

## Problem Solutions – Chapter 5

### Problem 5.1.1 Solution

The repair of each laptop can be viewed as an independent trial with four possible outcomes corresponding to the four types of needed repairs.

- (a) Since the four types of repairs are mutually exclusive choices and since 4 laptops are returned for repair, the joint distribution of  $N_1, \dots, N_4$  is the multinomial PMF

$$P_{N_1, \dots, N_4}(n_1, \dots, n_4) = \binom{4}{n_1, n_2, n_3, n_4} p_1^{n_1} p_2^{n_2} p_3^{n_3} p_4^{n_4} \quad (1)$$

$$= \begin{cases} \frac{4!}{n_1! n_2! n_3! n_4!} \left(\frac{8}{15}\right)^{n_1} \left(\frac{4}{15}\right)^{n_2} \left(\frac{2}{15}\right)^{n_3} \left(\frac{1}{15}\right)^{n_4} & n_1 + \dots + n_4 = 4; n_i \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

- (b) Let  $L_2$  denote the event that exactly two laptops need LCD repairs. Thus  $P[L_2] = P_{N_1}(2)$ . Since each laptop requires an LCD repair with probability  $p_1 = 8/15$ , the number of LCD repairs,  $N_1$ , is a binomial  $(4, 8/15)$  random variable with PMF

$$P_{N_1}(n_1) = \binom{4}{n_1} (8/15)^{n_1} (7/15)^{4-n_1} \quad (3)$$

The probability that two laptops need LCD repairs is

$$P_{N_1}(2) = \binom{4}{2} (8/15)^2 (7/15)^2 = 0.3717 \quad (4)$$

- (c) A repair is type (2) with probability  $p_2 = 4/15$ . A repair is type (3) with probability  $p_3 = 2/15$ ; otherwise a repair is type “other” with probability  $p_o = 9/15$ . Define  $X$  as the number of “other” repairs needed. The joint PMF of  $X, N_2, N_3$  is the multinomial PMF

$$P_{N_2, N_3, X}(n_2, n_3, x) = \binom{4}{n_2, n_3, x} \left(\frac{4}{15}\right)^{n_2} \left(\frac{2}{15}\right)^{n_3} \left(\frac{9}{15}\right)^x \quad (5)$$

However, Since  $X + 4 - N_2 - N_3$ , we observe that

$$P_{N_2, N_3}(n_2, n_3) = P_{N_2, N_3, X}(n_2, n_3, 4 - n_2 - n_3) \quad (6)$$

$$= \binom{4}{n_2, n_3, 4 - n_2 - n_3} \left(\frac{4}{15}\right)^{n_2} \left(\frac{2}{15}\right)^{n_3} \left(\frac{9}{15}\right)^{4 - n_2 - n_3} \quad (7)$$

$$= \left(\frac{9}{15}\right)^4 \binom{4}{n_2, n_3, 4 - n_2 - n_3} \left(\frac{4}{9}\right)^{n_2} \left(\frac{2}{9}\right)^{n_3} \quad (8)$$

Similarly, since each repair is a motherboard repair with probability  $p_2 = 4/15$ , the number of motherboard repairs has binomial PMF

$$P_{N_2}(n_2) = \binom{4}{n_2} \left(\frac{4}{15}\right)^{n_2} \left(\frac{11}{15}\right)^{4-n_2} \quad (9)$$

Finally, the probability that more laptops require motherboard repairs than keyboard repairs is

$$P[N_2 > N_3] = P_{N_2, N_3}(1, 0) + P_{N_2, N_3}(2, 0) + P_{N_2, N_3}(2, 1) + P_{N_2}(3) + P_{N_2}(4) \quad (10)$$

where we use the fact that if  $N_2 = 3$  or  $N_2 = 4$ , then we must have  $N_2 > N_3$ . Inserting the various probabilities, we obtain

$$P[N_2 > N_3] = P_{N_2, N_3}(1, 0) + P_{N_2, N_3}(2, 0) + P_{N_2, N_3}(2, 1) + P_{N_2}(3) + P_{N_2}(4) \quad (11)$$

Plugging in the various probabilities yields  $P[N_2 > N_3] = 8,656/16,875 \approx 0.5129$ .

### Problem 5.1.2 Solution

Whether a computer has feature  $i$  is a Bernoulli trial with success probability  $p_i = 2^{-i}$ . Given that  $n$  computers were sold, the number of computers sold with feature  $i$  has the binomial PMF

$$P_{N_i}(n_i) = \begin{cases} \binom{n}{n_i} p_i^{n_i} (1 - p_i)^{n - n_i} & n_i = 0, 1, \dots, n \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Since a computer has feature  $i$  with probability  $p_i$  independent of whether any other feature is on the computer, the number  $N_i$  of computers with feature  $i$  is independent of the number of computers with any other features. That is,  $N_1, \dots, N_4$  are mutually independent and have joint PMF

$$P_{N_1, \dots, N_4}(n_1, \dots, n_4) = P_{N_1}(n_1) P_{N_2}(n_2) P_{N_3}(n_3) P_{N_4}(n_4) \quad (2)$$

### Problem 5.1.3 Solution

(a) In terms of the joint PDF, we can write joint CDF as

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{X_1, \dots, X_n}(y_1, \dots, y_n) dy_1 \cdots dy_n \quad (1)$$

However, simplifying the above integral depends on the values of each  $x_i$ . In particular,  $f_{X_1, \dots, X_n}(y_1, \dots, y_n) = 1$  if and only if  $0 \leq y_i \leq 1$  for each  $i$ . Since  $F_{X_1, \dots, X_n}(x_1, \dots, x_n) = 0$  if any  $x_i < 0$ , we limit, for the moment, our attention to the case where  $x_i \geq 0$  for all  $i$ . In this case, some thought will show that we can write the limits in the following way:

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \int_0^{\max(1, x_1)} \cdots \int_0^{\min(1, x_n)} dy_1 \cdots dy_n \quad (2)$$

$$= \min(1, x_1) \min(1, x_2) \cdots \min(1, x_n) \quad (3)$$

A complete expression for the CDF of  $X_1, \dots, X_n$  is

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \begin{cases} \prod_{i=1}^n \min(1, x_i) & 0 \leq x_i, i = 1, 2, \dots, n \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

(b) For  $n = 3$ ,

$$1 - P \left[ \min_i X_i \leq 3/4 \right] = P \left[ \min_i X_i > 3/4 \right] \quad (5)$$

$$= P [X_1 > 3/4, X_2 > 3/4, X_3 > 3/4] \quad (6)$$

$$= \int_{3/4}^1 \int_{3/4}^1 \int_{3/4}^1 dx_1 dx_2 dx_3 \quad (7)$$

$$= (1 - 3/4)^3 = 1/64 \quad (8)$$

Thus  $P[\min_i X_i \leq 3/4] = 63/64$ .

