1/2] = \frac{(1 - 1/2)^2}{12} = \frac{1}{48} \tag{11}$$

Quiz 4.10

(A) (1) For random variables X and Y from Example 4.1, we observe that $P_Y(1) = 0.09$ and $P_X(0) = 0.01$. However,

$$P_{X,Y}(0, 1) = 0 \neq P_X(0)\, P_Y(1) \tag{1}$$

Since we have found a pair x, y such that $P_{X,Y}(x, y) \neq P_X(x)P_Y(y)$, we can conclude that X and Y are dependent. Note that whenever $P_{X,Y}(x, y) = 0$, independence requires that either $P_X(x) = 0$ or $P_Y(y) = 0$.

(2) For random variables Q and G from Quiz 4.2, it is not obvious whether they are independent. Unlike X and Y in part (a), there are no obvious pairs q, g that fail the independence requirement. In this case, we calculate the marginal PMFs from the table of the joint PMF $P_{Q,G}(q, g)$ in Quiz 4.2.

$P_{Q,G}(q,g)$	$g = 0$	$g = 1$	$g = 2$	$g = 3$	$P_Q(q)$
$q = 0$	0.06	0.18	0.24	0.12	0.60
$q = 1$	0.04	0.12	0.16	0.08	0.40
$P_G(g)$	0.10	0.30	0.40	0.20	

Careful study of the table will verify that $P_{Q,G}(q, g) = P_Q(q)P_G(g)$ for every pair q, g. Hence Q and G are independent.

(B) (1) Since X_1 and X_2 are independent,

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2) \tag{2}$$

$$= \begin{cases} (1 - x_1/2)(1 - x_2/2) & 0 \le x_1 \le 2, 0 \le x_2 \le 2 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

(2) Let $F_X(x)$ denote the CDF of both X_1 and X_2. The CDF of $Z = \max(X_1, X_2)$ is found by observing that $Z \leq z$ iff $X_1 \leq z$ and $X_2 \leq z$. That is,

$$P[Z \leq z] = P[X_1 \leq z, X_2 \leq z] \tag{4}$$

$$= P[X_1 \leq z]\, P[X_2 \leq z] = [F_X(z)]^2 \tag{5}$$

To complete the problem, we need to find the CDF of each X_i. From the PDF $f_X(x)$, the CDF is

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy = \begin{cases} 0 & x < 0 \\ x - x^2/4 & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases} \tag{6}$$

Thus for $0 \leq z \leq 2$,

$$F_Z(z) = (z - z^2/4)^2 \tag{7}$$

The complete expression for the CDF of Z is

$$F_Z(z) = \begin{cases} 0 & z < 0 \\ (z - z^2/4)^2 & 0 \leq z \leq 2 \\ 1 & z > 1 \end{cases} \tag{8}$$

Quiz 4.11

This problem just requires identifying the various terms in Definition 4.17 and Theorem 4.29. Specifically, from the problem statement, we know that $\rho = 1/2$,

$$\mu_1 = \mu_X = 0, \qquad \mu_2 = \mu_Y = 0, \tag{1}$$

and that

$$\sigma_1 = \sigma_X = 1, \qquad \sigma_2 = \sigma_Y = 1. \tag{2}$$

(1) Applying these facts to Definition 4.17, we have

$$f_{X,Y}(x, y) = \frac{1}{\sqrt{3\pi^2}} e^{-2(x^2 - xy + y^2)/3}. \tag{3}$$

(2) By Theorem 4.30, the conditional expected value and standard deviation of X given $Y = y$ are

$$E[X|Y = y] = y/2 \qquad \tilde{\sigma}_X = \sigma_1^2(1 - \rho^2) = \sqrt{3/4}. \tag{4}$$

When $Y = y = 2$, we see that $E[X|Y = 2] = 1$ and $\text{Var}[X|Y = 2] = 3/4$. The conditional PDF of X given $Y = 2$ is simply the Gaussian PDF

$$f_{X|Y}(x|2) = \frac{1}{\sqrt{3\pi/2}} e^{-2(x-1)^2/3}. \tag{5}$$

Quiz 4.12

One straightforward method is to follow the approach of Example 4.28. Instead, we use an alternate approach. First we observe that X has the discrete uniform $(1, 4)$ PMF. Also, given $X = x$, Y has a discrete uniform $(1, x)$ PMF. That is,

$$P_X(x) = \begin{cases} 1/4 & x = 1, 2, 3, 4, \\ 0 & \text{otherwise}, \end{cases} \qquad P_{Y|X}(y|x) = \begin{cases} 1/x & y = 1, \ldots, x \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Given $X = x$, and an independent uniform $(0, 1)$ random variable U, we can generate a sample value of Y with a discrete uniform $(1, x)$ PMF via $Y = \lceil xU \rceil$. This observation prompts the following program:

```
function xy=dtrianglerv(m)
sx=[1;2;3;4];
px=0.25*ones(4,1);
x=finiterv(sx,px,m);
y=ceil(x.*rand(m,1));
xy=[x';y'];
```

Quiz Solutions – Chapter 5

Quiz 5.1

We find $P[C]$ by integrating the joint PDF over the region of interest. Specifically,

$$P[C] = \int_0^{1/2} dy_2 \int_0^{y_2} dy_1 \int_0^{1/2} dy_4 \int_0^{y_4} 4dy_3 \quad (1)$$

$$= 4\left(\int_0^{1/2} y_2\, dy_2\right)\left(\int_0^{1/2} y_4\, dy_4\right) = 1/4. \quad (2)$$

Quiz 5.2

By definition of $\mathbf{A}$, $Y_1 = X_1$, $Y_2 = X_2 - X_1$ and $Y_3 = X_3 - X_2$. Since $0 < X_1 < X_2 < X_3$, each Y_i must be a strictly positive integer. Thus, for $y_1, y_2, y_3 \in \{1, 2, \ldots\}$,

$$P_{\mathbf{Y}}(\mathbf{y}) = P[Y_1 = y_1, Y_2 = y_2, Y_3 = y_3] \quad (1)$$

$$= P[X_1 = y_1, X_2 - X_1 = y_2, X_3 - X_2 = y_3] \quad (2)$$

$$= P[X_1 = y_1, X_2 = y_2 + y_1, X_3 = y_3 + y_2 + y_1] \quad (3)$$

$$= (1-p)^3 p^{y_1+y_2+y_3} \quad (4)$$

By defining the vector $\mathbf{a} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}'$, the complete expression for the joint PMF of $\mathbf{Y}$ is

$$P_{\mathbf{Y}}(\mathbf{y}) = \begin{cases} (1-p)p^{\mathbf{a}'\mathbf{y}} & y_1, y_2, y_3 \in \{1, 2, \ldots\} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Quiz 5.3

First we note that each marginal PDF is nonzero only if any subset of the x_i obeys the ordering contraints $0 \leq x_1 \leq x_2 \leq x_3 \leq 1$. Within these constraints, we have

$$f_{X_1,X_2}(x_1, x_2) = \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\, dx_3 = \int_{x_2}^{1} 6\, dx_3 = 6(1 - x_2), \quad (1)$$

$$f_{X_2,X_3}(x_2, x_3) = \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\, dx_1 = \int_{0}^{x_2} 6\, dx_1 = 6x_2, \quad (2)$$

$$f_{X_1,X_3}(x_1, x_3) = \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})\, dx_2 = \int_{x_1}^{x_3} 6\, dx_2 = 6(x_3 - x_1). \quad (3)$$

In particular, we must keep in mind that $f_{X_1,X_2}(x_1, x_2) = 0$ unless $0 \leq x_1 \leq x_2 \leq 1$, $f_{X_2,X_3}(x_2, x_3) = 0$ unless $0 \leq x_2 \leq x_3 \leq 1$, and that $f_{X_1,X_3}(x_1, x_3) = 0$ unless $0 \leq x_1 \leq$

$x_3 \leq 1$. The complete expressions are

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} 6(1-x_2) & 0 \leq x_1 \leq x_2 \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$f_{X_2,X_3}(x_2,x_3) = \begin{cases} 6x_2 & 0 \leq x_2 \leq x_3 \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$$f_{X_1,X_3}(x_1,x_3) = \begin{cases} 6(x_3-x_1) & 0 \leq x_1 \leq x_3 \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Now we can find the marginal PDFs. When $0 \leq x_i \leq 1$ for each x_i,

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\, dx_2 = \int_{x_1}^{1} 6(1-x_2)\, dx_2 = 3(1-x_1)^2 \tag{7}$$

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_2,X_3}(x_2,x_3)\, dx_3 = \int_{x_2}^{1} 6x_2\, dx_3 = 6x_2(1-x_2) \tag{8}$$

$$f_{X_3}(x_3) = \int_{-\infty}^{\infty} f_{X_2,X_3}(x_2,x_3)\, dx_2 = \int_{0}^{x_3} 6x_2\, dx_2 = 3x_3^2 \tag{9}$$

The complete expressions are

$$f_{X_1}(x_1) = \begin{cases} 3(1-x_1)^2 & 0 \leq x_1 \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

$$f_{X_2}(x_2) = \begin{cases} 6x_2(1-x_2) & 0 \leq x_2 \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

$$f_{X_3}(x_3) = \begin{cases} 3x_3^2 & 0 \leq x_3 \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

Quiz 5.4

In the PDF $f_{\mathbf{Y}}(\mathbf{y})$, the components have dependencies as a result of the ordering constraints $Y_1 \leq Y_2$ and $Y_3 \leq Y_4$. We can separate these constraints by creating the vectors

$$\mathbf{V} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}, \qquad \mathbf{W} = \begin{bmatrix} Y_3 \\ Y_4 \end{bmatrix}. \tag{1}$$

The joint PDF of **V** and **W** is

$$f_{\mathbf{V},\mathbf{W}}(\mathbf{v},\mathbf{w}) = \begin{cases} 4 & 0 \leq v_1 \leq v_2 \leq 1, 0 \leq w_1 \leq w_2 \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

We must verify that $\mathbf{V}$ and $\mathbf{W}$ are independent. For $0 \leq v_1 \leq v_2 \leq 1$,

$$f_{\mathbf{V}}(\mathbf{v}) = \iint f_{\mathbf{V},\mathbf{W}}(\mathbf{v}, \mathbf{w})\, dw_1\, dw_2 \tag{3}$$

$$= \int_0^1 \left(\int_{w_1}^1 4\, dw_2 \right) dw_1 \tag{4}$$

$$= \int_0^1 4(1 - w_1)\, dw_1 = 2 \tag{5}$$

Similarly, for $0 \leq w_1 \leq w_2 \leq 1$,

$$f_{\mathbf{W}}(\mathbf{w}) = \iint f_{\mathbf{V},\mathbf{W}}(\mathbf{v}, \mathbf{w})\, dv_1\, dv_2 \tag{6}$$

$$= \int_0^1 \left(\int_{v_1}^1 4\, dv_2 \right) dv_1 = 2 \tag{7}$$

It follows that $\mathbf{V}$ and $\mathbf{W}$ have PDFs

$$f_{\mathbf{V}}(\mathbf{v}) = \begin{cases} 2 & 0 \leq v_1 \leq v_2 \leq 1 \\ 0 & \text{otherwise} \end{cases}, \qquad f_{\mathbf{W}}(\mathbf{w}) = \begin{cases} 2 & 0 \leq w_1 \leq w_2 \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

It is easy to verify that $f_{\mathbf{V},\mathbf{W}}(\mathbf{v}, \mathbf{w}) = f_{\mathbf{V}}(\mathbf{v})\, f_{\mathbf{W}}(\mathbf{w})$, confirming that $\mathbf{V}$ and $\mathbf{W}$ are independent vectors.

Quiz 5.5

(A) Referring to Theorem 1.19, each test is a subexperiment with three possible outcomes: L, A and R. In five trials, the vector $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}'$ indicating the number of outcomes of each subexperiment has the multinomial PMF

$$P_{\mathbf{X}}(\mathbf{x}) = \begin{cases} \binom{5}{x_1, x_2, x_3} (0.3)^{x_1} (0.6)^{x_2} (0.1)^{x_3} & x_1 + x_2 + x_3 = 5; \\ & x_1, x_2, x_3 \in \{0, 1, \ldots, 5\} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

We can find the marginal PMF for each X_i from the joint PMF $P_{\mathbf{X}}(\mathbf{x})$; however it is simpler to just start from first principles and observe that X_1 is the number of occurrences of L in five independent tests. If we view each test as a trial with success probability $P[L] = 0.3$, we see that X_1 is a binomial $(n, p) = (5, 0.3)$ random variable. Similarly, X_2 is a binomial $(5, 0.6)$ random variable and X_3 is a binomial $(5, 0.1)$ random variable. That is, for $p_1 = 0.3$, $p_2 = 0.6$ and $p_3 = 0.1$,

$$P_{X_i}(x) = \begin{cases} \binom{5}{x} p_i^x (1 - p_i)^{5-x} & x = 0, 1, \ldots, 5 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

From the marginal PMFs, we see that X_1, X_2 and X_3 are not independent. Hence, we must use Theorem 5.6 to find the PMF of W. In particular, since $X_1 + X_2 + X_3 = 5$ and since each X_i is non-negative, $P_W(0) = P_W(1) = 0$. Furthermore,

$$\begin{aligned} P_W(2) &= P_{\mathbf{X}}(1,2,2) + P_{\mathbf{X}}(2,1,2) + P_{\mathbf{X}}(2,2,1) && (3)\\ &= \frac{5![0.3(0.6)^2(0.1)^2 + 0.3^2(0.6)(0.1)^2 + 0.3^2(0.6)^2(0.1)]}{2!2!1!} && (4)\\ &= 0.1458 && (5)\end{aligned}$$

In addition, for $w = 3$, $w = 4$, and $w = 5$, the event $W = w$ occurs if and only if one of the mutually exclusive events $X_1 = w$, $X_2 = w$, or $X_3 = w$ occurs. Thus,

$$\begin{aligned} P_W(3) &= P_{X_1}(3) + P_{X_2}(3) + P_{X_3}(3) = 0.486 && (6)\\ P_W(4) &= P_{X_1}(4) + P_{X_2}(4) + P_{X_3}(4) = 0.288 && (7)\\ P_W(5) &= P_{X_1}(5) + P_{X_2}(5) + P_{X_3}(5) = 0.0802 && (8)\end{aligned}$$

(B) Since each $Y_i = 2X_i + 4$, we can apply Theorem 5.10 to write

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= \frac{1}{2^3} f_{\mathbf{X}}\left(\frac{y_1-4}{2}, \frac{y_2-4}{2}, \frac{y_3-4}{2}\right) && (9)\\ &= \begin{cases} (1/8)e^{-(y_3-4)/2} & 4 \le y_1 \le y_2 \le y_3 \\ 0 & \text{otherwise} \end{cases} && (10)\end{aligned}$$

Note that for other matrices $\mathbf{A}$, the constraints on $\mathbf{y}$ resulting from the constraints $0 \le X_1 \le X_2 \le X_3$ can be much more complicated.

Quiz 5.6

We start by finding the components $E[X_i] = \int_{-\infty}^{\infty} x f_{X_i}(x)\,dx$ of $\boldsymbol{\mu}_X$. To do so, we use the marginal PDFs $f_{X_i}(x)$ found in Quiz 5.3:

$$E[X_1] = \int_0^1 3x(1-x)^2\,dx = 1/4, \qquad (1)$$

$$E[X_2] = \int_0^1 6x^2(1-x)\,dx = 1/2, \qquad (2)$$

$$E[X_3] = \int_0^1 3x^3\,dx = 3/4. \qquad (3)$$

To find the correlation matrix $\mathbf{R}_X$, we need to find $E[X_iX_j]$ for all i and j. We start with the second moments:

$$E\left[X_1^2\right] = \int_0^1 3x^2(1-x)^2\,dx = 1/10. \tag{4}$$

$$E\left[X_2^2\right] = \int_0^1 6x^3(1-x)\,dx = 3/10. \tag{5}$$

$$E\left[X_3^2\right] = \int_0^1 3x^4\,dx = 3/5. \tag{6}$$

Using marginal PDFs from Quiz 5.3, the cross terms are

$$E\left[X_1X_2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2\,f_{X_1,X_2}\left(x_1,x_2\right),\ dx_1\,dx_2 \tag{7}$$

$$= \int_0^1\left(\int_{x_1}^1 6x_1x_2(1-x_2)\,dx_2\right)dx_1 \tag{8}$$

$$= \int_0^1 [x_1 - 3x_1^3 + 2x_1^4]\,dx_1 = 3/20. \tag{9}$$

$$E\left[X_2X_3\right] = \int_0^1\int_{x_2}^1 6x_2^2x_3\,dx_3\,dx_2 \tag{10}$$

$$= \int_0^1 [3x_2^2 - 3x_2^4]\,dx_2 = 2/5 \tag{11}$$

$$E\left[X_1X_3\right] = \int_0^1\int_{x_1}^1 6x_1x_3(x_3-x_1)\,dx_3\,dx_1. \tag{12}$$

$$= \int_0^1\left(\left.(2x_1x_3^3 - 3x_1^2x_3^2)\right|_{x_3=x_1}^{x_3=1}\right)dx_1 \tag{13}$$

$$= \int_0^1 [2x_1 - 3x_1^2 + x_1^4]\,dx_1 = 1/5. \tag{14}$$

Summarizing the results, **X** has correlation matrix

$$\mathbf{R}_X = \begin{bmatrix} 1/10 & 3/20 & 1/5 \\ 3/20 & 3/10 & 2/5 \\ 1/5 & 2/5 & 3/5 \end{bmatrix}. \tag{15}$$

Vector **X** has covariance matrix

$$\mathbf{C}_X = \mathbf{R}_X - E\left[\mathbf{X}\right]E\left[\mathbf{X}\right]' \tag{16}$$

$$= \begin{bmatrix} 1/10 & 3/20 & 1/5 \\ 3/20 & 3/10 & 2/5 \\ 1/5 & 2/5 & 3/5 \end{bmatrix} - \begin{bmatrix} 1/4 \\ 1/2 \\ 3/4 \end{bmatrix}\begin{bmatrix} 1/4 & 1/2 & 3/4 \end{bmatrix} \tag{17}$$

$$= \begin{bmatrix} 1/10 & 3/20 & 1/5 \\ 3/20 & 3/10 & 2/5 \\ 1/5 & 2/5 & 3/5 \end{bmatrix} - \begin{bmatrix} 1/16 & 1/8 & 3/16 \\ 1/8 & 1/4 & 3/8 \\ 3/16 & 3/8 & 9/16 \end{bmatrix} = \frac{1}{80}\begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}. \tag{18}$$

This problem shows that even for fairly simple joint PDFs, computing the covariance matrix by calculus can be a time consuming task.

Quiz 5.7

We observe that $\mathbf{X} = \mathbf{A}\mathbf{Z} + \mathbf{b}$ where

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}. \tag{1}$$

It follows from Theorem 5.18 that $\boldsymbol{\mu}_X = \mathbf{b}$ and that

$$\mathbf{C}_X = \mathbf{A}\mathbf{A}' = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 1 & 2 \end{bmatrix}. \tag{2}$$

Quiz 5.8

First, we observe that $Y = \mathbf{A}\mathbf{T}$ where $\mathbf{A} = \begin{bmatrix} 1/31 & 1/31 & \cdots & 1/31 \end{bmatrix}'$. Since $\mathbf{T}$ is a Gaussian random vector, Theorem 5.16 tells us that Y is a 1 dimensional Gaussian vector, i.e., just a Gaussian random variable. The expected value of Y is $\mu_Y = \mu_T = 80$. The covariance matrix of Y is 1×1 and is just equal to $\text{Var}[Y]$. Thus, by Theorem 5.16, $\text{Var}[Y] = \mathbf{A}\mathbf{C}_T\mathbf{A}'$.

```
function p=julytemps(T);
[D1 D2]=ndgrid((1:31),(1:31));
CT=36./(1+abs(D1-D2));
A=ones(31,1)/31.0;
CY=(A')*CT*A;
p=phi((T-80)/sqrt(CY));
```

In `julytemps.m`, the first two lines generate the 31×31 covariance matrix `CT`, or $\mathbf{C}_T$. Next we calculate $\text{Var}[Y]$. The final step is to use the $\Phi(\cdot)$ function to calculate $P[Y < T]$.

Here is the output of `julytemps.m`:

```
>> julytemps([70 75 80 85 90 95])
ans =
    0.0000    0.0221    0.5000    0.9779    1.0000    1.0000
```

Note that $P[T \leq 70]$ is not actually zero and that $P[T \leq 90]$ is not actually 1.0000. Its just that the MATLAB's short format output, invoked with the command `format short`, rounds off those probabilities. Here is the long format output:

```
>> format long
>> julytemps([70 75 80 85 90 95])
ans =
 Columns 1 through 4
  0.00002844263128  0.02207383067604  0.50000000000000  0.97792616932396
 Columns 5 through 6
  0.99997155736872  0.99999999922010
```

The `ndgrid` function is a useful to way calculate many covariance matrices. However, in this problem, $\mathbf{C}_X$ has a special structure; the i, jth element is

$$C_{\mathbf{T}}(i, j) = c_{|i-j|} = \frac{36}{1 + |i - j|}. \tag{1}$$

If we write out the elements of the covariance matrix, we see that

$$\mathbf{C}_{\mathbf{T}} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{30} \\ c_1 & c_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_1 \\ c_{30} & \cdots & c_1 & c_0 \end{bmatrix}. \tag{2}$$

This covariance matrix is known as a symmetric Toeplitz matrix. We will see in Chapters 9 and 11 that Toeplitz covariance matrices are quite common. In fact, MATLAB has a `toeplitz` function for generating them. The function `julytemps2` use the `toeplitz` to generate the correlation matrix $\mathbf{C}_{\mathbf{T}}$.

```
function p=julytemps2(T);
c=36./(1+abs(0:30));
CT=toeplitz(c);
A=ones(31,1)/31.0;
CY=(A')*CT*A;
p=phi((T-80)/sqrt(CY));
```

Quiz Solutions – Chapter 6

Quiz 6.1

Let $K_1, \ldots, K_n$ denote a sequence of iid random variables each with PMF

$$P_K(k) = \begin{cases} 1/4 & k = 1, \ldots, 4 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

We can write W_n in the form of $W_n = K_1 + \cdots + K_n$. First, we note that the first two moments of K_i are

$$E[K_i] = (1 + 2 + 3 + 4)/4 = 2.5 \tag{2}$$

$$E\left[K_i^2\right] = (1^2 + 2^2 + 3^2 + 4^2)/4 = 7.5 \tag{3}$$

Thus the variance of K_i is

$$\text{Var}[K_i] = E\left[K_i^2\right] - (E[K_i])^2 = 7.5 - (2.5)^2 = 1.25 \tag{4}$$

Since $E[K_i] = 2.5$, the expected value of W_n is

$$E[W_n] = E[K_1] + \cdots + E[K_n] = nE[K_i] = 2.5n \tag{5}$$

Since the rolls are independent, the random variables $K_1, \ldots, K_n$ are independent. Hence, by Theorem 6.3, the variance of the sum equals the sum of the variances. That is,

$$\text{Var}[W_n] = \text{Var}[K_1] + \cdots + \text{Var}[K_n] = 1.25n \tag{6}$$

Quiz 6.2

Random variables X and Y have PDFs

$$f_X(x) = \begin{cases} 3e^{-3x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad f_Y(y) = \begin{cases} 2e^{-2y} & y \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Since X and Y are nonnegative, $W = X + Y$ is nonnegative. By Theorem 6.5, the PDF of $W = X + Y$ is

$$f_W(w) = \int_{-\infty}^{\infty} f_X(w - y)\, f_Y(y)\, dy = 6 \int_0^w e^{-3(w-y)} e^{-2y}\, dy \tag{2}$$

Fortunately, this integral is easy to evaluate. For $w > 0$,

$$f_W(w) = e^{-3w} e^y \Big|_0^w = 6\left(e^{-2w} - e^{-3w}\right) \tag{3}$$

Since $f_W(w) = 0$ for $w < 0$, a conmplete expression for the PDF of W is

$$f_W(w) = \begin{cases} 6e^{-2w}\left(1 - e^{-w}\right) & w \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Quiz 6.3

The MGF of K is

$$\phi_K(s) = E\left[e^{sK}\right] == \sum_{k=0}^{4}(0.2)e^{sk} = 0.2\left(1 + e^s + e^{2s} + e^{3s} + e^{4s}\right) \quad (1)$$

We find the moments by taking derivatives. The first derivative of $\phi_K(s)$ is

$$\frac{d\phi_K(s)}{ds} = 0.2(e^s + 2e^{2s} + 3e^{3s} + 4e^{4s}) \quad (2)$$

Evaluating the derivative at $s = 0$ yields

$$E[K] = \left.\frac{d\phi_K(s)}{ds}\right|_{s=0} = 0.2(1 + 2 + 3 + 4) = 2 \quad (3)$$

To find higher-order moments, we continue to take derivatives:

$$E\left[K^2\right] = \left.\frac{d^2\phi_K(s)}{ds^2}\right|_{s=0} = 0.2(e^s + 4e^{2s} + 9e^{3s} + 16e^{4s})\Big|_{s=0} = 6 \quad (4)$$

$$E\left[K^3\right] = \left.\frac{d^3\phi_K(s)}{ds^3}\right|_{s=0} = 0.2(e^s + 8e^{2s} + 27e^{3s} + 64e^{4s})\Big|_{s=0} = 20 \quad (5)$$

$$E\left[K^4\right] = \left.\frac{d^4\phi_K(s)}{ds^4}\right|_{s=0} = 0.2(e^s + 16e^{2s} + 81e^{3s} + 256e^{4s})\Big|_{s=0} = 70.8 \quad (6)$$

$$(7)$$

Quiz 6.4

(A) Each K_i has MGF

$$\phi_K(s) = E\left[e^{sK_i}\right] = \frac{e^s + e^{2s} + \cdots + e^{ns}}{n} = \frac{e^s(1 - e^{ns})}{n(1 - e^s)} \quad (1)$$

Since the sequence of K_i is independent, Theorem 6.8 says the MGF of J is

$$\phi_J(s) = (\phi_K(s))^m = \frac{e^{ms}(1 - e^{ns})^m}{n^m(1 - e^s)^m} \quad (2)$$

(B) Since the set of $\alpha^j X_j$ are independent Gaussian random variables, Theorem 6.10 says that W is a Gaussian random variable. Thus to find the PDF of W, we need only find the expected value and variance. Since the expectation of the sum equals the sum of the expectations:

$$E[W] = \alpha E[X_1] + \alpha^2 E[X_2] + \cdots + \alpha^n E[X_n] = 0 \quad (3)$$

Since the $\alpha^j X_j$ are independent, the variance of the sum equals the sum of the variances:

$$\begin{aligned} \text{Var}[W] &= \alpha^2 \text{Var}[X_1] + \alpha^4 \text{Var}[X_2] + \cdots + \alpha^{2n} \text{Var}[X_n] && (4) \\ &= \alpha^2 + 2(\alpha^2)^2 + 3(\alpha^2)^3 + \cdots + n(\alpha^2)^n && (5) \end{aligned}$$

Defining $q = \alpha^2$, we can use Math Fact B.6 to write

$$\text{Var}[W] = \frac{\alpha^2 - \alpha^{2n+2}[1 + n(1 - \alpha^2)]}{(1 - \alpha^2)^2} \tag{6}$$

With $E[W] = 0$ and $\sigma_W^2 = \text{Var}[W]$, we can write the PDF of W as

$$f_W(w) = \frac{1}{\sqrt{2\pi\sigma_W^2}} e^{-w^2/2\sigma_W^2} \tag{7}$$

Quiz 6.5

(1) From Table 6.1, each X_i has MGF $\phi_X(s)$ and random variable N has MGF $\phi_N(s)$ where

$$\phi_X(s) = \frac{1}{1 - s}, \qquad \phi_N(s) = \frac{\frac{1}{5}e^s}{1 - \frac{4}{5}e^s}. \tag{1}$$

From Theorem 6.12, R has MGF

$$\phi_R(s) = \phi_N(\ln \phi_X(s)) = \frac{\frac{1}{5}\phi_X(s)}{1 - \frac{4}{5}\phi_X(s)} \tag{2}$$

Substituting the expression for $\phi_X(s)$ yields

$$\phi_R(s) = \frac{\frac{1}{5}}{\frac{1}{5} - s}. \tag{3}$$

(2) From Table 6.1, we see that R has the MGF of an exponential (1/5) random variable. The corresponding PDF is

$$f_R(r) = \begin{cases} (1/5)e^{-r/5} & r \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

This quiz is an example of the general result that a geometric sum of exponential random variables is an exponential random variable.

Quiz 6.6

(1) The expected access time is

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^{12} \frac{x}{12}\, dx = 6 \text{ msec} \tag{1}$$

(2) The second moment of the access time is

$$E\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \int_0^{12} \frac{x^2}{12}\, dx = 48 \tag{2}$$

The variance of the access time is $\text{Var}[X] = E[X^2] - (E[X])^2 = 48 - 36 = 12$.

(3) Using X_i to denote the access time of block i, we can write

$$A = X_1 + X_2 + \cdots + X_{12} \tag{3}$$

Since the expectation of the sum equals the sum of the expectations,

$$E[A] = E[X_1] + \cdots + E[X_{12}] = 12E[X] = 72 \text{ msec} \tag{4}$$

(4) Since the X_i are independent,

$$\text{Var}[A] = \text{Var}[X_1] + \cdots + \text{Var}[X_{12}] = 12\,\text{Var}[X] = 144 \tag{5}$$

Hence, the standard deviation of A is $\sigma_A = 12$

(5) To use the central limit theorem, we write

$$\begin{aligned}
P[A > 75] &= 1 - P[A \leq 75] && (6)\\
&= 1 - P\left[\frac{A - E[A]}{\sigma_A} \leq \frac{75 - E[A]}{\sigma_A}\right] && (7)\\
&\approx 1 - \Phi\left(\frac{75-72}{12}\right) && (8)\\
&= 1 - 0.5987 = 0.4013 && (9)
\end{aligned}$$

Note that we used Table 3.1 to look up $\Phi(0.25)$.

(6) Once again, we use the central limit theorem and Table 3.1 to estimate

$$\begin{aligned}
P[A < 48] &= P\left[\frac{A - E[A]}{\sigma_A} < \frac{48 - E[A]}{\sigma_A}\right] && (10)\\
&\approx \Phi\left(\frac{48-72}{12}\right) && (11)\\
&= 1 - \Phi(2) = 1 - 0.9773 = 0.0227 && (12)
\end{aligned}$$

Quiz 6.7

Random variable K_n has a binomial distribution for n trials and success probability $P[V] = 3/4$.

(1) The expected number of voice calls out of 48 calls is $E[K_{48}] = 48P[V] = 36$.

(2) The variance of K_{48} is

$$\text{Var}[K_{48}] = 48P\,[V]\,(1 - P\,[V]) = 48(3/4)(1/4) = 9 \tag{1}$$

Thus K_{48} has standard deviation $\sigma_{K_{48}} = 3$.