### Problem 5.2.1 Solution

This problem is very simple. In terms of the vector  $\mathbf{X}$ , the PDF is

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} 1 & \mathbf{0} \leq \mathbf{x} \leq \mathbf{1} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

However, just keep in mind that the inequalities  $\mathbf{0} \leq \mathbf{x}$  and  $\mathbf{x} \leq \mathbf{1}$  are vector inequalities that must hold for every component  $x_i$ .

### Problem 5.2.2 Solution

In this problem, we find the constant  $c$  from the requirement that that the integral of the vector PDF over all possible values is 1. That is,  $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x}) dx_1 \cdots dx_n = 1$ . Since  $f_{\mathbf{X}}(\mathbf{x}) = c \mathbf{a}' \mathbf{x} = c \sum_{i=1}^n a_i x_i$ , we have that

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x}) dx_1 \cdots dx_n = c \int_0^1 \cdots \int_0^1 \left( \sum_{i=1}^n a_i x_i \right) dx_1 \cdots dx_n \quad (1)$$

$$= c \sum_{i=1}^n \left( \int_0^1 \cdots \int_0^1 a_i x_i dx_1 \cdots dx_n \right) \quad (2)$$

$$= c \sum_{i=1}^n a_i \left[ \left( \int_0^1 dx_1 \right) \cdots \left( \int_0^1 x_i dx_i \right) \cdots \left( \int_0^1 dx_n \right) \right] \quad (3)$$

$$= c \sum_{i=1}^n a_i \left( \frac{x_i^2}{2} \Big|_0^1 \right) = c \sum_{i=1}^n \frac{a_i}{2} \quad (4)$$

The requirement that the PDF integrate to unity thus implies

$$c = \frac{2}{\sum_{i=1}^n a_i} \quad (5)$$

### Problem 5.3.1 Solution

Here we solve the following problem:<sup>1</sup>

Given  $f_{\mathbf{X}}(\mathbf{x})$  with  $c = 2/3$  and  $a_1 = a_2 = a_3 = 1$  in Problem 5.2.2, find the marginal PDF  $f_{X_3}(x_3)$ .

---

<sup>1</sup>The wrong problem statement appears in the first printing.

Filling in the parameters in Problem 5.2.2, we obtain the vector PDF

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} \frac{2}{3}(x_1 + x_2 + x_3) & 0 \leq x_1, x_2, x_3 \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

In this case, for  $0 \leq x_3 \leq 1$ , the marginal PDF of  $X_3$  is

$$f_{X_3}(x_3) = \frac{2}{3} \int_0^1 \int_0^1 (x_1 + x_2 + x_3) dx_1 dx_2 \quad (2)$$

$$= \frac{2}{3} \int_0^1 \left( \frac{x_1^2}{2} + x_2 x_1 + x_3 x_1 \right) \Big|_{x_1=0}^{x_1=1} dx_2 \quad (3)$$

$$= \frac{2}{3} \int_0^1 \left( \frac{1}{2} + x_2 + x_3 \right) dx_2 \quad (4)$$

$$= \frac{2}{3} \left( \frac{x_2}{2} + \frac{x_2^2}{2} + x_3 x_2 \right) \Big|_{x_2=0}^{x_2=1} = \frac{2}{3} \left( \frac{1}{2} + \frac{1}{2} + x_3 \right) \quad (5)$$

The complete expression for the marginal PDF of  $X_3$  is

$$f_{X_3}(x_3) = \begin{cases} 2(1 + x_3)/3 & 0 \leq x_3 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

### Problem 5.3.2 Solution

Since  $J_1$ ,  $J_2$  and  $J_3$  are independent, we can write

$$P_{\mathbf{K}}(\mathbf{k}) = P_{J_1}(k_1) P_{J_2}(k_2 - k_1) P_{J_3}(k_3 - k_2) \quad (1)$$

Since  $P_{J_i}(j) > 0$  only for integers  $j > 0$ , we have that  $P_{\mathbf{K}}(\mathbf{k}) > 0$  only for  $0 < k_1 < k_2 < k_3$ ; otherwise  $P_{\mathbf{K}}(\mathbf{k}) = 0$ . Finally, for  $0 < k_1 < k_2 < k_3$ ,

$$P_{\mathbf{K}}(\mathbf{k}) = (1 - p)^{k_1 - 1} p (1 - p)^{k_2 - k_1 - 1} p (1 - p)^{k_3 - k_2 - 1} p \quad (2)$$

$$= (1 - p)^{k_3 - 3} p^3 \quad (3)$$

### Problem 5.3.3 Solution

The joint PMF is

$$P_{\mathbf{K}}(\mathbf{k}) = P_{K_1, K_2, K_3}(k_1, k_2, k_3) = \begin{cases} p^3(1 - p)^{k_3 - 3} & 1 \leq k_1 < k_2 < k_3 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

(a) We start by finding  $P_{K_1, K_2}(k_1, k_2)$ . For  $1 \leq k_1 < k_2$ ,

$$P_{K_1, K_2}(k_1, k_2) = \sum_{k_3=-\infty}^{\infty} P_{K_1, K_2, K_3}(k_1, k_2, k_3) \quad (2)$$

$$= \sum_{k_3=k_2+1}^{\infty} p^3(1 - p)^{k_3 - 3} \quad (3)$$

$$= p^3(1 - p)^{k_2 - 2} (1 + (1 - p) + (1 - p)^2 + \dots) \quad (4)$$

$$= p^2(1 - p)^{k_2 - 2} \quad (5)$$

The complete expression is

$$P_{K_1, K_2}(k_1, k_2) = \begin{cases} p^2(1-p)^{k_2-2} & 1 \leq k_1 < k_2 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Next we find  $P_{K_1, K_3}(k_1, k_3)$ . For  $k_1 \geq 1$  and  $k_3 \geq k_1 + 2$ , we have

$$P_{K_1, K_3}(k_1, k_3) = \sum_{k_2=-\infty}^{\infty} P_{K_1, K_2, K_3}(k_1, k_2, k_3) = \sum_{k_2=k_1+1}^{k_3-1} p^3(1-p)^{k_3-3} \quad (7)$$

$$= (k_3 - k_1 - 1)p^3(1-p)^{k_3-3} \quad (8)$$

The complete expression of the PMF of  $K_1$  and  $K_3$  is

$$P_{K_1, K_3}(k_1, k_3) = \begin{cases} (k_3 - k_1 - 1)p^3(1-p)^{k_3-3} & 1 \leq k_1, k_1 + 2 \leq k_3, \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