(3) Using the ordinary central limit theorem and Table 3.1 yields

$$P\,[30 \le K_{48} \le 42] \approx \Phi\left(\frac{42-36}{3}\right) - \Phi\left(\frac{30-36}{3}\right) = \Phi(2) - \Phi(-2) \tag{2}$$

Recalling that $\Phi(-x) = 1 - \Phi(x)$, we have

$$P\,[30 \le K_{48} \le 42] \approx 2\Phi(2) - 1 = 0.9545 \tag{3}$$

(4) Since K_{48} is a discrete random variable, we can use the De Moivre-Laplace approximation to estimate

$$P\,[30 \le K_{48} \le 42] \approx \Phi\left(\frac{42+0.5-36}{3}\right) - \Phi\left(\frac{30-0.5-36}{3}\right) \tag{4}$$

$$= 2\Phi(2.16666) - 1 = 0.9687 \tag{5}$$

Quiz 6.8

The train interarrival times X_1, X_2, X_3 are iid exponential (λ) random variables. The arrival time of the third train is

$$W = X_1 + X_2 + X_3. \tag{1}$$

In Theorem 6.11, we found that the sum of three iid exponential (λ) random variables is an Erlang ($n = 3, \lambda$) random variable. From Appendix A, we find that W has expected value and variance

$$E\,[W] = 3/\lambda = 6 \qquad \text{Var}[W] = 3/\lambda^2 = 12 \tag{2}$$

(1) By the Central Limit Theorem,

$$P\,[W > 20] = P\left[\frac{W-6}{\sqrt{12}} > \frac{20-6}{\sqrt{12}}\right] \approx Q(7/\sqrt{3}) = 2.66 \times 10^{-5} \tag{3}$$

(2) To use the Chernoff bound, we note that the MGF of W is

$$\phi_W(s) = \left(\frac{\lambda}{\lambda - s}\right)^3 = \frac{1}{(1-2s)^3} \tag{4}$$

The Chernoff bound states that

$$P[W > 20] \le \min_{s\ge 0} e^{-20s}\phi_X(s) = \min_{s\ge 0} \frac{e^{-20s}}{(1-2s)^3} \tag{5}$$

To minimize $h(s) = e^{-20s}/(1-2s)^3$, we set the derivative of $h(s)$ to zero:

$$\frac{dh(s)}{ds} = \frac{-20(1-2s)^3 e^{-20s} + 6e^{-20s}(1-2s)^2}{(1-2s)^6} = 0 \tag{6}$$

This implies $20(1-2s) = 6$ or $s = 7/20$. Applying $s = 7/20$ into the Chernoff bound yields

$$P[W > 20] \le \left.\frac{e^{-20s}}{(1-2s)^3}\right|_{s=7/20} = (10/3)^3 e^{-7} = 0.0338 \tag{7}$$

(3) Theorem 3.11 says that for any $w > 0$, the CDF of the Erlang $(\lambda, 3)$ random variable W satisfies

$$F_W(w) = 1 - \sum_{k=0}^{2} \frac{(\lambda w)^k e^{-\lambda w}}{k!} \tag{8}$$

Equivalently, for $\lambda = 1/2$ and $w = 20$,

$$P[W > 20] = 1 - F_W(20) \tag{9}$$

$$= e^{-10}\left(1 + \frac{10}{1!} + \frac{10^2}{2!}\right) = 61e^{-10} = 0.0028 \tag{10}$$

Although the Chernoff bound is relatively weak in that it overestimates the probability by roughly a factor of 12, it is a valid bound. By contrast, the Central Limit Theorem approximation grossly underestimates the true probability.

Quiz 6.9

One solution to this problem is to follow the approach of Example 6.19:

```
%unifbinom100.m
sx=0:100;sy=0:100;
px=binomialpmf(100,0.5,sx); py=duniformpmf(0,100,sy);
[SX,SY]=ndgrid(sx,sy); [PX,PY]=ndgrid(px,py);
SW=SX+SY; PW=PX.*PY;
sw=unique(SW); pw=finitepmf(SW,PW,sw);
pmfplot(sw,pw,'\itw','\itP_W(w)');
```

A graph of the PMF $P_W(w)$ appears in Figure 2 With some thought, it should be apparent that the `finitepmf` function is implementing the convolution of the two PMFs.

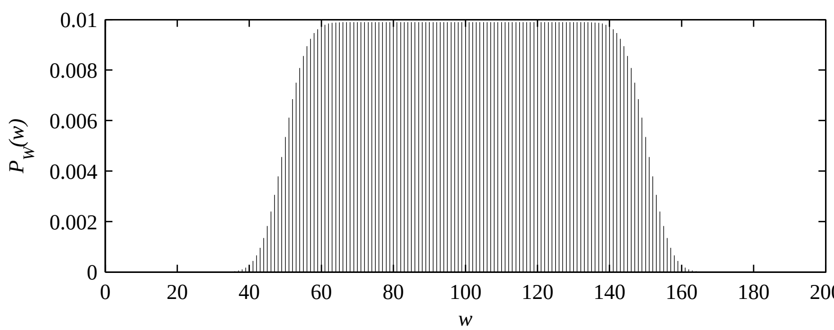


Figure 2: From Quiz 6.9, the PMF $P_W(w)$ of the independent sum of a binomial (100, 0.5) random variable and a discrete uniform (0, 100) random variable.

Quiz Solutions – Chapter 7

Quiz 7.1

An exponential random variable with expected value 1 also has variance 1. By Theorem 7.1, $M_n(X)$ has variance $\text{Var}[M_n(X)] = 1/n$. Hence, we need $n = 100$ samples.

Quiz 7.2

The arrival time of the third elevator is $W = X_1 + X_2 + X_3$. Since each X_i is uniform $(0, 30)$,

$$E[X_i] = 15, \qquad \text{Var}[X_i] = \frac{(30-0)^2}{12} = 75. \tag{1}$$

Thus $E[W] = 3E[X_i] = 45$, and $\text{Var}[W] = 3\,\text{Var}[X_i] = 225$.

(1) By the Markov inequality,

$$P[W > 75] \leq \frac{E[W]}{75} = \frac{45}{75} = \frac{3}{5} \tag{2}$$

(2) By the Chebyshev inequality,

$$P[W > 75] = P[W - E[W] > 30] \tag{3}$$

$$\leq P[|W - E[W]| > 30] \leq \frac{\text{Var}[W]}{30^2} = \frac{225}{900} = \frac{1}{4} \tag{4}$$

Quiz 7.3

Define the random variable $W = (X - \mu_X)^2$. Observe that $V_{100}(X) = M_{100}(W)$. By Theorem 7.6, the mean square error is

$$E\left[(M_{100}(W) - \mu_W)^2\right] = \frac{\text{Var}[W]}{100} \tag{1}$$

Observe that $\mu_X = 0$ so that $W = X^2$. Thus,

$$\mu_W = E\left[X^2\right] = \int_{-1}^{1} x^2 f_X(x)\, dx = 1/3 \tag{2}$$

$$E\left[W^2\right] = E\left[X^4\right] = \int_{-1}^{1} x^4 f_X(x)\, dx = 1/5 \tag{3}$$

Therefore $\text{Var}[W] = E[W^2] - \mu_W^2 = 1/5 - (1/3)^2 = 4/45$ and the mean square error is $4/4500 = 0.000889$.

Quiz 7.4

Assuming the number n of samples is large, we can use a Gaussian approximation for $M_n(X)$. Since$E[X] = p$ and $\text{Var}[X] = p(1-p)$, we apply Theorem 7.13 which says that the interval estimate

$$M_n(X) - c \leq p \leq M_n(X) + c \tag{1}$$

has confidence coefficient $1-\alpha$ where

$$\alpha = 2 - 2\Phi\left(\frac{c\sqrt{n}}{p(1-p)}\right). \tag{2}$$

We must ensure for every value of p that $1-\alpha \geq 0.9$ or $\alpha \leq 0.1$. Equivalently, we must have

$$\Phi\left(\frac{c\sqrt{n}}{p(1-p)}\right) \geq 0.95 \tag{3}$$

for every value of p. Since $\Phi(x)$ is an increasing function of x, we must satisfy $c\sqrt{n} \geq 1.65p(1-p)$. Since $p(1-p) \leq 1/4$ for all p, we require that

$$c \geq \frac{1.65}{4\sqrt{n}} = \frac{0.41}{\sqrt{n}}. \tag{4}$$

The 0.9 confidence interval estimate of p is

$$M_n(X) - \frac{0.41}{\sqrt{n}} \leq p \leq M_n(X) + \frac{0.41}{\sqrt{n}}. \tag{5}$$

For the 0.99 confidence interval, we have $\alpha \leq 0.01$, implying $\Phi(c\sqrt{n}/(p(1-p))) \geq 0.995$. This implies $c\sqrt{n} \geq 2.58p(1-p)$. Since $p(1-p) \leq 1/4$ for all p, we require that $c \geq (0.25)(2.58)/\sqrt{n}$. In this case, the 0.99 confidence interval estimate is

$$M_n(X) - \frac{0.645}{\sqrt{n}} \leq p \leq M_n(X) + \frac{0.645}{\sqrt{n}}. \tag{6}$$

Note that if $M_{100}(X) = 0.4$, then the 0.99 confidence interval estimate is

$$0.3355 \leq p \leq 0.4645. \tag{7}$$

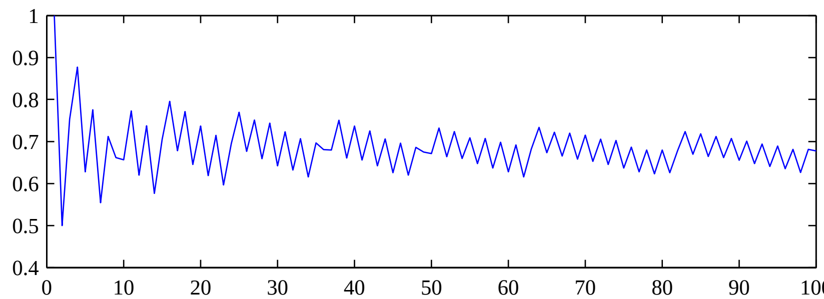
The interval is wide because the 0.99 confidence is high.

Quiz 7.5

Following the approach of `bernoullitraces.m`, we generate $m = 1000$ sample paths, each sample path having $n = 100$ Bernoulli traces. at time k, `OK(k)` counts the fraction of sample paths that have sample mean within one standard error of p. The program `bernoullisample.m` generates graphs the number of traces within one standard error as a function of the time, i.e. the number of trials in each trace.

```
function OK=bernoullisample(n,m,p);
x=reshape(bernoullirv(p,m*n),n,m);
nn=(1:n)'*ones(1,m);
MN=cumsum(x)./nn;
stderr=sqrt(p*(1-p))./sqrt((1:n)');
stderrmat=stderr*ones(1,m);
OK=sum(abs(MN-p)<stderrmat,2)/m;
plot(1:n,OK,'-s');
```

The following graph was generated by `bernoullisample(100,5000,0.5)`:



As we would expect, as m gets large, the fraction of traces within one standard error approaches $2\Phi(1) - 1 \approx 0.68$. The unusual sawtooth pattern, though perhaps unexpected, is examined in Problem 7.5.2.

Quiz Solutions – Chapter 8

Quiz 8.1

From the problem statement, each X_i has PDF and CDF

$$f_{X_i}(x) = \begin{cases} e^{-x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad F_{X_i}(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-x} & x \geq 0 \end{cases} \tag{1}$$

Hence, the CDF of the maximum of $X_1, \ldots, X_{15}$ obeys

$$F_X(x) = P[X \leq x] = P[X_1 \leq x, X_2 \leq x, \cdots, X_{15} \leq x] = [P[X_i \leq x]]^{15}. \tag{2}$$

This implies that for $x \geq 0$,

$$F_X(x) = \left[F_{X_i}(x)\right]^{15} = \left[1 - e^{-x}\right]^{15} \tag{3}$$

To design a significance test, we must choose a rejection region for X. A reasonable choice is to reject the hypothesis if X is too small. That is, let $R = \{X \leq r\}$. For a significance level of $\alpha = 0.01$, we obtain

$$\alpha = P[X \leq r] = (1 - e^{-r})^{15} = 0.01 \tag{4}$$

It is straightforward to show that

$$r = -\ln\left[1 - (0.01)^{1/15}\right] = 1.33 \tag{5}$$

Hence, if we observe $X < 1.33$, then we reject the hypothesis.

Quiz 8.2

From the problem statement, the conditional PMFs of K are

$$P_{K|H_0}(k) = \begin{cases} \frac{10^{4k} e^{-10^4}}{k!} & k = 0, 1, \ldots \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$P_{K|H_1}(k) = \begin{cases} \frac{10^{6k} e^{-10^6}}{k!} & k = 0, 1, \ldots \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Since the two hypotheses are equally likely, the MAP and ML tests are the same. From Theorem 8.6, the ML hypothesis rule is

$$k \in A_0 \text{ if } P_{K|H_0}(k) \geq P_{K|H_1}(k); \qquad k \in A_1 \text{ otherwise.} \tag{3}$$

This rule simplifies to

$$k \in A_0 \text{ if } k \leq k^* = \frac{10^6 - 10^4}{\ln 100} = 214,975.7; \qquad k \in A_1 \text{ otherwise.} \tag{4}$$

Thus if we observe at least 214, 976 photons, then we accept hypothesis H_1.

Quiz 8.3

For the QPSK system, a symbol error occurs when s_i is transmitted but $(X_1, X_2) \in A_j$ for some $j \neq i$. For a QPSK system, it is easier to calculate the probability of a correct decision. Given H_0, the conditional probability of a correct decision is

$$P[C|H_0] = P[X_1 > 0, X_2 > 0|H_0] = P\left[\sqrt{E/2} + N_1 > 0, \sqrt{E/2} + N_2 > 0\right] \tag{1}$$

Because of the symmetry of the signals, $P[C|H_0] = P[C|H_i]$ for all i. This implies the probability of a correct decision is $P[C] = P[C|H_0]$. Since N_1 and N_2 are iid Gaussian $(0, \sigma)$ random variables, we have

$$P[C] = P[C|H_0] = P\left[\sqrt{E/2} + N_1 > 0\right] P\left[\sqrt{E/2} + N_2 > 0\right] \tag{2}$$

$$= \left(P\left[N_1 > -\sqrt{E/2}\right]\right)^2 \tag{3}$$

$$= \left[1 - \Phi\left(\frac{-\sqrt{E/2}}{\sigma}\right)\right]^2 \tag{4}$$

Since $\Phi(-x) = 1 - \Phi(x)$, we have $P[C] = \Phi^2(\sqrt{E/2\sigma^2})$. Equivalently, the probability of error is

$$P_{\text{ERR}} = 1 - P[C] = 1 - \Phi^2\left(\sqrt{\frac{E}{2\sigma^2}}\right) \tag{5}$$

Quiz 8.4

To generate the ROC, the existing program `sqdistor` already calculates this miss probability $P_{\text{MISS}} = P_{01}$ and the false alarm probability $P_{\text{FA}} = P_{10}$. The modified program, `sqdistroc.m` is essentially the same as `sqdistor` except the output is a matrix `FM` whose columns are the false alarm and miss probabilities. Next, the program `sqdistrocplot.m` calls `sqdistroc` three times to generate a plot that compares the receiver performance for the three requested values of d. Here is the modified code:

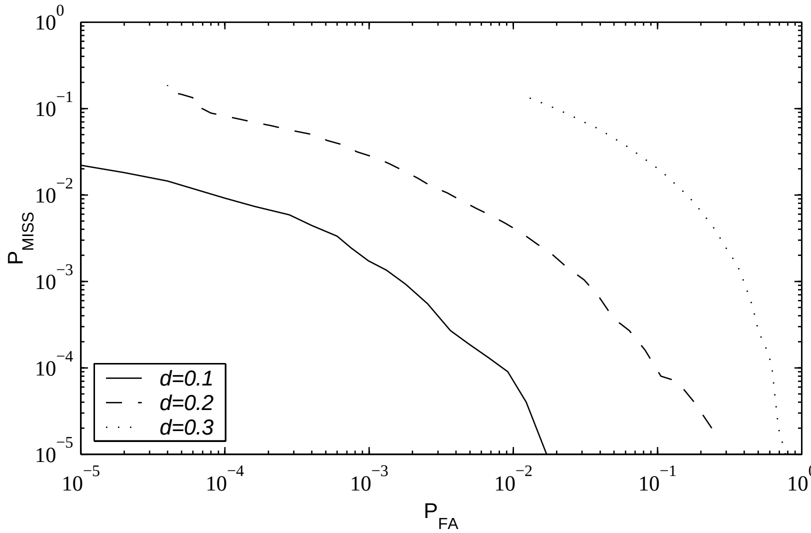
```
function FM=sqdistroc(v,d,m,T)
%square law distortion recvr
%P(error) for m bits tested
%transmit v volts or -v volts,
%add N volts, N is Gauss(0,1)
%add d(v+N)^2 distortion
%receive 1 if x>T, otherwise 0
%FM = [P(FA) P(MISS)]
x=(v+randn(m,1));
[XX,TT]=ndgrid(x,T(:));
P01=sum((XX+d*(XX.^2)< TT),1)/m;
x= -v+randn(m,1);
[XX,TT]=ndgrid(x,T(:));
P10=sum((XX+d*(XX.^2)>TT),1)/m;
FM=[P10(:) P01(:)];
```

```
function FM=sqdistrocplot(v,m,T);
FM1=sqdistroc(v,0.1,m,T);
FM2=sqdistroc(v,0.2,m,T);
FM5=sqdistroc(v,0.3,m,T);
FM=[FM1 FM2 FM5];
loglog(FM1(:,1),FM1(:,2),'-k', ...
    FM2(:,1),FM2(:,2),'--k', ...
    FM5(:,1),FM5(:,2),':k');
legend('\it d=0.1','\it d=0.2',...
    '\it d=0.3',3)
ylabel('P_{MISS}');
xlabel('P_{FA}');
```

To see the effect of d, the commands

```
T=-3:0.1:3; sqdistrocplot(3,100000,T);
```

generated the plot shown in Figure 3.