The next marginal PMF is

$$P_{K_2, K_3}(k_2, k_3) = \sum_{k_1=-\infty}^{\infty} P_{K_1, K_2, K_3}(k_1, k_2, k_3) = \sum_{k_1=1}^{k_2-1} p^3(1-p)^{k_3-3} \quad (10)$$

$$= (k_2 - 1)p^3(1-p)^{k_3-3} \quad (11)$$

The complete expression of the PMF of  $K_2$  and  $K_3$  is

$$P_{K_2, K_3}(k_2, k_3) = \begin{cases} (k_2 - 1)p^3(1-p)^{k_3-3} & 1 \leq k_2 < k_3, \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

- (b) Going back to first principles, we note that  $K_n$  is the number of trials up to and including the  $n$ th success. Thus  $K_1$  is a geometric ( $p$ ) random variable,  $K_2$  is an Pascal ( $2, p$ ) random variable, and  $K_3$  is an Pascal ( $3, p$ ) random variable. We could write down the respective marginal PMFs of  $K_1$ ,  $K_2$  and  $K_3$  just by looking up the Pascal ( $n, p$ ) PMF. Nevertheless, it is instructive to derive these PMFs from the joint PMF  $P_{K_1, K_2, K_3}(k_1, k_2, k_3)$ .

For  $k_1 \geq 1$ , we can find  $P_{K_1}(k_1)$  via

$$P_{K_1}(k_1) = \sum_{k_2=-\infty}^{\infty} P_{K_1, K_2}(k_1, k_2) = \sum_{k_2=k_1+1}^{\infty} p^2(1-p)^{k_2-2} \quad (13)$$

$$= p^2(1-p)^{k_1-1}[1 + (1-p) + (1-p)^2 + \dots] \quad (14)$$

$$= p(1-p)^{k_1-1} \quad (15)$$

The complete expression for the PMF of  $K_1$  is the usual geometric PMF

$$P_{K_1}(k_1) = \begin{cases} p(1-p)^{k_1-1} & k_1 = 1, 2, \dots, \\ 0 & \text{otherwise.} \end{cases} \quad (16)$$

Following the same procedure, the marginal PMF of  $K_2$  is

$$P_{K_2}(k_2) = \sum_{k_1=-\infty}^{\infty} P_{K_1, K_2}(k_1, k_2) = \sum_{k_1=1}^{k_2-1} p^2(1-p)^{k_2-2} \quad (17)$$

$$= (k_2 - 1)p^2(1-p)^{k_2-2} \quad (18)$$

Since  $P_{K_2}(k_2) = 0$  for  $k_2 < 2$ , the complete PMF is the Pascal  $(2, p)$  PMF

$$P_{K_2}(k_2) = \binom{k_2 - 1}{1} p^2 (1 - p)^{k_2 - 2} \quad (19)$$

Finally, for  $k_3 \geq 3$ , the PMF of  $K_3$  is

$$P_{K_3}(k_3) = \sum_{k_2=-\infty}^{\infty} P_{K_2, K_3}(k_2, k_3) = \sum_{k_2=2}^{k_3-1} (k_2 - 1) p^3 (1 - p)^{k_3 - 3} \quad (20)$$

$$= [1 + 2 + \cdots + (k_3 - 2)] p^3 (1 - p)^{k_3 - 3} \quad (21)$$

$$= \frac{(k_3 - 2)(k_3 - 1)}{2} p^3 (1 - p)^{k_3 - 3} \quad (22)$$

Since  $P_{K_3}(k_3) = 0$  for  $k_3 < 3$ , the complete expression for  $P_{K_3}(k_3)$  is the Pascal  $(3, p)$  PMF

$$P_{K_3}(k_3) = \binom{k_3 - 1}{2} p^3 (1 - p)^{k_3 - 3}. \quad (23)$$

#### Problem 5.3.4 Solution

For  $0 \leq y_1 \leq y_4 \leq 1$ , the marginal PDF of  $Y_1$  and  $Y_4$  satisfies

$$f_{Y_1, Y_4}(y_1, y_4) = \iint f_{\mathbf{Y}}(\mathbf{y}) \, dy_2 \, dy_3 \quad (1)$$

$$= \int_{y_1}^{y_4} \left( \int_{y_2}^{y_4} 24 \, dy_3 \right) \, dy_2 \quad (2)$$

$$= \int_{y_1}^{y_4} 24(y_4 - y_2) \, dy_2 \quad (3)$$

$$= -12(y_4 - y_2)^2 \Big|_{y_2=y_1}^{y_2=y_4} = 12(y_4 - y_1)^2 \quad (4)$$

The complete expression for the joint PDF of  $Y_1$  and  $Y_4$  is

$$f_{Y_1, Y_4}(y_1, y_4) = \begin{cases} 12(y_4 - y_1)^2 & 0 \leq y_1 \leq y_4 \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

For  $0 \leq y_1 \leq y_2 \leq 1$ , the marginal PDF of  $Y_1$  and  $Y_2$  is

$$f_{Y_1, Y_2}(y_1, y_2) = \iint f_{\mathbf{Y}}(\mathbf{y}) \, dy_3 \, dy_4 \quad (6)$$

$$= \int_{y_2}^1 \left( \int_{y_3}^1 24 \, dy_4 \right) \, dy_3 \quad (7)$$

$$= \int_{y_2}^1 24(1 - y_3) \, dy_3 = 12(1 - y_2)^2 \quad (8)$$

The complete expression for the joint PDF of  $Y_1$  and  $Y_2$  is

$$f_{Y_1, Y_2}(y_1, y_2) = \begin{cases} 12(1 - y_2)^2 & 0 \leq y_1 \leq y_2 \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

For  $0 \leq y_1 \leq 1$ , the marginal PDF of  $Y_1$  can be found from

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{Y_1, Y_2}(y_1, y_2) dy_2 = \int_{y_1}^1 12(1 - y_2)^2 dy_2 = 4(1 - y_1)^3 \quad (10)$$

The complete expression of the PDF of  $Y_1$  is

$$f_{Y_1}(y_1) = \begin{cases} 4(1 - y_1)^3 & 0 \leq y_1 \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

Note that the integral  $f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{Y_1, Y_4}(y_1, y_4) dy_4$  would have yielded the same result. This is a good way to check our derivations of  $f_{Y_1, Y_4}(y_1, y_4)$  and  $f_{Y_1, Y_2}(y_1, y_2)$ .