T=-3:0.1:3; sqdistrocplot(3,100000,T);

Figure 3: The receiver operating curve for the communications system of Quiz 8.4 with squared distortion.

Quiz Solutions – Chapter 9

Quiz 9.1

(1) First, we calculate the marginal PDF for $0 \le y \le 1$:

$$f_Y(y) = \int_0^y 2(y+x)\,dx = 2xy + x^2\Big|_{x=0}^{x=y} = 3y^2 \tag{1}$$

This implies the conditional PDF of X given Y is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{2}{3y} + \frac{2x}{3y^2} & 0 \le x \le y \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

(2) The minimum mean square error estimate of X given $Y = y$ is

$$\hat{x}_M(y) = E[X|Y=y] = \int_0^y \left(\frac{2x}{3y} + \frac{2x^2}{3y^2}\right) dx = 5y/9 \tag{3}$$

Thus the MMSE estimator of X given Y is $\hat{X}_M(Y) = 5Y/9$.

(3) To obtain the conditional PDF $f_{Y|X}(y|x)$, we need the marginal PDF $f_X(x)$. For $0 \le x \le 1$,

$$f_X(x) = \int_x^1 2(y+x)\,dy = y^2 + 2xy\Big|_{y=x}^{y=1} = 1 + 2x - 3x^2 \tag{4}$$

$$\tag{5}$$

For $0 \le x \le 1$, the conditional PDF of Y given X is

$$f_{Y|X}(y|x) = \begin{cases} \frac{2(y+x)}{1+2x-3x^2} & x \le y \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

(4) The MMSE estimate of Y given $X = x$ is

$$\hat{y}_M(x) = E[Y|X=x] = \int_x^1 \frac{2y^2 + 2xy}{1+2x-3x^2}\,dy \tag{7}$$

$$= \frac{2y^3/3 + xy^2}{1+2x-3x^2}\Bigg|_{y=x}^{y=1} \tag{8}$$

$$= \frac{2+3x-5x^3}{3+6x-9x^2} \tag{9}$$

Quiz 9.2

(1) Since the expectation of the sum equals the sum of the expectations,

$$E[R] = E[T] + E[X] = 0 \tag{1}$$

(2) Since T and X are independent, the variance of the sum $R = T + X$ is

$$\text{Var}[R] = \text{Var}[T] + \text{Var}[X] = 9 + 3 = 12 \tag{2}$$

(3) Since T and R have expected values $E[R] = E[T] = 0$,

$$\text{Cov}[T, R] = E[TR] = E[T(T + X)] = E\left[T^2\right] + E[TX] \tag{3}$$

Since T and X are independent and have zero expected value, $E[TX] = E[T]E[X] = 0$ and $E[T^2] = \text{Var}[T]$. Thus $\text{Cov}[T, R] = \text{Var}[T] = 9$.

(4) From Definition 4.8, the correlation coefficient of T and R is

$$\rho_{T,R} = \frac{\text{Cov}[T, R]}{\sqrt{\text{Var}[R]\,\text{Var}[T]}} = \frac{\sigma_T}{\sigma_R} = \sqrt{3}/2 \tag{4}$$

(5) From Theorem 9.4, the optimum linear estimate of T given R is

$$\hat{T}_L(R) = \rho_{T,R}\frac{\sigma_T}{\sigma_R}(R - E[R]) + E[T] \tag{5}$$

Since $E[R] = E[T] = 0$ and $\rho_{T,R} = \sigma_T/\sigma_R$,

$$\hat{T}_L(R) = \frac{\sigma_T^2}{\sigma_R^2}R = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_X^2}R = \frac{3}{4}R \tag{6}$$

Hence $a^* = 3/4$ and $b^* = 0$.

(6) By Theorem 9.4, the mean square error of the linear estimate is

$$e_L^* = \text{Var}[T](1 - \rho_{T,R}^2) = 9(1 - 3/4) = 9/4 \tag{7}$$

Quiz 9.3

When $R = r$, the conditional PDF of $X = Y - 40 - 40\log_{10} r$ is Gaussian with expected value $-40 - 40\log_{10} r$ and variance 64. The conditional PDF of X given R is

$$f_{X|R}(x|r) = \frac{1}{\sqrt{128\pi}}e^{-(x+40+40\log_{10} r)^2/128} \tag{1}$$

From the conditional PDF $f_{X|R}(x|r)$, we can use Definition 9.2 to write the ML estimate of R given $X = x$ as

$$\hat{r}_{\mathrm{ML}}(x) = \arg\max_{r \geq 0} f_{X|R}(x|r) \tag{2}$$

We observe that $f_{X|R}(x|r)$ is maximized when the exponent $(x + 40 + 40\log_{10} r)^2$ is minimized. This minimum occurs when the exponent is zero, yielding

$$\log_{10} r = -1 - x/40 \tag{3}$$

or

$$\hat{r}_{\mathrm{ML}}(x) = (0.1)10^{-x/40} \text{ m} \tag{4}$$

If the result doesn't look correct, note that a typical figure for the signal strength might be $x = -120$ dB. This corresponds to a distance estimate of $\hat{r}_{\mathrm{ML}}(-120) = 100$ m.

For the MAP estimate, we observe that the joint PDF of X and R is

$$f_{X,R}(x,r) = f_{X|R}(x|r)\, f_R(r) = \frac{1}{10^6\sqrt{32\pi}} r e^{-(x+40+40\log_{10} r)^2/128} \tag{5}$$

From Theorem 9.6, the MAP estimate of R given $X = x$ is the value of r that maximizes $f_{X,R}(x,r)$. That is,

$$\hat{r}_{\mathrm{MAP}}(x) = \arg\max_{0 \leq r \leq 1000} f_{X,R}(x,r) \tag{6}$$

Note that we have included the constraint $r \leq 1000$ in the maximization to highlight the fact that under our probability model, $R \leq 1000$ m. Setting the derivative of $f_{X,R}(x,r)$ with respect to r to zero yields

$$e^{-(x+40+40\log_{10} r)^2/128}\left[1 - \frac{80\log_{10} e}{128}(x + 40 + 40\log_{10} r)\right] = 0 \tag{7}$$

Solving for r yields

$$r = 10^{\left(\frac{1}{25\log_{10} e} - 1\right)} 10^{-x/40} = (0.1236)10^{-x/40} \tag{8}$$

This is the MAP estimate of R given $X = x$ as long as $r \leq 1000$ m. When $x \leq -156.3$ dB, the above estimate will exceed 1000 m, which is not possible in our probability model. Hence, the complete description of the MAP estimate is

$$\hat{r}_{\mathrm{MAP}}(x) = \begin{cases} 1000 & x < -156.3 \\ (0.1236)10^{-x/40} & x \geq -156.3 \end{cases} \tag{9}$$

For example, if $x = -120$dB, then $\hat{r}_{\mathrm{MAP}}(-120) = 123.6$ m. When the measured signal strength is not too low, the MAP estimate is 23.6% larger than the ML estimate. This reflects the fact that large values of R are a priori more probable than small values. However, for very low signal strengths, the MAP estimate takes into account that the distance can never exceed 1000 m.

Quiz 9.4

(1) From Theorem 9.4, the LMSE estimate of X_2 given Y_2 is $\hat{X}_2(Y_2) = a^* Y_2 + b^*$ where

$$a^* = \frac{\text{Cov}[X_2, Y_2]}{\text{Var}[Y_2]}, \qquad b^* = \mu_{X_2} - a^* \mu_{Y_2}. \tag{1}$$

Because $E[\mathbf{X}] = E[\mathbf{Y}] = \mathbf{0}$,

$$\text{Cov}[X_2, Y_2] = E[X_2 Y_2] = E[X_2(X_2 + W_2)] = E\left[X_2^2\right] = 1 \tag{2}$$

$$\text{Var}[Y_2] = \text{Var}[X_2] + \text{Var}[W_2] = E\left[X_2^2\right] + E\left[W_2^2\right] = 1.1 \tag{3}$$

It follows that $a^* = 1/1.1$. Because $\mu_{X_2} = \mu_{Y_2} = 0$, it follows that $b^* = 0$. Finally, to compute the expected square error, we calculate the correlation coefficient

$$\rho_{X_2,Y_2} = \frac{\text{Cov}[X_2, Y_2]}{\sigma_{X_2}\sigma_{Y_2}} = \frac{1}{\sqrt{1.1}} \tag{4}$$

The expected square error is

$$e_L^* = \text{Var}[X_2](1 - \rho_{X_2,Y_2}^2) = 1 - \frac{1}{1.1} = \frac{1}{11} = 0.0909 \tag{5}$$

(2) Since $\mathbf{Y} = \mathbf{X} + \mathbf{W}$ and $E[\mathbf{X}] = E[\mathbf{W}] = \mathbf{0}$, it follows that $E[\mathbf{Y}] = \mathbf{0}$. Thus we can apply Theorem 9.7. Note that $\mathbf{X}$ and $\mathbf{W}$ have correlation matrices

$$\mathbf{R_X} = \begin{bmatrix} 1 & -0.9 \\ -0.9 & 1 \end{bmatrix}, \qquad \mathbf{R_W} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}. \tag{6}$$

In terms of Theorem 9.7, $n = 2$ and we wish to estimate X_2 given the observation vector $\mathbf{Y} = \begin{bmatrix} Y_1 & Y_2 \end{bmatrix}'$. To apply Theorem 9.7, we need to find $\mathbf{R_Y}$ and $\mathbf{R}_{\mathbf{Y}X_2}$.

$$\mathbf{R_Y} = E\left[\mathbf{YY}'\right] = E\left[(\mathbf{X} + \mathbf{W})(\mathbf{X}' + \mathbf{W}')\right] \tag{7}$$

$$= E\left[\mathbf{XX}' + \mathbf{XW}' + \mathbf{WX}' + \mathbf{WW}'\right]. \tag{8}$$

Because $\mathbf{X}$ and $\mathbf{W}$ are independent, $E[\mathbf{XW}'] = E[\mathbf{X}]E[\mathbf{W}'] = \mathbf{0}$. Similarly, $E[\mathbf{WX}'] = \mathbf{0}$. This implies

$$\mathbf{R_Y} = E\left[\mathbf{XX}'\right] + E\left[\mathbf{WW}'\right] = \mathbf{R_X} + \mathbf{R_W} = \begin{bmatrix} 1.1 & -0.9 \\ -0.9 & 1.1 \end{bmatrix}. \tag{9}$$

In addition, we need to find

$$\mathbf{R}_{\mathbf{Y}X_2} = E[\mathbf{Y}X_2] = \begin{bmatrix} E[Y_1 X_2] \\ E[Y_2 X_2] \end{bmatrix} = \begin{bmatrix} E[(X_1 + W_1)X_2] \\ E[(X_2 + W_2)X_2] \end{bmatrix}. \tag{10}$$

Since $\mathbf{X}$ and $\mathbf{W}$ are independent vectors, $E[W_1X_2] = E[W_1]E[X_2] = 0$ and $E[W_2X_2] = 0$. Thus

$$\mathbf{R}_{\mathbf{Y}X_2} = \begin{bmatrix} E[X_1X_2] \\ E\left[X_2^2\right] \end{bmatrix} = \begin{bmatrix} -0.9 \\ 1 \end{bmatrix}. \tag{11}$$

By Theorem 9.7,

$$\hat{\mathbf{a}} = \mathbf{R}_{\mathbf{Y}}^{-1}\mathbf{R}_{\mathbf{Y}X_2} = \begin{bmatrix} -0.225 \\ 0.725 \end{bmatrix} \tag{12}$$

Therefore, the optimum linear estimator of X_2 given Y_1 and Y_2 is

$$\hat{X}_L = \hat{\mathbf{a}}'\mathbf{Y} = -0.225Y_1 + 0.725Y_2. \tag{13}$$

The mean square error is

$$\text{Var}\,[X_2] - \hat{\mathbf{a}}'\mathbf{R}_{\mathbf{Y}X_2} = \text{Var}\,[X] - a_1 r_{Y_1,X_2} - a_2 r_{Y_2,X_2} = 0.0725. \tag{14}$$

Quiz 9.5

Since X and $\mathbf{W}$ have zero expected value, $\mathbf{Y}$ also has zero expected value. Thus, by Theorem 9.7, $\hat{X}_L(\mathbf{Y}) = \hat{\mathbf{a}}'\mathbf{Y}$ where $\hat{\mathbf{a}} = \mathbf{R}_{\mathbf{Y}}^{-1}\mathbf{R}_{\mathbf{Y}X}$. Since X and $\mathbf{W}$ are independent, $E[\mathbf{W}X] = \mathbf{0}$ and $E[X\mathbf{W}'] = \mathbf{0}'$. This implies

$$\mathbf{R}_{\mathbf{Y}X} = E\,[\mathbf{Y}X] = E\,[(\mathbf{1}X + \mathbf{W})X] = \mathbf{1}E\left[X^2\right] = \mathbf{1}. \tag{1}$$

By the same reasoning, the correlation matrix of $\mathbf{Y}$ is

$$\begin{aligned} \mathbf{R}_{\mathbf{Y}} = E\left[\mathbf{Y}\mathbf{Y}'\right] &= E\left[(\mathbf{1}X + \mathbf{W})(\mathbf{1}'X + \mathbf{W}')\right] &(2)\\ &= \mathbf{1}\mathbf{1}'E\left[X^2\right] + \mathbf{1}E\left[X\mathbf{W}'\right] + E\,[\mathbf{W}X]\,\mathbf{1}' + E\left[\mathbf{W}\mathbf{W}'\right] &(3)\\ &= \mathbf{1}\mathbf{1}' + \mathbf{R}_{\mathbf{W}} &(4) \end{aligned}$$

Note that $\mathbf{1}\mathbf{1}'$ is a 20×20 matrix with every entry equal to 1. Thus,

$$\hat{\mathbf{a}} = \mathbf{R}_{\mathbf{Y}}^{-1}\mathbf{R}_{\mathbf{Y}X} = \left(\mathbf{1}\mathbf{1}' + \mathbf{R}_W\right)^{-1}\mathbf{1} \tag{5}$$

and the optimal linear estimator is

$$\hat{X}_L(\mathbf{Y}) = \mathbf{1}'\left(\mathbf{1}\mathbf{1}' + \mathbf{R}_{\mathbf{W}}\right)^{-1}\mathbf{Y} \tag{6}$$

The mean square error is

$$e_L^* = \text{Var}[X] - \hat{\mathbf{a}}'\mathbf{R}_{\mathbf{Y}X} = 1 - \mathbf{1}'\left(\mathbf{1}\mathbf{1}' + \mathbf{R}_{\mathbf{W}}\right)^{-1}\mathbf{1} \tag{7}$$

Now we note that $\mathbf{R}_W$ has i, jth entry $R_{\mathbf{W}}(i, j) = c^{|i-j|-1}$. The question we must address is what value c minimizes e_L^*. This problem is atypical in that one does not usually get to choose the correlation structure of the noise. However, we will see that the answer is somewhat instructive.

We note that the answer is not obviously apparent from Equation (7). In particular, we observe that $\text{Var}[W_i] = R_{\mathbf{W}}(i,i) = 1/c$. Thus, when c is small, the noises W_i have high variance and we would expect our estimator to be poor. On the other hand, if c is large W_i and W_j are highly correlated and the separate measurements of X are very dependent. This would suggest that large values of c will also result in poor MSE. If this argument is not clear, consider the extreme case in which every W_i and W_j have correlation coefficient $\rho_{ij} = 1$. In this case, our 20 measurements will be all the same and one measurement is as good as 20 measurements.

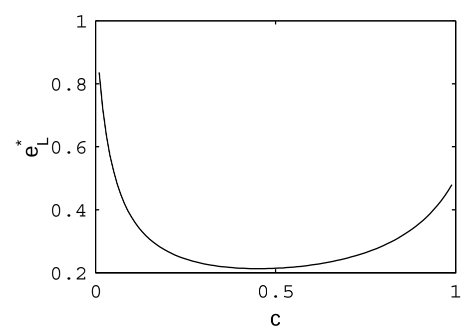
To find the optimal value of c, we write a MATLAB function `mquiz9(c)` to calculate the MSE for a given c and second function that finds plots the MSE for a range of values of c.

```
function [mse,af]=mquiz9(c);
v1=ones(20,1);
RW=toeplitz(c.^((0:19)-1));
RY=(v1*(v1')) +RW;
af=(inv(RY))*v1;
mse=1-((v1')*af);
```

```
function cmin=mquiz9minc(c);
msec=zeros(size(c));
for k=1:length(c),
    [msec(k),af]=mquiz9(c(k));
end
plot(c,msec);
xlabel('c');ylabel('e_L^*');
[msemin,optk]=min(msec);
cmin=c(optk);
```

Note in `mquiz9` that `v1` corresponds to the vector **1** of all ones. The following commands finds the minimum c and also produces the following graph:

```
>> c=0.01:0.01:0.99;
>> mquiz9minc(c)
ans =
    0.4500
```



As we see in the graph, both small values and large values of c result in large MSE.

Quiz Solutions – Chapter 10

Quiz 10.1

There are many correct answers to this question. A correct answer specifies enough random variables to specify the sample path exactly. One choice for an alternate set of random variables that would specify $m(t, s)$ is

- $m(0, s)$, the number of ongoing calls at the start of the experiment
- N, the number of new calls that arrive during the experiment
- $X_1, \ldots, X_N$, the interarrival times of the N new arrivals
- H, the number of calls that hang up during the experiment
- $D_1, \ldots, D_H$, the call completion times of the H calls that hang up

Quiz 10.2

(1) We obtain a continuous time, continuous valued process when we record the temperature as a continuous waveform over time.

(2) If at every moment in time, we round the temperature to the nearest degree, then we obtain a continuous time, discrete valued process.

(3) If we sample the process in part (a) every T seconds, then we obtain a discrete time, continuous valued process.

(4) Rounding the samples in part (c) to the nearest integer degree yields a discrete time, discrete valued process.

Quiz 10.3

(1) Each resistor has resistance R in ohms with uniform PDF

$$f_R(r) = \begin{cases} 0.01 & 950 \leq r \leq 1050 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The probability that a test produces a 1% resistor is

$$p = P[990 \leq R \leq 1010] = \int_{990}^{1010} (0.01)\, dr = 0.2 \tag{2}$$

(2) In t seconds, exactly t resistors are tested. Each resistor is a 1% resistor with probability p, independent of any other resistor. Consequently, the number of 1% resistors found has the binomial PMF

$$P_{N(t)}(n) = \begin{cases} \binom{t}{n} p^n (1-p)^{t-n} & n = 0, 1, \ldots, t \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

(3) First we will find the PMF of T_1. This problem is easy if we view each resistor test as an independent trial. A success occurs on a trial with probability p if we find a 1% resistor. The first 1% resistor is found at time $T_1 = t$ if we observe failures on trials $1, \ldots, t-1$ followed by a success on trial t. Hence, just as in Example 2.11, T_1 has the geometric PMF

$$P_{T_1}(t) = \begin{cases} (1-p)^{t-1} p & t = 1, 2, \ldots \\ 9 & \text{otherwise} \end{cases} \tag{4}$$

Since $p = 0.2$, the probability the first 1% resistor is found in exactly five seconds is $P_{T_1}(5) = (0.8)^4(0.2) = 0.08192$.

(4) From Theorem 2.5, a geometric random variable with success probability p has expected value $1/p$. In this problem, $E[T_1] = 1/p = 5$.

(5) Note that once we find the first 1% resistor, the number of additional trials needed to find the second 1% resistor once again has a geometric PMF with expected value $1/p$ since each independent trial is a success with probability p. That is, $T_2 = T_1 + T'$ where T' is independent and identically distributed to T_1. Thus

$$E[T_2|T_1 = 10] = E[T_1|T_1 = 10] + E\left[T'|T_1 = 10\right] \tag{5}$$

$$= 10 + E\left[T'\right] = 10 + 5 = 15 \tag{6}$$

Quiz 10.4

Since each X_i is a $N(0, 1)$ random variable, each X_i has PDF

$$f_{X(i)}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{1}$$

By Theorem 10.1, the joint PDF of $\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_n \end{bmatrix}'$ is

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X(1),\ldots,X(n)}(x_1, \ldots, x_n) = \prod_{i=1}^{k} f_X(x_i) = \frac{1}{(2\pi)^{n/2}} e^{-(x_1^2+\cdots+x_n^2)/2} \tag{2}$$

Quiz 10.5

The first and second hours are nonoverlapping intervals. Since one hour equals 3600 sec and the Poisson process has a rate of 10 packets/sec, the expected number of packets in each hour is $E[M_i] = \alpha = 36,000$. This implies M_1 and M_2 are independent Poisson random variables each with PMF

$$P_{M_i}(m) = \begin{cases} \frac{\alpha^m e^{-\alpha}}{m!} & m = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Since M_1 and M_2 are independent, the joint PMF of M_1 and M_2 is

$$P_{M_1,M_2}(m_1, m_2) = P_{M_1}(m_1)\, P_{M_2}(m_2) = \begin{cases} \frac{\alpha^{m_1+m_2} e^{-2\alpha}}{m_1! m_2!} & m_1 = 0, 1, \ldots; \\ & m_2 = 0, 1, \ldots, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Quiz 10.6

To answer whether $N'(t)$ is a Poisson process, we look at the interarrival times. Let $X_1, X_2, \ldots$ denote the interarrival times of the $N(t)$ process. Since we count only even-numbered arrival for $N'(t)$, the time until the first arrival of the $N'(t)$ is $Y_1 = X_1 + X_2$. Since X_1 and X_2 are independent exponential (λ) random variables, Y_1 is an Erlang $(n = 2, \lambda)$ random variable; see Theorem 6.11. Since $Y_i(t)$, the ith interarrival time of the $N'(t)$ process, has the same PDF as $Y_1(t)$, we can conclude that the interarrival times of $N'(t)$ are not exponential random variables. Thus $N'(t)$ is *not* a Poisson process.

Quiz 10.7

First, we note that for $t > s$,

$$X(t) - X(s) = \frac{W(t) - W(s)}{\sqrt{\alpha}} \tag{1}$$

Since $W(t) - W(s)$ is a Gaussian random variable, Theorem 3.13 states that $W(t) - W(s)$ is Gaussian with expected value

$$E[X(t) - X(s)] = \frac{E[W(t) - W(s)]}{\sqrt{\alpha}} = 0 \tag{2}$$

and variance

$$E\left[(W(t) - W(s))^2\right] = \frac{E\left[(W(t) - W(s))^2\right]}{\alpha} = \frac{\alpha(t-s)}{\alpha} \tag{3}$$

Consider $s' \leq s < t$. Since $s \geq s'$, $W(t) - W(s)$ is independent of $W(s')$. This implies $[W(t) - W(s)]/\sqrt{\alpha}$ is independent of $W(s')/\sqrt{\alpha}$ for all $s \geq s'$. That is, $X(t) - X(s)$ is independent of $X(s')$ for all $s \geq s'$. Thus $X(t)$ is a Brownian motion process with variance $\text{Var}[X(t)] = t$.

Quiz 10.8

First we find the expected value

$$\mu_Y(t) = \mu_X(t) + \mu_N(t) = \mu_X(t). \tag{1}$$

To find the autocorrelation, we observe that since $X(t)$ and $N(t)$ are independent and since $N(t)$ has zero expected value, $E[X(t)N(t')] = E[X(t)]E[N(t')] = 0$. Since $R_Y(t,\tau) = E[Y(t)Y(t+\tau)]$, we have

$$R_Y(t,\tau) = E\left[(X(t)+N(t))\left(X(t+\tau)+N(t+\tau)\right)\right] \tag{2}$$

$$= E\left[X(t)X(t+\tau)\right] + E\left[X(t)N(t+\tau)\right]$$
$$+ E\left[X(t+\tau)N(t)\right] + E\left[N(t)N(t+\tau)\right] \tag{3}$$

$$= R_X(t,\tau) + R_N(t,\tau). \tag{4}$$

Quiz 10.9

From Definition 10.14, $X_1, X_2, \ldots$ is a stationary random sequence if for all sets of time instants $n_1, \ldots, n_m$ and time offset k,

$$f_{X_{n_1},\ldots,X_{n_m}}(x_1,\ldots,x_m) = f_{X_{n_1+k},\ldots,X_{n_m+k}}(x_1,\ldots,x_m) \tag{1}$$

Since the random sequence is iid,

$$f_{X_{n_1},\ldots,X_{n_m}}(x_1,\ldots,x_m) = f_X(x_1)\, f_X(x_2)\cdots f_X(x_m) \tag{2}$$

Similarly, for time instants $n_1+k,\ldots,n_m+k$,

$$f_{X_{n_1+k},\ldots,X_{n_m+k}}(x_1,\ldots,x_m) = f_X(x_1)\, f_X(x_2)\cdots f_X(x_m) \tag{3}$$

We can conclude that the iid random sequence is stationary.

Quiz 10.10

We must check whether each function $R(\tau)$ meets the conditions of Theorem 10.12:

$$R(\tau) \geq 0 \qquad R(\tau) = R(-\tau) \qquad |R(\tau)| \leq R(0) \tag{1}$$

(1) $R_1(\tau) = e^{-|\tau|}$ meets all three conditions and thus is valid.

(2) $R_2(\tau) = e^{-\tau^2}$ also is valid.

(3) $R_3(\tau) = e^{-\tau}\cos\tau$ is not valid because

$$R_3(-2\pi) = e^{2\pi}\cos 2\pi = e^{2\pi} > 1 = R_3(0) \tag{2}$$

(4) $R_4(\tau) = e^{-\tau^2}\sin\tau$ also cannot be an autocorrelation function because

$$R_4(\pi/2) = e^{-\pi/2}\sin\pi/2 = e^{-\pi/2} > 0 = R_4(0) \tag{3}$$

Quiz 10.11

(1) The autocorrelation of $Y(t)$ is

$$R_Y(t,\tau) = E[Y(t)Y(t+\tau)] \tag{1}$$
$$= E[X(-t)X(-t-\tau)] \tag{2}$$
$$= R_X(-t-(-t-\tau)) = R_X(\tau) \tag{3}$$

Since $E[Y(t)] = E[X(-t)] = \mu_X$, we can conclude that $Y(t)$ is a wide sense stationary process. In fact, we see that by viewing a process backwards in time, we see the same second order statistics.

(2) Since $X(t)$ and $Y(t)$ are both wide sense stationary processes, we can check whether they are jointly wide sense stationary by seeing if $R_{XY}(t,\tau)$ is just a function of τ. In this case,

$$R_{XY}(t,\tau) = E[X(t)Y(t+\tau)] \tag{4}$$
$$= E[X(t)X(-t-\tau)] \tag{5}$$
$$= R_X(t-(-t-\tau)) = R_X(2t+\tau) \tag{6}$$

Since $R_{XY}(t,\tau)$ depends on both t and τ, we conclude that $X(t)$ and $Y(t)$ are not jointly wide sense stationary. To see why this is, suppose $R_X(\tau) = e^{-|\tau|}$ so that samples of $X(t)$ far apart in time have almost no correlation. In this case, as t gets larger, $Y(t) = X(-t)$ and $X(t)$ become less and less correlated.

Quiz 10.12

From the problem statement,

$$E[X(t)] = E[X(t+1)] = 0 \tag{1}$$
$$E[X(t)X(t+1)] = 1/2 \tag{2}$$
$$\mathrm{Var}[X(t)] = \mathrm{Var}[X(t+1)] = 1 \tag{3}$$

The Gaussian random vector $\mathbf{X} = \begin{bmatrix} X(t) & X(t+1) \end{bmatrix}'$ has covariance matrix and corresponding inverse

$$\mathbf{C_X} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \qquad \mathbf{C}_{\mathbf{X}}^{-1} = \frac{4}{3}\begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} \tag{4}$$

Since

$$\mathbf{x}'\mathbf{C}_{\mathbf{X}}^{-1}\mathbf{x} = \begin{bmatrix} x_0 & x_1 \end{bmatrix}' \frac{4}{3}\begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix}\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \frac{4}{3}\left(x_0^2 - x_0 x_+ x_1^2\right) \tag{5}$$

the joint PDF of $X(t)$ and $X(t+1)$ is the Gaussian vector PDF

$$f_{X(t),X(t+1)}(x_0,x_1) = \frac{1}{(2\pi)^{n/2}[\det(\mathbf{C_X})]^{1/2}} \exp\left(-\frac{1}{2}\mathbf{x}'\mathbf{C}_{\mathbf{X}}^{-1}\mathbf{x}\right) \tag{6}$$
$$= \frac{1}{\sqrt{3\pi^2}} e^{-\frac{2}{3}\left(x_0^2 - x_0 x_1 + x_1^2\right)} \tag{7}$$

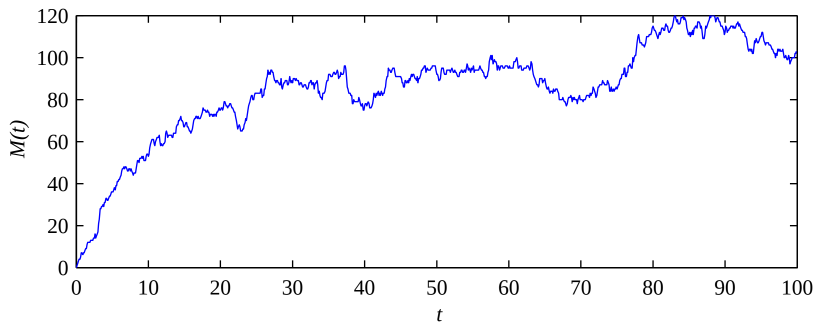


Figure 4: Sample path of 100 minutes of the blocking switch of Quiz 10.13.

Quiz 10.13

The simple structure of the switch simulation of Example 10.28 admits a deceptively simple solution in terms of the vector of arrivals `A` and the vector of departures `D`. With the introduction of call blocking. we cannot generate these vectors all at once. In particular, when an arrival occurs at time t, we need to know that $M(t)$, the number of ongoing calls, satisfies $M(t) < c = 120$. Otherwise, when $M(t) = c$, we must block the call. Call blocking can be implemented by setting the service time of the call to zero so that the call departs as soon as it arrives.

The blocking switch is an example of a discrete event system. The system evolves via a sequence of discrete events, namely arrivals and departures, at discrete time instances. A simulation of the system moves from one time instant to the next by maintaining a chronological schedule of future events (arrivals and departures) to be executed. The program simply executes the event at the head of the schedule. The logic of such a simulation is

1. Start at time $t = 0$ with an empty system. Schedule the first arrival to occur at S_1, an exponential (λ) random variable.

2. Examine the head-of-schedule event.

   - When the head-of-schedule event is the kth arrival is at time t, check the state $M(t)$.
     - If $M(t) < c$, admit the arrival, increase the system state n by 1, and schedule a departure to occur at time $t + S_n$, where S_k is an exponential (λ) random variable.
     - If $M(t) = c$, block the arrival, do not schedule a departure event.
   - If the head of schedule event is a departure, reduce the system state n by 1.

3. Delete the head-of-schedule event and go to step 2.

After the head-of-schedule event is completed and any new events (departures in this system) are scheduled, we know the system state cannot change until the next scheduled event.

Thus we know that $M(t)$ will stay the same until then. In our simulation, we use the vector `t` as the set of time instances at which we inspect the system state. Thus for all times `t(i)` between the current head-of-schedule event and the next, we set `m(i)` to the current switch state.