### Problem 5.3.5 Solution

The value of each byte is an independent experiment with 255 possible outcomes. Each byte takes on the value  $b_i$  with probability  $p_i = p = 1/255$ . The joint PMF of  $N_0, \dots, N_{255}$  is the multinomial PMF

$$P_{N_0, \dots, N_{255}}(n_0, \dots, n_{255}) = \frac{10000!}{n_0! n_1! \dots n_{255}!} p^{n_0} p^{n_1} \dots p^{n_{255}} \quad n_0 + \dots + n_{255} = 10000 \quad (1)$$

$$= \frac{10000!}{n_0! n_1! \dots n_{255}!} (1/255)^{10000} \quad n_0 + \dots + n_{255} = 10000 \quad (2)$$

To evaluate the joint PMF of  $N_0$  and  $N_1$ , we define a new experiment with three categories:  $b_0$ ,  $b_1$  and “other.” Let  $\hat{N}$  denote the number of bytes that are “other.” In this case, a byte is in the “other” category with probability  $\hat{p} = 253/255$ . The joint PMF of  $N_0$ ,  $N_1$ , and  $\hat{N}$  is

$$P_{N_0, N_1, \hat{N}}(n_0, n_1, \hat{n}) = \frac{10000!}{n_0! n_1! \hat{n}!} \left(\frac{1}{255}\right)^{n_0} \left(\frac{1}{255}\right)^{n_1} \left(\frac{253}{255}\right)^{\hat{n}} \quad n_0 + n_1 + \hat{n} = 10000 \quad (3)$$

Now we note that the following events are one in the same:

$$\{N_0 = n_0, N_1 = n_1\} = \{N_0 = n_0, N_1 = n_1, \hat{N} = 10000 - n_0 - n_1\} \quad (4)$$

Hence, for non-negative integers  $n_0$  and  $n_1$  satisfying  $n_0 + n_1 \leq 10000$ ,

$$P_{N_0, N_1}(n_0, n_1) = P_{N_0, N_1, \hat{N}}(n_0, n_1, 10000 - n_0 - n_1) \quad (5)$$

$$= \frac{10000!}{n_0! n_1! (10000 - n_0 - n_1)!} \left(\frac{1}{255}\right)^{n_0 + n_1} \left(\frac{253}{255}\right)^{10000 - n_0 - n_1} \quad (6)$$

### Problem 5.3.6 Solution

In Example 5.1, random variables  $N_1, \dots, N_r$  have the multinomial distribution

$$P_{N_1, \dots, N_r}(n_1, \dots, n_r) = \binom{n}{n_1, \dots, n_r} p_1^{n_1} \dots p_r^{n_r} \quad (1)$$

where  $n > r > 2$ .

- (a) To evaluate the joint PMF of  $N_1$  and  $N_2$ , we define a new experiment with mutually exclusive events:  $s_1$ ,  $s_2$  and “other”. Let  $\hat{N}$  denote the number of trial outcomes that are “other”. In this case, a trial is in the “other” category with probability  $\hat{p} = 1 - p_1 - p_2$ . The joint PMF of  $N_1$ ,  $N_2$ , and  $\hat{N}$  is

$$P_{N_1, N_2, \hat{N}}(n_1, n_2, \hat{n}) = \frac{n!}{n_1! n_2! \hat{n}!} p_1^{n_1} p_2^{n_2} (1 - p_1 - p_2)^{\hat{n}} \quad n_1 + n_2 + \hat{n} = n \quad (2)$$

Now we note that the following events are one in the same:

$$\{N_1 = n_1, N_2 = n_2\} = \{N_1 = n_1, N_2 = n_2, \hat{N} = n - n_1 - n_2\} \quad (3)$$

Hence, for non-negative integers  $n_1$  and  $n_2$  satisfying  $n_1 + n_2 \leq n$ ,

$$P_{N_1, N_2}(n_1, n_2) = P_{N_1, N_2, \hat{N}}(n_1, n_2, n - n_1 - n_2) \quad (4)$$

$$= \frac{n!}{n_1! n_2! (n - n_1 - n_2)!} p_1^{n_1} p_2^{n_2} (1 - p_1 - p_2)^{n - n_1 - n_2} \quad (5)$$

- (b) We could find the PMF of  $T_i$  by summing the joint PMF  $P_{N_1, \dots, N_r}(n_1, \dots, n_r)$ . However, it is easier to start from first principles. Suppose we say a success occurs if the outcome of the trial is in the set  $\{s_1, s_2, \dots, s_i\}$  and otherwise a failure occurs. In this case, the success probability is  $q_i = p_1 + \dots + p_i$  and  $T_i$  is the number of successes in  $n$  trials. Thus,  $T_i$  has the binomial PMF

$$P_{T_i}(t) = \begin{cases} \binom{n}{t} q_i^t (1 - q_i)^{n-t} & t = 0, 1, \dots, n \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

- (c) The joint PMF of  $T_1$  and  $T_2$  satisfies

$$P_{T_1, T_2}(t_1, t_2) = P[N_1 = t_1, N_1 + N_2 = t_2] \quad (7)$$

$$= P[N_1 = t_1, N_2 = t_2 - t_1] \quad (8)$$

$$= P_{N_1, N_2}(t_1, t_2 - t_1) \quad (9)$$

By the result of part (a),

$$P_{T_1, T_2}(t_1, t_2) = \frac{n!}{t_1! (t_2 - t_1)! (n - t_2)!} p_1^{t_1} p_2^{t_2 - t_1} (1 - p_1 - p_2)^{n - t_2} \quad 0 \leq t_1 \leq t_2 \leq n \quad (10)$$

### Problem 5.3.7 Solution

- (a) Note that  $Z$  is the number of three-page faxes. In principle, we can sum the joint PMF  $P_{X, Y, Z}(x, y, z)$  over all  $x, y$  to find  $P_Z(z)$ . However, it is better to realize that each fax has 3 pages with probability  $1/6$ , independent of any other fax. Thus,  $Z$  has the binomial PMF

$$P_Z(z) = \begin{cases} \binom{5}{z} (1/6)^z (5/6)^{5-z} & z = 0, 1, \dots, 5 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

- (b) From the properties of the binomial distribution given in Appendix A, we know that  $E[Z] = 5(1/6)$ .



- (c) We want to find the conditional PMF of the number  $X$  of 1-page faxes and number  $Y$  of 2-page faxes given  $Z = 2$  3-page faxes. Note that given  $Z = 2$ ,  $X + Y = 3$ . Hence for non-negative integers  $x, y$  satisfying  $x + y = 3$ ,

$$P_{X,Y|Z}(x, y|2) = \frac{P_{X,Y,Z}(x, y, 2)}{P_Z(2)} = \frac{\frac{5!}{x!y!2!}(1/3)^x(1/2)^y(1/6)^2}{\binom{5}{2}(1/6)^2(5/6)^3} \quad (2)$$

With some algebra, the complete expression of the conditional PMF is

$$P_{X,Y|Z}(x, y|2) = \begin{cases} \frac{3!}{x!y!}(2/5)^x(3/5)^y & x + y = 3, x \geq 0, y \geq 0; x, y \text{ integer} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

In the above expression, we note that if  $Z = 2$ , then  $Y = 3 - X$  and

$$P_{X|Z}(x|2) = P_{X,Y|Z}(x, 3 - x|2) = \begin{cases} \binom{3}{x}(2/5)^x(3/5)^{3-x} & x = 0, 1, 2, 3 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

That is, given  $Z = 2$ , there are 3 faxes left, each of which independently could be a 1-page fax. The conditional PMF of the number of 1-page faxes is binomial where  $2/5$  is the conditional probability that a fax has 1 page given that it either has 1 page or 2 pages. Moreover given  $X = x$  and  $Z = 2$  we must have  $Y = 3 - x$ .

- (d) Given  $Z = 2$ , the conditional PMF of  $X$  is binomial for 3 trials and success probability  $2/5$ . The conditional expectation of  $X$  given  $Z = 2$  is  $E[X|Z = 2] = 3(2/5) = 6/5$ .
- (e) There are several ways to solve this problem. The most straightforward approach is to realize that for integers  $0 \leq x \leq 5$  and  $0 \leq y \leq 5$ , the event  $\{X = x, Y = y\}$  occurs iff  $\{X = x, Y = y, Z = 5 - (x + y)\}$ . For the rest of this problem, we assume  $x$  and  $y$  are non-negative integers so that

$$P_{X,Y}(x, y) = P_{X,Y,Z}(x, y, 5 - (x + y)) \quad (5)$$

$$= \begin{cases} \frac{5!}{x!y!(5-x-y)!} \left(\frac{1}{3}\right)^x \left(\frac{1}{2}\right)^y \left(\frac{1}{6}\right)^{5-x-y} & 0 \leq x + y \leq 5, x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

The above expression may seem unwieldy and it isn't even clear that it will sum to 1. To simplify the expression, we observe that

$$P_{X,Y}(x, y) = P_{X,Y,Z}(x, y, 5 - x - y) = P_{X,Y|Z}(x, y|5 - x - y) P_Z(5 - x - y) \quad (7)$$

Using  $P_Z(z)$  found in part (c), we can calculate  $P_{X,Y|Z}(x, y|5 - x - y)$  for  $0 \leq x + y \leq 5$ , integer valued.

$$P_{X,Y|Z}(x, y|5 - x - y) = \frac{P_{X,Y,Z}(x, y, 5 - x - y)}{P_Z(5 - x - y)} \quad (8)$$

$$= \binom{x+y}{x} \left(\frac{1/3}{1/2 + 1/3}\right)^x \left(\frac{1/2}{1/2 + 1/3}\right)^y \quad (9)$$

$$= \binom{x+y}{x} \left(\frac{2}{5}\right)^x \left(\frac{3}{5}\right)^{(x+y)-x} \quad (10)$$

In the above expression, it is wise to think of  $x + y$  as some fixed value. In that case, we see that given  $x + y$  is a fixed value,  $X$  and  $Y$  have a joint PMF given by a binomial distribution

in  $x$ . This should not be surprising since it is just a generalization of the case when  $Z = 2$ . That is, given that there were a fixed number of faxes that had either one or two pages, each of those faxes is a one page fax with probability  $(1/3)/(1/2 + 1/3)$  and so the number of one page faxes should have a binomial distribution. Moreover, given the number  $X$  of one page faxes, the number  $Y$  of two page faxes is completely specified.

Finally, by rewriting  $P_{X,Y}(x, y)$  given above, the complete expression for the joint PMF of  $X$  and  $Y$  is

$$P_{X,Y}(x, y) = \begin{cases} \binom{5}{5-x-y} \left(\frac{1}{6}\right)^{5-x-y} \left(\frac{5}{6}\right)^{x+y} \binom{x+y}{x} \left(\frac{2}{5}\right)^x \left(\frac{3}{5}\right)^y & x, y \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

### Problem 5.3.8 Solution

In Problem 5.3.2, we found that the joint PMF of  $\mathbf{K} = [K_1 \ K_2 \ K_3]'$  is

$$P_{\mathbf{K}}(\mathbf{k}) = \begin{cases} p^3(1-p)^{k_3-3} & k_1 < k_2 < k_3 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

In this problem, we generalize the result to  $n$  messages.

(a) For  $k_1 < k_2 < \dots < k_n$ , the joint event

$$\{K_1 = k_1, K_2 = k_2, \dots, K_n = k_n\} \quad (2)$$

occurs if and only if all of the following events occur

- $A_1$   $k_1 - 1$  failures, followed by a successful transmission
- $A_2$   $(k_2 - 1) - k_1$  failures followed by a successful transmission
- $A_3$   $(k_3 - 1) - k_2$  failures followed by a successful transmission
- $\vdots$
- $A_n$   $(k_n - 1) - k_{n-1}$  failures followed by a successful transmission

Note that the events  $A_1, A_2, \dots, A_n$  are independent and

$$P[A_j] = (1-p)^{k_j - k_{j-1} - 1} p. \quad (3)$$

Thus

$$P_{K_1, \dots, K_n}(k_1, \dots, k_n) = P[A_1] P[A_2] \dots P[A_n] \quad (4)$$

$$= p^n (1-p)^{(k_1-1) + (k_2-k_1-1) + (k_3-k_2-1) + \dots + (k_n-k_{n-1}-1)} \quad (5)$$

$$= p^n (1-p)^{k_n - n} \quad (6)$$

To clarify subsequent results, it is better to rename  $\mathbf{K}$  as  $\mathbf{K}_n = [K_1 \ K_2 \ \dots \ K_n]'$ . We see that

$$P_{\mathbf{K}_n}(\mathbf{k}_n) = \begin{cases} p^n (1-p)^{k_n - n} & 1 \leq k_1 < k_2 < \dots < k_n, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