The complete program is shown in Figure 5. In most programming languages, it is common to implement the event schedule as a linked list where each item in the list has a data structure indicating an event timestamp and the type of the event. In MATLAB, a simple (but not elegant) way to do this is to have maintain two vectors: `time` is a list of timestamps of scheduled events and `event` is a the list of event types. In this case, `event(i)=1` if the ith scheduled event is an arrival, or `event(i)=-1` if the ith scheduled event is a departure.

When the program is passed a vector `t`, the output `[m a b]` is such that `m(i)` is the number of ongoing calls at time `t(i)` while `a` and `b` are the number of admits and blocks. The following instructions

```
t=0:0.1:5000;
[m,a,b]=simblockswitch(10,0.1,120,t);
plot(t,m);
```

generated a simulation lasting 5,000 minutes. A sample path of the first 100 minutes of that simulation is shown in Figure 4. The 5,000 minute full simulation produced `a=49658` admitted calls and `b=239` blocked calls. We can estimate the probability a call is blocked as

$$\hat{P}_b = \frac{b}{a+b} = 0.0048. \tag{1}$$

In Chapter 12, we will learn that the exact blocking probability is given by Equation (12.93), a result known as the "Erlang-B formula." From the Erlang-B formula, we can calculate that the exact blocking probability is $P_b = 0.0057$. One reason our simulation underestimates the blocking probability is that in a 5,000 minute simulation, roughly the first 100 minutes are needed to load up the switch since the switch is idle when the simulation starts at time $t = 0$. However, this says that roughly the first two percent of the simulation time was unusual. Thus this would account for only part of the disparity. The rest of the gap between 0.0048 and 0.0057 is that a simulation that includes only 239 blocks is not all that likely to give a very accurate result for the blocking probability.

Note that in Chapter 12, we will learn that the blocking switch is an example of an $M/M/c/c$ queue, a kind of Markov chain. Chapter 12 develops techniques for analyzing and simulating systems described by Markov chains that are much simpler than the discrete event simulation technique shown here. Nevertheless, for very complicated systems, the discrete event simulation is widely-used and often very efficient simulation method.

```
function [M,admits,blocks]=simblockswitch(lam,mu,c,t);
blocks=0; %total # blocks
admits=0; %total # admits
M=zeros(size(t));
n=0; % # in system
time=[ exponentialrv(lam,1) ];
event=[ 1 ]; %first event is an arrival
timenow=0;
tmax=max(t);
while (timenow<tmax)
   M((timenow<=t)&(t<time(1)))=n;
   timenow=time(1);
   eventnow=event(1);
   event(1)=[ ]; time(1)= [ ]; % clear current event
   if (eventnow==1) % arrival
       arrival=timenow+exponentialrv(lam,1); % next arrival
       b4arrival=time<arrival;
       event=[event(b4arrival) 1 event(~b4arrival)];
       time=[time(b4arrival) arrival time(~b4arrival)];
       if n<c %call admitted
         admits=admits+1;
         n=n+1;
         depart=timenow+exponentialrv(mu,1);
         b4depart=time<depart;
         event=[event(b4depart) -1 event(~b4depart)];
         time=[time(b4depart) depart time(~b4depart)];
       else
         blocks=blocks+1; %one more block, immed departure
         disp(sprintf('Time %10.3d Admits %10d Blocks %10d',...
             timenow,admits,blocks));
       end
   elseif (eventnow==-1) %departure
       n=n-1;
   end
end
```

Figure 5: Discrete event simulation of the blocking switch of Quiz 10.13.

Quiz Solutions – Chapter 11

Quiz 11.1

By Theorem 11.2,

$$\mu_Y = \mu_X \int_{-\infty}^{\infty} h(t)dt = 2\int_0^{\infty} e^{-t}\,dt = 2 \tag{1}$$

Since $R_X(\tau) = \delta(\tau)$, the autocorrelation function of the output is

$$R_Y(\tau) = \int_{-\infty}^{\infty} h(u)\int_{-\infty}^{\infty} h(v)\delta(\tau + u - v)\,dv\,du = \int_{-\infty}^{\infty} h(u)h(\tau + u)\,du \tag{2}$$

For $\tau > 0$, we have

$$R_Y(\tau) = \int_0^{\infty} e^{-u}e^{-\tau-u}\,du = e^{-\tau}\int_0^{\infty} e^{-2u}\,du = \frac{1}{2}e^{-\tau} \tag{3}$$

For $\tau < 0$, we can deduce that $R_Y(\tau) = \frac{1}{2}e^{-|\tau|}$ by symmetry. Just to be safe though, we can double check. For $\tau < 0$,

$$R_Y(\tau) = \int_{-\tau}^{\infty} h(u)h(\tau + u)\,du = \int_{-\tau}^{\infty} e^{-u}e^{-\tau-u}\,du = \frac{1}{2}e^{\tau} \tag{4}$$

Hence,

$$R_Y(\tau) = \frac{1}{2}e^{-|\tau|} \tag{5}$$

Quiz 11.2

The expected value of the output is

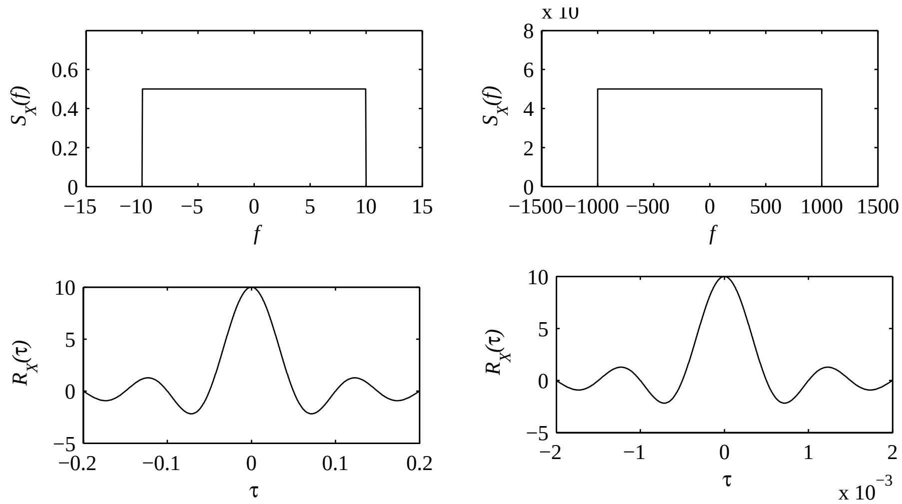
$$\mu_Y = \mu_X \sum_{n=-\infty}^{\infty} h_n = 0.5(1 + -1) = 0 \tag{1}$$

The autocorrelation of the output is

$$R_Y[n] = \sum_{i=0}^{1}\sum_{j=0}^{1} h_i h_j R_X[n + i - j] \tag{2}$$

$$= 2R_X[n] - R_X[n-1] - R_X[n+1] = \begin{cases} 1 & n = 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Since $\mu_Y = 0$, The variance of Y_n is $\text{Var}[Y_n] = E[Y_n^2] = R_Y[0] = 1$.



(a) $W = 10$ **(b)** $W = 1000$

Figure 6: The autocorrelation $R_X(\tau)$ and power spectral density $S_X(f)$ for process $X(t)$ in Quiz 11.5.

Quiz 11.3

By Theorem 11.8, $\mathbf{Y} = \begin{bmatrix} Y_{33} & Y_{34} & Y_{35} \end{bmatrix}'$ is a Gaussian random vector since X_n is a Gaussian random process. Moreover, by Theorem 11.5, each Y_n has expected value $E[Y_n] = \mu_X \sum_{n=-\infty}^{\infty} h_n = 0$. Thus $E[\mathbf{Y}] = \mathbf{0}$. Fo find the PDF of the Gaussian vector $\mathbf{Y}$, we need to find the covariance matrix $\mathbf{C_Y}$, which equals the correlation matrix $\mathbf{R_Y}$ since $\mathbf{Y}$ has zero expected value. One way to find the $\mathbf{R_Y}$ is to observe that $\mathbf{R_Y}$ has the Toeplitz structure of Theorem 11.6 and to use Theorem 11.5 to find the autocorrelation function

$$R_Y[n] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h_i h_j R_X[n+i-j]. \tag{1}$$

Despite the fact that $R_X[k]$ is an impulse, using Equation (1) is surprisingly tedious because we still need to sum over all i and j such that $n+i-j=0$.

In this problem, it is simpler to observe that $\mathbf{Y} = \mathbf{HX}$ where

$$\mathbf{X} = \begin{bmatrix} X_{30} & X_{31} & X_{32} & X_{33} & X_{34} & X_{35} \end{bmatrix}' \tag{2}$$

and

$$\mathbf{H} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{3}$$

In this case, following Theorem 11.7, or by directly applying Theorem 5.13 with $\boldsymbol{\mu}_\mathbf{X} = \mathbf{0}$ and $\mathbf{A} = \mathbf{H}$, we obtain $\mathbf{R_Y} = \mathbf{H R_X H}'$. Since $R_X[n] = \delta_n$, $\mathbf{R_X} = \mathbf{I}$, the identity matrix.

Thus

$$\mathbf{C_Y} = \mathbf{R_Y} = \mathbf{H}\mathbf{H}' = \frac{1}{16}\begin{bmatrix} 4 & 3 & 2 \\ 3 & 4 & 3 \\ 2 & 3 & 4 \end{bmatrix}. \tag{4}$$

It follows (very quickly if you use MATLAB for 3×3 matrix inversion) that

$$\mathbf{C}_{\mathbf{Y}}^{-1} = 16\begin{bmatrix} 7/12 & -1/2 & 1/12 \\ -1/2 & 1 & -1/2 \\ 1/12 & -1/2 & 7/12 \end{bmatrix}. \tag{5}$$

Thus, the PDF of $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{3/2}[\det(\mathbf{C_Y})]^{1/2}} \exp\left(-\frac{1}{2}\mathbf{y}'\mathbf{C}_{\mathbf{Y}}^{-1}\mathbf{y}\right). \tag{6}$$

A disagreeable amount of algebra will show $\det(\mathbf{C_Y}) = 3/1024$ and that the PDF can be "simplified" to

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{16}{\sqrt{6\pi^3}} \exp\left[-8\left(\frac{7}{12}y_{33}^2 + y_{34}^2 + \frac{7}{12}y_{35}^2 - y_{33}y_{34} + \frac{1}{6}y_{33}y_{35} - y_{34}y_{35}\right)\right]. \tag{7}$$

Equation (7) shows that one of the nicest features of the multivariate Gaussian distribution is that $\mathbf{y}'\mathbf{C}_{\mathbf{Y}}^{-1}\mathbf{y}$ is a very concise representation of the cross-terms in the exponent of $f_{\mathbf{Y}}(\mathbf{y})$.

Quiz 11.4

This quiz is solved using Theorem 11.9 for the case of $k = 1$ and $M = 2$. In this case, $\mathbf{X}_n = \begin{bmatrix} X_{n-1} & X_n \end{bmatrix}'$ and

$$\mathbf{R}_{\mathbf{X}_n} = \begin{bmatrix} R_X[0] & R_X[1] \\ R_X[1] & R_X[0] \end{bmatrix} = \begin{bmatrix} 1.1 & 0.9 \\ 0.9 & 1.1 \end{bmatrix} \tag{1}$$

and

$$\mathbf{R}_{\mathbf{X}_n X_{n+1}} = E\left[\begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix} X_{n+1}\right] = \begin{bmatrix} R_X[2] \\ R_X[1] \end{bmatrix} = \begin{bmatrix} 0.81 \\ 0.9 \end{bmatrix}. \tag{2}$$

The MMSE linear first order filter for predicting X_{n+1} at time n is the filter $\mathbf{h}$ such that

$$\overleftarrow{\mathbf{h}} = \mathbf{R}_{\mathbf{X}_n}^{-1}\mathbf{R}_{\mathbf{X}_n X_{n+1}} = \begin{bmatrix} 1.1 & 0.9 \\ 0.9 & 1.1 \end{bmatrix}^{-1}\begin{bmatrix} 0.81 \\ 0.9 \end{bmatrix} = \frac{1}{400}\begin{bmatrix} 81 \\ 261 \end{bmatrix}. \tag{3}$$

It follows that the filter is $\mathbf{h} = \begin{bmatrix} 261/400 & 81/400 \end{bmatrix}'$ and the MMSE linear predictor is

$$\hat{X}_{n+1} = \frac{81}{400}X_{n-1} + \frac{261}{400}X_n. \tag{4}$$

to find the mean square error, one approach is to follow the method of Example 11.13 and to directly calculate

$$e_L^* = E\left[(X_{n+1} - \hat{X}_{n+1})^2\right]. \tag{5}$$

This method is workable for this simple problem but becomes increasingly tedious for higher order filters. Instead, we can derive the mean square error for an arbitary prediction filter $\mathbf{h}$. Since $\hat{X}_{n+1} = \overleftarrow{\mathbf{h}}'\mathbf{X}_n$,

$$e_L^* = E\left[\left(X_{n+1} - \overleftarrow{\mathbf{h}}'\mathbf{X}_n\right)^2\right] \tag{6}$$

$$= E\left[(X_{n+1} - \overleftarrow{\mathbf{h}}'\mathbf{X}_n)(X_{n+1} - \overleftarrow{\mathbf{h}}'\mathbf{X}_n)'\right] \tag{7}$$

$$= E\left[(X_{n+1} - \overleftarrow{\mathbf{h}}'\mathbf{X}_n)(X_{n+1} - \mathbf{X}_n'\overleftarrow{\mathbf{h}})\right] \tag{8}$$

After a bit of algebra, we obtain

$$e_L^* = R_X[0] - 2\overleftarrow{\mathbf{h}}'\mathbf{R}_{\mathbf{X}_n X_{n+1}} + \overleftarrow{\mathbf{h}}'\mathbf{R}_{\mathbf{X}_n}\overleftarrow{\mathbf{h}} \tag{9}$$

$$\tag{10}$$

with the substitution $\overleftarrow{\mathbf{h}} = \mathbf{R}_{\mathbf{X}_n}^{-1}\mathbf{R}_{\mathbf{X}_n X_{n+1}}$, we obtain

$$e_L^* = R_X[0] - \mathbf{R}_{\mathbf{X}_n X_{n+1}}'\mathbf{R}_{\mathbf{X}_n}^{-1}\mathbf{R}_{\mathbf{X}_n X_{n+1}} \tag{11}$$

$$= R_X[0] - \overleftarrow{\mathbf{h}}'\mathbf{R}_{\mathbf{X}_n X_{n+1}} \tag{12}$$

Note that this is essentially the same result as Theorem 9.7 with $\mathbf{Y} = \mathbf{X}_n$, $X = X_{n+1}$ and $\hat{\mathbf{a}}' = \overleftarrow{\mathbf{h}}'$. It is noteworthy that the result is derived in a much simpler way in the proof of Theorem 9.7 by using the orthoginality property of the LMSE estimator.

In any case, the mean square error is

$$e_L^* = R_X[0] - \overleftarrow{\mathbf{h}}'\mathbf{R}_{\mathbf{X}_n X_{n+1}} = 1.1 - \frac{1}{400}\begin{bmatrix}81 & 261\end{bmatrix}\begin{bmatrix}0.81\\0.9\end{bmatrix} = \frac{506}{1451} = 0.3487. \tag{13}$$

recalling that the blind estimate would yield a mean square error of $\text{Var}[X] = 1.1$, we see that observing X_{n-1} and X_n improves the accuracy of our prediction of X_{n+1}.

Quiz 11.5

(1) By Theorem 11.13(b), the average power of $X(t)$ is

$$E\left[X^2(t)\right] = \int_{-\infty}^{\infty} S_X(f)\,df = \int_{-W}^{W}\frac{5}{W}\,df = 10 \text{ Watts} \tag{1}$$

(2) The autocorrelation function is the inverse Fourier transform of $S_X(f)$. Consulting Table 11.1, we note that

$$S_X(f) = 10\frac{1}{2W}\text{rect}\left(\frac{f}{2W}\right) \tag{2}$$

It follows that the inverse transform of $S_X(f)$ is

$$R_X(\tau) = 10\,\text{sinc}(2W\tau) = 10\frac{\sin(2\pi W\tau)}{2\pi W\tau} \tag{3}$$

(3) For $W = 10$ Hz and $W = 1$ kHZ, graphs of $S_X(f)$ and $R_X(\tau)$ appear in Figure 6.