(b) For  $j < n$ ,

$$P_{K_1, K_2, \dots, K_j}(k_1, k_2, \dots, k_j) = P_{\mathbf{K}_j}(\mathbf{k}_j). \quad (8)$$

Since  $\mathbf{K}_j$  is just  $\mathbf{K}_n$  with  $n = j$ , we have

$$P_{\mathbf{K}_j}(\mathbf{k}_j) = \begin{cases} p^j(1-p)^{k_j-j} & 1 \leq k_1 < k_2 < \dots < k_j, \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

(c) Rather than try to deduce  $P_{K_i}(k_i)$  from the joint PMF  $P_{\mathbf{K}_n}(\mathbf{k}_n)$ , it is simpler to return to first principles. In particular,  $K_i$  is the number of trials up to and including the  $i$ th success and has the Pascal  $(i, p)$  PMF

$$P_{K_i}(k_i) = \binom{k_i-1}{i-1} p^i (1-p)^{k_i-i}. \quad (10)$$

### Problem 5.4.1 Solution

For  $i \neq j$ ,  $X_i$  and  $X_j$  are independent and  $E[X_i X_j] = E[X_i]E[X_j] = 0$  since  $E[X_i] = 0$ . Thus the  $i, j$ th entry in the covariance matrix  $\mathbf{C}_X$  is

$$C_{\mathbf{X}}(i, j) = E[X_i X_j] = \begin{cases} \sigma_i^2 & i = j, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Thus for random vector  $\mathbf{X} = [X_1 \ X_2 \ \dots \ X_n]'$ , all the off-diagonal entries in the covariance matrix are zero and the covariance matrix is

$$\mathbf{C}_X = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix}. \quad (2)$$

### Problem 5.4.2 Solution

The random variables  $N_1, N_2, N_3$  and  $N_4$  are *dependent*. To see this we observe that  $P_{N_i}(4) = p_i^4$ . However,

$$P_{N_1, N_2, N_3, N_4}(4, 4, 4, 4) = 0 \neq p_1^4 p_2^4 p_3^4 p_4^4 = P_{N_1}(4) P_{N_2}(4) P_{N_3}(4) P_{N_4}(4). \quad (1)$$

### Problem 5.4.3 Solution

We will use the PDF

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} 1 & 0 \leq x_i \leq 1, i = 1, 2, 3, 4 \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

to find the marginal PDFs  $f_{X_i}(x_i)$ . In particular, for  $0 \leq x_1 \leq 1$ ,

$$f_{X_1}(x_1) = \int_0^1 \int_0^1 \int_0^1 f_{\mathbf{X}}(\mathbf{x}) \, dx_2 \, dx_3 \, dx_4 \quad (2)$$

$$= \left( \int_0^1 dx_2 \right) \left( \int_0^1 dx_3 \right) \left( \int_0^1 dx_4 \right) = 1. \quad (3)$$

Thus,

$$f_{X_1}(x_1) = \begin{cases} 1 & 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Following similar steps, one can show that

$$f_{X_1}(x) = f_{X_2}(x) = f_{X_3}(x) = f_{X_4}(x) = \begin{cases} 1 & 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Thus

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x) f_{X_2}(x) f_{X_3}(x) f_{X_4}(x). \quad (6)$$

We conclude that  $X_1$ ,  $X_2$ ,  $X_3$  and  $X_4$  are independent.

#### Problem 5.4.4 Solution

We will use the PDF

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} 6e^{-(x_1+2x_2+3x_3)} & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

to find the marginal PDFs  $f_{X_i}(x_i)$ . In particular, for  $x_1 \geq 0$ ,

$$f_{X_1}(x_1) = \int_0^\infty \int_0^\infty f_{\mathbf{X}}(\mathbf{x}) dx_2 dx_3 \quad (2)$$

$$= 6e^{-x_1} \left( \int_0^\infty e^{-2x_2} dx_2 \right) \left( \int_0^\infty e^{-3x_3} dx_3 \right) \quad (3)$$

$$= 6e^{-x_1} \left( -\frac{1}{2}e^{-2x_2} \Big|_0^\infty \right) \left( -\frac{1}{3}e^{-3x_3} \Big|_0^\infty \right) = e^{-x_1}. \quad (4)$$

Thus,

$$f_{X_1}(x_1) = \begin{cases} e^{-x_1} & x_1 \geq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Following similar steps, one can show that

$$f_{X_2}(x_2) = \int_0^\infty \int_0^\infty f_{\mathbf{X}}(\mathbf{x}) dx_1 dx_3 = \begin{cases} 2^{-2x_2} & x_2 \geq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

$$f_{X_3}(x_3) = \int_0^\infty \int_0^\infty f_{\mathbf{X}}(\mathbf{x}) dx_1 dx_2 = \begin{cases} 3^{-3x_3} & x_3 \geq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

Thus

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x_1) f_{X_2}(x_2) f_{X_3}(x_3). \quad (8)$$

We conclude that  $X_1$ ,  $X_2$ , and  $X_3$  are independent.

#### Problem 5.4.5 Solution

This problem can be solved without any real math. Some thought should convince you that for any  $x_i > 0$ ,  $f_{X_i}(x_i) > 0$ . Thus,  $f_{X_1}(10) > 0$ ,  $f_{X_2}(9) > 0$ , and  $f_{X_3}(8) > 0$ . Thus  $f_{X_1}(10)f_{X_2}(9)f_{X_3}(8) > 0$ . However, from the definition of the joint PDF

$$f_{X_1, X_2, X_3}(10, 9, 8) = 0 \neq f_{X_1}(10) f_{X_2}(9) f_{X_3}(8). \quad (1)$$

It follows that  $X_1$ ,  $X_2$  and  $X_3$  are dependent. Readers who find this quick answer dissatisfying are invited to confirm this conclusions by solving Problem 5.4.6 for the exact expressions for the marginal PDFs  $f_{X_1}(x_1)$ ,  $f_{X_2}(x_2)$ , and  $f_{X_3}(x_3)$ .

### Problem 5.4.6 Solution

We find the marginal PDFs using Theorem 5.5. First we note that for  $x < 0$ ,  $f_{X_i}(x) = 0$ . For  $x_1 \geq 0$ ,

$$f_{X_1}(x_1) = \int_{x_1}^{\infty} \left( \int_{x_2}^{\infty} e^{-x_3} dx_3 \right) dx_2 = \int_{x_1}^{\infty} e^{-x_2} dx_2 = e^{-x_1} \quad (1)$$

Similarly, for  $x_2 \geq 0$ ,  $X_2$  has marginal PDF

$$f_{X_2}(x_2) = \int_0^{x_2} \left( \int_{x_2}^{\infty} e^{-x_3} dx_3 \right) dx_1 = \int_0^{x_2} e^{-x_2} dx_1 = x_2 e^{-x_2} \quad (2)$$

Lastly,

$$f_{X_3}(x_3) = \int_0^{x_3} \left( \int_{x_1}^{x_3} e^{-x_3} dx_2 \right) dx_1 = \int_0^{x_3} (x_3 - x_1) e^{-x_3} dx_1 \quad (3)$$

$$= -\frac{1}{2}(x_3 - x_1)^2 e^{-x_3} \Big|_{x_1=0}^{x_1=x_3} = \frac{1}{2} x_3^2 e^{-x_3} \quad (4)$$

The complete expressions for the three marginal PDFs are

$$f_{X_1}(x_1) = \begin{cases} e^{-x_1} & x_1 \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

$$f_{X_2}(x_2) = \begin{cases} x_2 e^{-x_2} & x_2 \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$$f_{X_3}(x_3) = \begin{cases} (1/2)x_3^2 e^{-x_3} & x_3 \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

In fact, each  $X_i$  is an Erlang  $(n, \lambda) = (i, 1)$  random variable.

### Problem 5.4.7 Solution

Since  $U_1, \dots, U_n$  are iid uniform  $(0, 1)$  random variables,

$$f_{U_1, \dots, U_n}(u_1, \dots, u_n) = \begin{cases} 1/T^n & 0 \leq u_i \leq 1; i = 1, 2, \dots, n \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Since  $U_1, \dots, U_n$  are continuous,  $P[U_i = U_j] = 0$  for all  $i \neq j$ . For the same reason,  $P[X_i = X_j] = 0$  for  $i \neq j$ . Thus we need only to consider the case when  $x_1 < x_2 < \dots < x_n$ .

To understand the claim, it is instructive to start with the  $n = 2$  case. In this case,  $(X_1, X_2) = (x_1, x_2)$  (with  $x_1 < x_2$ ) if either  $(U_1, U_2) = (x_1, x_2)$  or  $(U_1, U_2) = (x_2, x_1)$ . For infinitesimal  $\Delta$ ,

$$f_{X_1, X_2}(x_1, x_2) \Delta^2 = P[x_1 < X_1 \leq x_1 + \Delta, x_2 < X_2 \leq x_2 + \Delta] \quad (2)$$

$$= P[x_1 < U_1 \leq x_1 + \Delta, x_2 < U_2 \leq x_2 + \Delta] \\ + P[x_2 < U_1 \leq x_2 + \Delta, x_1 < U_2 \leq x_1 + \Delta] \quad (3)$$

$$= f_{U_1, U_2}(x_1, x_2) \Delta^2 + f_{U_1, U_2}(x_2, x_1) \Delta^2 \quad (4)$$

We see that for  $0 \leq x_1 < x_2 \leq 1$  that

$$f_{X_1, X_2}(x_1, x_2) = 2/T^n. \quad (5)$$

For the general case of  $n$  uniform random variables, we define  $\boldsymbol{\pi} = [\pi(1) \ \dots \ \pi(n)]'$  as a permutation vector of the integers  $1, 2, \dots, n$  and  $\Pi$  as the set of  $n!$  possible permutation vectors. In this case, the event  $\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\}$  occurs if

$$U_1 = x_{\pi(1)}, U_2 = x_{\pi(2)}, \dots, U_n = x_{\pi(n)} \quad (6)$$

for any permutation  $\boldsymbol{\pi} \in \Pi$ . Thus, for  $0 \leq x_1 < x_2 < \dots < x_n \leq 1$ ,

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) \Delta^n = \sum_{\boldsymbol{\pi} \in \Pi} f_{U_1, \dots, U_n}(x_{\pi(1)}, \dots, x_{\pi(n)}) \Delta^n. \quad (7)$$

Since there are  $n!$  permutations and  $f_{U_1, \dots, U_n}(x_{\pi(1)}, \dots, x_{\pi(n)}) = 1/T^n$  for each permutation  $\boldsymbol{\pi}$ , we can conclude that

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = n!/T^n. \quad (8)$$

Since the order statistics are necessarily ordered,  $f_{X_1, \dots, X_n}(x_1, \dots, x_n) = 0$  unless  $x_1 < \dots < x_n$ .

### Problem 5.5.1 Solution

For discrete random vectors, it is true in general that

$$P_{\mathbf{Y}}(\mathbf{y}) = P[\mathbf{Y} = \mathbf{y}] = P[\mathbf{A}\mathbf{X} + \mathbf{b} = \mathbf{y}] = P[\mathbf{A}\mathbf{X} = \mathbf{y} - \mathbf{b}]. \quad (1)$$

For an arbitrary matrix  $\mathbf{A}$ , the system of equations  $\mathbf{A}\mathbf{x} = \mathbf{y} - \mathbf{b}$  may have no solutions (if the columns of  $\mathbf{A}$  do not span the vector space), multiple solutions (if the columns of  $\mathbf{A}$  are linearly dependent), or, when  $\mathbf{A}$  is invertible, exactly one solution. In the invertible case,

$$P_{\mathbf{Y}}(\mathbf{y}) = P[\mathbf{A}\mathbf{X} = \mathbf{y} - \mathbf{b}] = P[\mathbf{X} = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})] = P_{\mathbf{X}}(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})). \quad (2)$$

As an aside, we note that when  $\mathbf{A}\mathbf{x} = \mathbf{y} - \mathbf{b}$  has multiple solutions, we would need to do some bookkeeping to add up the probabilities  $P_{\mathbf{X}}(\mathbf{x})$  for all vectors  $\mathbf{x}$  satisfying  $\mathbf{A}\mathbf{x} = \mathbf{y} - \mathbf{b}$ . This can get disagreeably complicated.