Quiz 11.6

In a sampled system, the discrete time impulse $\delta[n]$ has a flat discrete Fourier transform. That is, if $R_X[n] = 10\delta[n]$, then

$$S_X(\phi) = \sum_{n=-\infty}^{\infty} 10\delta[n]e^{-j2\pi\phi n} = 10 \tag{1}$$

Thus, $R_X[n] = 10\delta[n]$. (This quiz is really lame!)

Quiz 11.7

Since $Y(t) = X(t - t_0)$,

$$R_{XY}(t, \tau) = E[X(t)Y(t + \tau)] = E[X(t)X(t + \tau - t_0)] = R_X(\tau - t_0) \tag{1}$$

We see that $R_{XY}(t, \tau) = R_{XY}(\tau) = R_X(\tau - t_0)$. From Table 11.1, we recall the property that $g(\tau - \tau_0)$ has Fourier transform $G(f)e^{-j2\pi f\tau_0}$. Thus the Fourier transform of $R_{XY}(\tau) = R_X(\tau - t_0) = g(\tau - t_0)$ is

$$S_{XY}(f) = S_X(f)e^{-j2\pi f t_0}. \tag{2}$$

Quiz 11.8

We solve this quiz using Theorem 11.17. First we need some preliminary facts. Let $a_0 = 5{,}000$ so that

$$R_X(\tau) = \frac{1}{a_0}a_0 e^{-a_0|\tau|}. \tag{1}$$

Consulting with the Fourier transforms in Table 11.1, we see that

$$S_X(f) = \frac{1}{a_0}\frac{2a_0^2}{a_0^2 + (2\pi f)^2} = \frac{2a_0}{a_0^2 + (2\pi f)^2} \tag{2}$$

The RC filter has impulse response $h(t) = a_1 e^{-a_1 t}u(t)$, where $u(t)$ is the unit step function and $a_1 = 1/RC$ where $RC = 10^{-4}$ is the filter time constant. From Table 11.1,

$$H(f) = \frac{a_1}{a_1 + j2\pi f} \tag{3}$$

(1) Theorem 11.17,

$$S_{XY}(f) = H(f)S_X(f) = \frac{2a_0a_1}{[a_1 + j2\pi f]\left[a_0^2 + (2\pi f)^2\right]}. \tag{4}$$

(2) Again by Theorem 11.17,

$$S_Y(f) = H^*(f)S_{XY}(f) = |H(f)|^2 S_X(f). \tag{5}$$

Note that

$$|H(f)|^2 = H(f)H^*(f) = \frac{a_1}{(a_1 + j2\pi f)}\frac{a_1}{(a_1 - j2\pi f)} = \frac{a_1^2}{a_1^2 + (2\pi f)^2} \tag{6}$$

Thus,

$$S_Y(f) = |H(f)|^2\, S_X(f) = \frac{2a_0 a_1^2}{\left[a_1^2 + (2\pi f)^2\right]\left[a_0^2 + (2\pi f)^2\right]} \tag{7}$$

(3) To find the average power at the filter output, we can either use basic calculus and calculate $\int_{-\infty}^{\infty} S_Y(f)\,df$ directly or we can find $R_Y(\tau)$ as an inverse transform of $S_Y(f)$. Using partial fractions and the Fourier transform table, the latter method is actually less algebra. In particular, some algebra will show that

$$S_Y(f) = \frac{K_0}{a_0^2 + (2\pi f)^2} + \frac{K_1}{a_1 + (2\pi f)^2} \tag{8}$$

where

$$K_0 = \frac{2a_0 a_1^2}{a_1^2 - a_0^2}, \qquad K_1 = \frac{-2a_0 a_1^2}{a_1^2 - a_0^2}. \tag{9}$$

Thus,

$$S_Y(f) = \frac{K_0}{2a_0^2}\frac{2a_0^2}{a_0^2 + (2\pi f)^2} + \frac{K_1}{2a_1^2}\frac{2a_1^2}{a_1 + (2\pi f)^2}. \tag{10}$$

Consulting with Table 11.1, we see that

$$R_Y(\tau) = \frac{K_0}{2a_0^2} a_0 e^{-a_0|\tau|} + \frac{K_1}{2a_1^2} a_1 e^{-a_1|\tau|} \tag{11}$$

Substituting the values of K_0 and K_1, we obtain

$$R_Y(\tau) = \frac{a_1^2 e^{-a_0|\tau|} - a_0 a_1 e^{-a_1|\tau|}}{a_1^2 - a_0^2}. \tag{12}$$

The average power of the $Y(t)$ process is

$$R_Y(0) = \frac{a_1}{a_1 + a_0} = \frac{2}{3}. \tag{13}$$

Note that the input signal has average power $R_X(0) = 1$. Since the RC filter has a 3dB bandwidth of 10,000 rad/sec and the signal $X(t)$ has most of its its signal energy below 5,000 rad/sec, the output signal has almost as much power as the input.

Quiz 11.9

This quiz implements an example of Equations (11.146) and (11.147) for a system in which we filter $Y(t) = X(t) + N(t)$ to produce an optimal linear estimate of $X(t)$. The solution to this quiz is just to find the filter $\hat{H}(f)$ using Equation (11.146) and to calculate the mean square error e_L* using Equation (11.147).

Comment: Since the text omitted the derivations of Equations (11.146) and (11.147), we note that Example 10.24 showed that

$$R_Y(\tau) = R_X(\tau) + R_N(\tau), \qquad R_{YX}(\tau) = R_X(\tau). \tag{1}$$

Taking Fourier transforms, it follows that

$$S_Y(f) = S_X(f) + S_N(f), \qquad S_{YX}(f) = S_X(f). \tag{2}$$

Now we can go on to the quiz, at peace with the derivations.

(1) Since $\mu_N = 0$, $R_N(0) = \text{Var}[N] = 1$. This implies

$$R_N(0) = \int_{-\infty}^{\infty} S_N(f)\,df = \int_{-B}^{B} N_0\,df = 2N_0B \tag{3}$$

Thus $N_0 = 1/(2B)$. Because the noise process $N(t)$ has constant power $R_N(0) = 1$, decreasing the single-sided bandwidth B increases the power spectral density of the noise over frequencies $|f| < B$.

(2) Since $R_X(\tau) = \text{sinc}(2W\tau)$, where $W = 5{,}000$ Hz, we see from Table 11.1 that

$$S_X(f) = \frac{1}{10^4}\,\text{rect}\left(\frac{f}{10^4}\right). \tag{4}$$

The noise power spectral density can be written as

$$S_N(f) = N_0\,\text{rect}\left(\frac{f}{2B}\right) = \frac{1}{2B}\,\text{rect}\left(\frac{f}{2B}\right), \tag{5}$$

From Equation (11.146), the optimal filter is

$$\hat{H}(f) = \frac{S_X(f)}{S_X(f) + S_N(f)} = \frac{\frac{1}{10^4}\,\text{rect}\left(\frac{f}{10^4}\right)}{\frac{1}{10^4}\,\text{rect}\left(\frac{f}{10^4}\right) + \frac{1}{2B}\,\text{rect}\left(\frac{f}{2B}\right)}. \tag{6}$$

(3) We produce the output $\hat{X}(t)$ by passing the noisy signal $Y(t)$ through the filter $\hat{H}(f)$. From Equation (11.147), the mean square error of the estimate is

$$e_L^* = \int_{-\infty}^{\infty} \frac{S_X(f)S_N(f)}{S_X(f)+S_N(f)}\,df \tag{7}$$

$$= \int_{-\infty}^{\infty} \frac{\frac{1}{10^4}\operatorname{rect}\left(\frac{f}{10^4}\right)\frac{1}{2B}\operatorname{rect}\left(\frac{f}{2B}\right)}{\frac{1}{10^4}\operatorname{rect}\left(\frac{f}{10^4}\right)+\frac{1}{2B}\operatorname{rect}\left(\frac{f}{2B}\right)}\,df. \tag{8}$$

To evaluate the MSE e_L^*, we need to whether $B \le W$. Since the problem asks us to find the largest possible B, let's suppose $B \le W$. We can go back and consider the case $B > W$ later. When $B \le W$, the MSE is

$$e_L^* = \int_{-B}^{B} \frac{\frac{1}{10^4}\frac{1}{2B}}{\frac{1}{10^4}+\frac{1}{2B}}\,df = \frac{\frac{1}{10^4}}{\frac{1}{10^4}+\frac{1}{2B}} = \frac{1}{1+\frac{5{,}000}{B}} \tag{9}$$

To obtain MSE $e_L^* \le 0.05$ requires $B \le 5{,}000/19 = 263.16$ Hz.

Although this completes the solution to the quiz, what is happening may not be obvious. The noise power is always $\operatorname{Var}[N] = 1$ Watt, for all values of B. As B is decreased, the PSD $S_N(f)$ becomes increasingly tall, but only over a bandwidth B that is decreasing. Thus as B descreases, the filter $\hat{H}(f)$ makes an increasingly deep and narrow notch at frequencies $|f| \le B$. Two examples of the filter $\hat{H}(f)$ are shown in Figure 7. As B shrinks, the filter suppresses less of the signal of $X(t)$. The result is that the MSE goes down.

Finally, we note that we can choose B very large and also achieve MSE $e_L^* = 0.05$. In particular, when $B > W = 5000$, $S_N(f) = 1/2B$ over frequencies $|f| < W$. In this case, the Wiener filter $\hat{H}(f)$ is an ideal (flat) lowpass filter

$$\hat{H}(f) = \begin{cases} \frac{\frac{1}{10^4}}{\frac{1}{10^4}+\frac{1}{2B}} & |f| < 5{,}000, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

Thus increasing B spreads the constant 1 watt of power of $N(t)$ over more bandwidth. The Wiener filter removes the noise that is outside the band of the desired signal. The mean square error is

$$e_L^* = \int_{-5000}^{5000} \frac{\frac{1}{10^4}\frac{1}{2B}}{\frac{1}{10^4}+\frac{1}{2B}}\,df = \frac{\frac{1}{2B}}{\frac{1}{10^4}+\frac{1}{2B}} = \frac{1}{\frac{B}{5000}+1} \tag{11}$$

In this case, $B \ge 9.5 \times 10^4$ guarantees $e_L^* \le 0.05$.

Quiz 11.10

It is fairly straightforward to find $S_X(\phi)$ and $S_Y(\phi)$. The only thing to keep in mind is to use `fftc` to transform the autocorrelation $R_X[f]$ into the power spectral density $S_X(\phi)$. The following MATLAB program generates and plots the functions shown in Figure 8

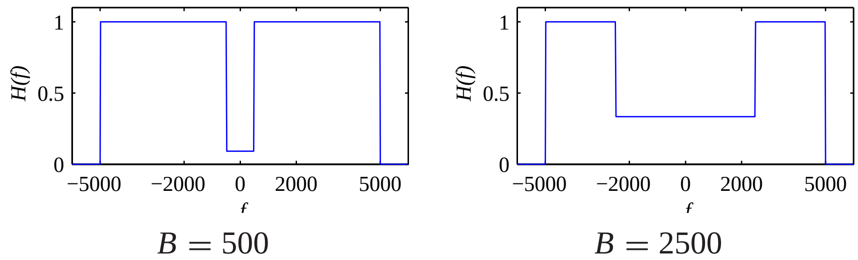


Figure 7: Wiener filter for Quiz 11.9.

```
%mquiz11.m
N=32;
rx=[2 4 2]; SX=fftc(rx,N); %autocorrelation and PSD
stem(0:N-1,abs(sx));
xlabel('n');ylabel('S_X(n/N)');
h2=0.5*[1 1]; H2=fft(h2,N); %impulse/filter response: M=2
SY2=SX.* ((abs(H2)).^2);
figure; stem(0:N-1,abs(SY2)); %PSD of Y for M=2
xlabel('n');ylabel('S_{Y_2}(n/N)');
h10=0.1*ones(1,10); H10=fft(h10,N); %impulse/filter response: M=10
SY10=sx.*((abs(H10)).^2);
figure; stem(0:N-1,abs(SY10));
xlabel('n');ylabel('S_{Y_{10}}(n/N)');
```

Relative to $M = 2$, when $M = 10$, the filter $H(\phi)$ filters out almost all of the high frequency components of $X(t)$. In the context of Example 11.26, the low pass moving average filter for $M = 10$ removes the high frquency components and results in a filter output that varies very slowly.

As an aside, note that the vectors `SX`, `SY2` and `SY10` in `mquiz11` should all be real-valued vectors. However, the finite numerical precision of MATLAB results in tiny imaginary parts. Although these imaginary parts have no computational significance, they tend to confuse the `stem` function. Hence, we generate stem plots of the magnitude of each power spectral density.

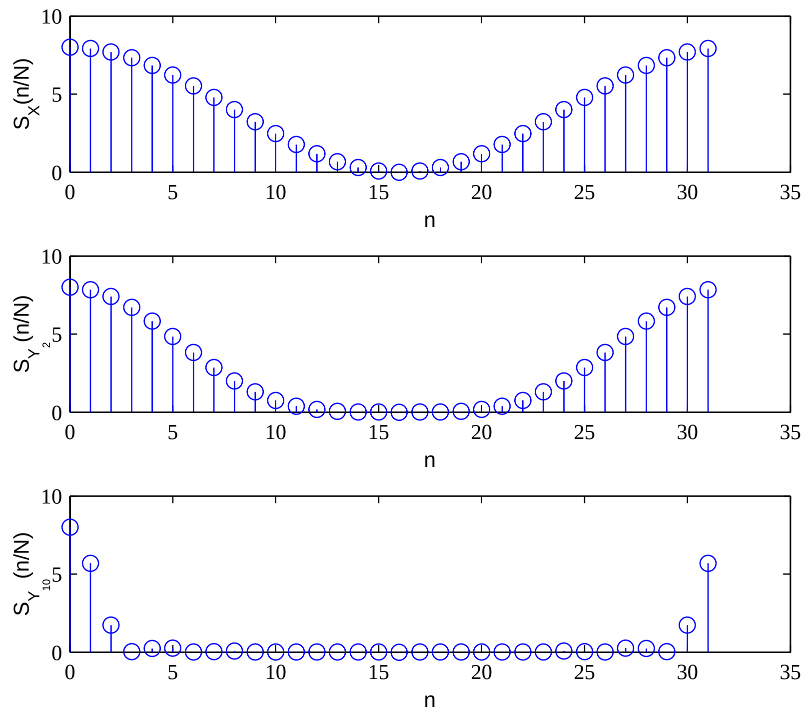


Figure 8: For Quiz 11.10, graphs of $S_X(\phi)$, $S_Y(n/N)$ for $M = 2$, and $S_\phi(n/N)$ for $M = 10$ using an $N = 32$ point DFT.

Quiz Solutions – Chapter 12

Quiz 12.1

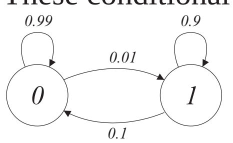
The system has two states depending on whether the previous packet was received in error. From the problem statement, we are given the conditional probabilities

$$P\left[X_{n+1} = 0|X_n = 0\right] = 0.99 \qquad P\left[X_{n+1} = 1|X_n = 1\right] = 0.9 \tag{1}$$

Since each X_n must be either 0 or 1, we can conclude that

$$P\left[X_{n+1} = 1|X_n = 0\right] = 0.01 \qquad P\left[X_{n+1} = 0|X_n = 1\right] = 0.1 \tag{2}$$

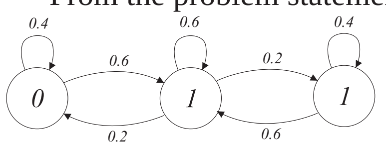
These conditional probabilities correspond to the transition matrix and Markov chain:



$$\mathbf{P} = \begin{bmatrix} 0.99 & 0.01 \\ 0.10 & 0.90 \end{bmatrix} \tag{3}$$

Quiz 12.2

From the problem statement, the Markov chain and the transition matrix are



$$\mathbf{P} = \begin{bmatrix} 0.4 & 0.6 & 0 \\ 0.2 & 0.6 & 0.2 \\ 0 & 0.6 & 0.4 \end{bmatrix} \tag{1}$$

The eigenvalues of **P** are

$$\lambda_1 = 0 \qquad \lambda_2 = 0.4 \qquad \lambda_3 = 1 \tag{2}$$

We can diagonalize **P** into

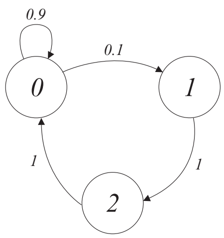
$$\mathbf{P} = \mathbf{S}^{-1}\mathbf{D}\mathbf{S} = \begin{bmatrix} -0.6 & 0.5 & 1 \\ 0.4 & 0 & 1 \\ -0.6 & -0.5 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} -0.5 & 1 & -0.5 \\ 1 & 0 & -1 \\ 0.2 & 0.6 & 0.2 \end{bmatrix} \tag{3}$$

where $\mathbf{s}_i$, the ith row of $\mathbf{S}$, is the left eigenvector of $\mathbf{P}$ satisfying $\mathbf{s}_i\mathbf{P} = \lambda_i\mathbf{s}_i$. Algebra will verify that the n-step transition matrix is

$$\mathbf{P}^n = \mathbf{S}^{-1}\mathbf{D}^n\mathbf{S} = \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.2 & 0.6 & 0.2 \end{bmatrix} + (0.4)^n \begin{bmatrix} 0.5 & 0 & -0.5 \\ 0 & 0 & 0 \\ -0.5 & 0 & 0.5 \end{bmatrix} \tag{4}$$

Quiz 12.3

The Markov chain describing the factory status and the corresponding state transition matrix are



$$\mathbf{P} = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \tag{1}$$

With $\boldsymbol{\pi} = \begin{bmatrix}\pi_0 & \pi_1 & \pi_2\end{bmatrix}'$, the system of equations $\boldsymbol{\pi}' = \boldsymbol{\pi}'\mathbf{P}$ yields $\pi_1 = 0.1\pi_0$ and $\pi_2 = \pi_1$. This implies

$$\pi_0 + \pi_1 + \pi_2 = \pi_0(1 + 0.1 + 0.1) = 1 \tag{2}$$

It follows that the limiting state probabilities are

$$\pi_0 = 5/6, \qquad \pi_1 = 1/12, \qquad \pi_2 = 1/12. \tag{3}$$

Quiz 12.4

The communicating classes are

$$C_1 = \{0, 1\} \qquad C_2 = \{2, 3\} \qquad C_3 = \{4, 5, 6\} \tag{1}$$

The states in C_1 and C_3 are aperiodic. The states in C_2 have period 2. Once the system enters a state in C_1, the class C_1 is never left. Thus the states in C_1 are recurrent. That is, C_1 is a recurrent class. Similarly, the states in C_3 are recurrent. On the other hand, the states in C_2 are transient. Once the system exits C_2, the states in C_2 are never reentered.

Quiz 12.5

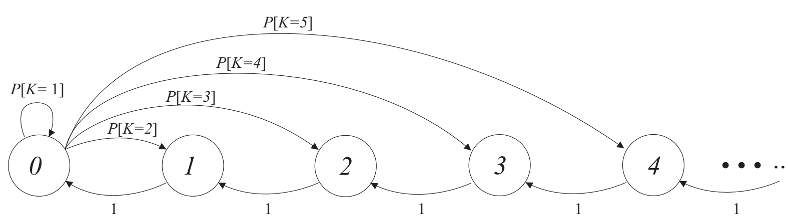
At any time t, the state n can take on the values $0, 1, 2, \ldots$. The state transition probabilities are

$$P_{n-1,n} = P\left[K > n | K > n-1\right] = \frac{P\left[K > n\right]}{P\left[K > n-1\right]} \tag{1}$$

$$P_{n-1,0} = P\left[K = n | K > n-1\right] = \frac{P\left[K = n\right]}{P\left[K > n-1\right]} \tag{2}$$

$$\tag{3}$$

The Markov chain resembles



The stationary probabilities satisfy

$$\pi_0 = \pi_0 P[K=1] + \pi_1, \tag{4}$$
$$\pi_1 = \pi_0 P[K=2] + \pi_2, \tag{5}$$
$$\vdots$$
$$\pi_{k-1} = \pi_0 P[K=k] + \pi_k, \qquad k = 1, 2, \ldots \tag{6}$$

From Equation (4), we obtain

$$\pi_1 = \pi_0 (1 - P[K=1]) = \pi_0 P[K>1] \tag{7}$$

Similarly, Equation (5) implies

$$\pi_2 = \pi_1 - \pi_0 P[K=2] = \pi_0 (P[K>1] - P[K=2]) = \pi_0 P[K>2] \tag{8}$$

This suggests that $\pi_k = \pi_0 P[K>k]$. We verify this pattern by showing that $\pi_k = \pi_0 P[K>k]$ satisfies Equation (6):

$$\pi_0 P[K>k-1] = \pi_0 P[K=k] + \pi_0 P[K>k]. \tag{9}$$

When we apply $\sum_{k=0}^{\infty} \pi_k = 1$, we obtain $\pi_0 \sum_{n=0}^{\infty} P[K>k] = 1$. From Problem 2.5.11, we recall that $\sum_{k=0}^{\infty} P[K>k] = E[K]$. This implies

$$\pi_n = \frac{P[K>n]}{E[K]} \tag{10}$$

This Markov chain models repeated random countdowns. The system state is the time until the counter expires. When the counter expires, the system is in state 0, and we randomly reset the counter to a new value $K = k$ and then we count down k units of time. Since we spend one unit of time in each state, including state 0, we have $k-1$ units of time left after the state 0 counter reset. If we have a random variable W such that the PMF of W satisfies $P_W(n) = \pi_n$, then W has a discrete PMF representing the remaining time of the counter at a time in the distant future.

Quiz 12.6

(1) By inspection, the number of transitions need to return to state 0 is always a multiple of 2. Thus the period of state 0 is $d = 2$.

(2) To find the stationary probabilities, we solve the system of equations $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$ and $\sum_{i=0}^{3} \pi_i = 1$:

$$\pi_0 = (3/4)\pi_1 + (1/4)\pi_3 \tag{1}$$
$$\pi_1 = (1/4)\pi_0 + (1/4)\pi_2 \tag{2}$$
$$\pi_2 = (1/4)\pi_1 + (3/4)\pi_3 \tag{3}$$
$$1 = \pi_0 + \pi_1 + \pi_2 + \pi_3 \tag{4}$$

Solving the second and third equations for π_2 and π_3 yields

$$\pi_2 = 4\pi_1 - \pi_0 \qquad \pi_3 = (4/3)\pi_2 - (1/3)\pi_1 = 5\pi_1 - (4/3)\pi_0 \tag{5}$$

Substituting π_3 back into the first equation yields

$$\pi_0 = (3/4)\pi_1 + (1/4)\pi_3 = (3/4)\pi_1 + (5/4)\pi_1 - (1/3)\pi_0 \tag{6}$$

This implies $\pi_1 = (2/3)\pi_0$. It follows from the first and second equations that $\pi_2 = (5/3)\pi_0$ and $\pi_3 = 2\pi_0$. Lastly, we choose π_0 so the state probabilities sum to 1:

$$1 = \pi_0 + \pi_1 + \pi_2 + \pi_3 = \pi_0 \left(1 + \frac{2}{3} + \frac{5}{3} + 2\right) = \frac{16}{3}\pi_0 \tag{7}$$

It follows that the state probabilities are

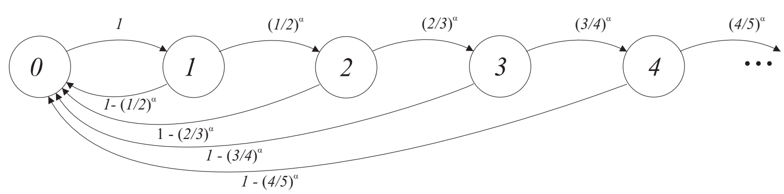
$$\pi_0 = \frac{3}{16} \quad \pi_1 = \frac{2}{16} \quad \pi_2 = \frac{5}{16} \quad \pi_3 = \frac{6}{16} \tag{8}$$

(3) Since the system starts in state 0 at time 0, we can use Theorem 12.14 to find the limiting probability that the system is in state 0 at time nd:

$$\lim_{n\to\infty} P_{00}(nd) = d\pi_0 = \frac{3}{8} \tag{9}$$

Quiz 12.7

The Markov chain has the same structure as that in Example 12.22. The only difference is the modified transition rates:



The event $T_{00} > n$ occurs if the system reaches state n before returning to state 0, which occurs with probability

$$P[T_{00} > n] = 1 \times \left(\frac{1}{2}\right)^\alpha \times \left(\frac{2}{3}\right)^\alpha \times \cdots \times \left(\frac{n-1}{n}\right)^\alpha = \left(\frac{1}{n}\right)^\alpha. \tag{1}$$

Thus the CDF of T_{00} satisfies $F_{T_{00}}(n) = 1 - P[T_{00} > n] = 1 - 1/n^\alpha$. To determine whether state 0 is recurrent, we observe that for all $\alpha > 0$

$$P[V_{00}] = \lim_{n\to\infty} F_{T_{00}}(n) = \lim_{n\to\infty} 1 - \frac{1}{n^\alpha} = 1. \tag{2}$$

Thus state 0 is recurrent for all $\alpha > 0$. Since the chain has only one communicating class, all states are recurrent. (We also note that if $\alpha = 0$, then all states are transient.)

To determine whether the chain is null recurrent or positive recurrent, we need to calculate $E[T_{00}]$. In Example 12.24, we did this by deriving the PMF $P_{T_{00}}(n)$. In this problem, it will be simpler to use the result of Problem 2.5.11 which says that $\sum_{k=0}^{\infty} P[K > k] = E[K]$ for any non-negative integer-valued random variable K. Applying this result, the expected time to return to state 0 is

$$E\left[T_{00}\right] = \sum_{n=0}^{\infty} P\left[T_{00} > n\right] = 1 + \sum_{n=1}^{\infty} \frac{1}{n^{\alpha}}. \tag{3}$$

For $0 < \alpha \leq 1$, $1/n^{\alpha} \geq 1/n$ and it follows that

$$E\left[T_{00}\right] \geq 1 + \sum_{n=1}^{\infty} \frac{1}{n} = \infty. \tag{4}$$

We conclude that the Markov chain is null recurrent for $0 < \alpha \leq 1$. On the other hand, for $\alpha > 1$,

$$E\left[T_{00}\right] = 2 + \sum_{n=2}^{\infty} \frac{1}{n^{\alpha}}. \tag{5}$$

Note that for all $n \geq 2$

$$\frac{1}{n^{\alpha}} \leq \int_{n-1}^{n} \frac{dx}{x^{\alpha}} \tag{6}$$

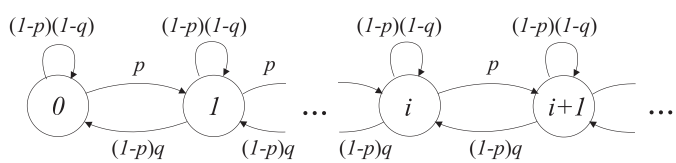
This implies

$$\begin{aligned} E\left[T_{00}\right] &\leq 2 + \sum_{n=2}^{\infty} \int_{n-1}^{n} \frac{dx}{x^{\alpha}} && (7)\\ &= 2 + \int_{1}^{\infty} \frac{dx}{x^{\alpha}} && (8)\\ &= 2 + \left. \frac{x^{-\alpha+1}}{-\alpha+1} \right|_{1}^{\infty} = 2 + \frac{1}{\alpha - 1} < \infty && (9) \end{aligned}$$

Thus for all $\alpha > 1$, the Markov chain is positive recurrent.

Quiz 12.8

The number of customers in the "friendly" store is given by the Markov chain



In the above chain, we note that $(1-p)q$ is the probability that no new customer arrives, an existing customer gets one unit of service and then departs the store.

By applying Theorem 12.13 with state space partitioned between $S = \{0, 1, \ldots, i\}$ and $S' = \{i+1, i+2, \ldots\}$, we see that for any state $i \geq 0$,

$$\pi_i p = \pi_{i+1}(1-p)q. \tag{1}$$

This implies

$$\pi_{i+1} = \frac{p}{(1-p)q}\pi_i. \tag{2}$$

Since Equation (2) holds for $i = 0, 1, \ldots$, we have that $\pi_i = \pi_0 \alpha^i$ where

$$\alpha = \frac{p}{(1-p)q}. \tag{3}$$

Requiring the state probabilities to sum to 1, we have that for $\alpha < 1$,

$$\sum_{i=0}^{\infty} \pi_i = \pi_0 \sum_{i=0}^{\infty} \alpha^i = \frac{\pi_0}{1-\alpha} = 1. \tag{4}$$

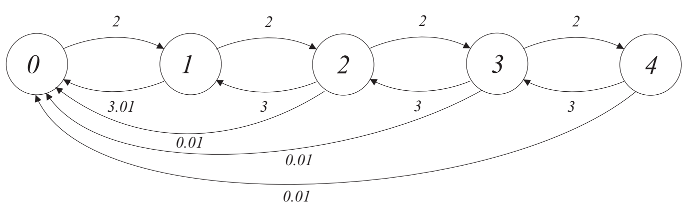
Thus for $\alpha < 1$, the limiting state probabilities are

$$\pi_i = (1-\alpha)\alpha^i, \qquad i = 0, 1, 2, \ldots \tag{5}$$

In addition, for $\alpha \geq 1$ or, equivalently, $p \geq q/(1-q)$, the limiting state probabilities do not exist.

Quiz 12.9

The continuous time Markov chain describing the processor is



Note that $q_{10} = 3.1$ since the task completes at rate 3 per msec and the processor reboots at rate 0.1 per msec and the rate to state 0 is the sum of those two rates. From the Markov chain, we obtain the following useful equations for the stationary distribution.

$$5.01p_1 = 2p_0 + 3p_2 \qquad 5.01p_2 = 2p_1 + 3p_3$$
$$5.01p_3 = 2p_2 + 3p_4 \qquad 3.01p_4 = 2p_3$$

We can solve these equations by working backward and solving for p_4 in terms of p_3, p_3 in terms of p_2 and so on, yielding

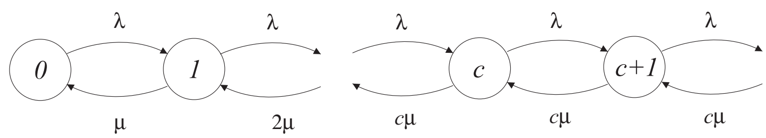
$$p_4 = \frac{20}{31}p_3 \quad p_3 = \frac{620}{981}p_2 \quad p_2 = \frac{19620}{31431}p_1 \quad p_1 = \frac{628,620}{1,014,381}p_0 \tag{1}$$

Applying $p_0 + p_1 + p_2 + p_3 + p_4 = 1$ yields $p_0 = 1,014,381/2,443,401$ and the stationary probabilities are

$$p_0 = 0.4151 \quad p_1 = 0.2573 \quad p_2 = 0.1606 \quad p_3 = 0.1015 \quad p_4 = 0.0655 \tag{2}$$

Quiz 12.10

The $M/M/c/\infty$ queue has Markov chain



From the Markov chain, the stationary probabilities must satisfy

$$p_n = \begin{cases} (\rho/n)p_{n-1} & n = 1, 2, \ldots, c \\ (\rho/c)p_{n-1} & n = c+1, c+2, \ldots \end{cases} \tag{1}$$

It is straightforward to show that this implies

$$p_n = \begin{cases} p_0 \rho^n/n! & n = 1, 2, \ldots, c \\ p_0 (\rho/c)^{n-c} \rho^c/c! & n = c+1, c+2, \ldots \end{cases} \tag{2}$$

The requirement that $\sum_{n=0}^{\infty} p_n = 1$ yields

$$p_0 = \left(\sum_{n=0}^{c} \rho^n/n! + \frac{\rho^c}{c!} \frac{\rho/c}{1-\rho/c} \right)^{-1} \tag{3}$$