### Problem 5.5.2 Solution

The random variable  $J_n$  is the number of times that message  $n$  is transmitted. Since each transmission is a success with probability  $p$ , independent of any other transmission, the number of transmissions of message  $n$  is independent of the number of transmissions of message  $m$ . That is, for  $m \neq n$ ,  $J_m$  and  $J_n$  are independent random variables. Moreover, because each message is transmitted over and over until it is transmitted successfully, each  $J_m$  is a geometric ( $p$ ) random variable with PMF

$$P_{J_m}(j) = \begin{cases} (1-p)^{j-1}p & j = 1, 2, \dots \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Thus the PMF of  $\mathbf{J} = [J_1 \ J_2 \ J_3]'$  is

$$P_{\mathbf{J}}(\mathbf{j}) = P_{J_1}(j_1) P_{J_2}(j_2) P_{J_3}(j_3) = \begin{cases} p^3(1-p)^{j_1+j_2+j_3-3} & j_i = 1, 2, \dots; \\ & i = 1, 2, 3 \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

### Problem 5.5.3 Solution

The response time  $X_i$  of the  $i$ th truck has PDF  $f_{X_i}(x_i)$  and CDF  $F_{X_i}(x_i)$  given by

$$f_{X_i}(x_i) = \begin{cases} \frac{1}{2}e^{-x/2} & x \geq 0, \\ 0 & \text{otherwise,} \end{cases} \quad F_{X_i}(x_i) = F_X(x_i) = \begin{cases} 1 - e^{-x/2} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Let  $R = \max(X_1, X_2, \dots, X_6)$  denote the maximum response time. From Theorem 5.7,  $R$  has PDF

$$F_R(r) = (F_X(r))^6. \quad (2)$$

(a) The probability that all six responses arrive within five seconds is

$$P[R \leq 5] = F_R(5) = (F_X(5))^6 = (1 - e^{-5/2})^6 = 0.5982. \quad (3)$$

(b) This question is worded in a somewhat confusing way. The “expected response time” refers to  $E[X_i]$ , the response time of an individual truck, rather than  $E[R]$ . If the expected response time of a truck is  $\tau$ , then each  $X_i$  has CDF

$$F_{X_i}(x) = F_X(x) = \begin{cases} 1 - e^{-x/\tau} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

The goal of this problem is to find the maximum permissible value of  $\tau$ . When each truck has expected response time  $\tau$ , the CDF of  $R$  is

$$F_R(r) = (F_X(x) r)^6 = \begin{cases} (1 - e^{-r/\tau})^6 & r \geq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

We need to find  $\tau$  such that

$$P[R \leq 3] = (1 - e^{-3/\tau})^6 = 0.9. \quad (6)$$

This implies

$$\tau = \frac{-3}{\ln(1 - (0.9)^{1/6})} = 0.7406 \text{ s.} \quad (7)$$

### Problem 5.5.4 Solution

Let  $X_i$  denote the finishing time of boat  $i$ . Since finishing times of all boats are iid Gaussian random variables with expected value 35 minutes and standard deviation 5 minutes, we know that each  $X_i$  has CDF

$$F_{X_i}(x) = P[X_i \leq x] = P\left[\frac{X_i - 35}{5} \leq \frac{x - 35}{5}\right] = \Phi\left(\frac{x - 35}{5}\right) \quad (1)$$

(a) The time of the winning boat is

$$W = \min(X_1, X_2, \dots, X_{10}) \quad (2)$$

To find the probability that  $W \leq 25$ , we will find the CDF  $F_W(w)$  since this will also be useful for part (c).

$$F_W(w) = P[\min(X_1, X_2, \dots, X_{10}) \leq w] \quad (3)$$

$$= 1 - P[\min(X_1, X_2, \dots, X_{10}) > w] \quad (4)$$

$$= 1 - P[X_1 > w, X_2 > w, \dots, X_{10} > w] \quad (5)$$

Since the  $X_i$  are iid,

$$F_W(w) = 1 - \prod_{i=1}^{10} P[X_i > w] = 1 - (1 - F_{X_i}(w))^{10} \quad (6)$$

$$= 1 - \left(1 - \Phi\left(\frac{w - 35}{5}\right)\right)^{10} \quad (7)$$

Thus,

$$P[W \leq 25] = F_W(25) = 1 - (1 - \Phi(-2))^{10} \quad (8)$$

$$= 1 - [\Phi(2)]^{10} = 0.2056. \quad (9)$$

- (b) The finishing time of the last boat is  $L = \max(X_1, \dots, X_{10})$ . The probability that the last boat finishes in more than 50 minutes is

$$P[L > 50] = 1 - P[L \leq 50] \quad (10)$$

$$= 1 - P[X_1 \leq 50, X_2 \leq 50, \dots, X_{10} \leq 50] \quad (11)$$

Once again, since the  $X_i$  are iid Gaussian (35, 5) random variables,

$$P[L > 50] = 1 - \prod_{i=1}^{10} P[X_i \leq 50] = 1 - (F_{X_i}(50))^{10} \quad (12)$$

$$= 1 - (\Phi([50 - 35]/5))^{10} \quad (13)$$

$$= 1 - (\Phi(3))^{10} = 0.0134 \quad (14)$$

- (c) A boat will finish in negative time if and only iff the winning boat finishes in negative time, which has probability

$$F_W(0) = 1 - (1 - \Phi(-35/5))^{10} = 1 - (1 - \Phi(-7))^{10} = 1 - (\Phi(7))^{10}. \quad (15)$$

Unfortunately, the tables in the text have neither  $\Phi(7)$  nor  $Q(7)$ . However, those with access to MATLAB, or a programmable calculator, can find out that  $Q(7) = 1 - \Phi(7) = 1.28 \times 10^{-12}$ . This implies that a boat finishes in negative time with probability

$$F_W(0) = 1 - (1 - 1.28 \times 10^{-12})^{10} = 1.28 \times 10^{-11}. \quad (16)$$

### Problem 5.5.5 Solution

Since 50 cents of each dollar ticket is added to the jackpot,

$$J_{i-1} = J_i + \frac{N_i}{2} \quad (1)$$

Given  $J_i = j$ ,  $N_i$  has a Poisson distribution with mean  $j$ . It follows that  $E[N_i | J_i = j] = j$  and that  $\text{Var}[N_i | J_i = j] = j$ . This implies

$$E[N_i^2 | J_i = j] = \text{Var}[N_i | J_i = j] + (E[N_i | J_i = j])^2 = j + j^2 \quad (2)$$



In terms of the conditional expectations given  $J_i$ , these facts can be written as

$$E[N_i|J_i] = J_i \quad E[N_i^2|J_i] = J_i + J_i^2 \quad (3)$$

This permits us to evaluate the moments of  $J_{i-1}$  in terms of the moments of  $J_i$ . Specifically,

$$E[J_{i-1}|J_i] = E[J_i|J_i] + \frac{1}{2}E[N_i|J_i] = J_i + \frac{J_i}{2} = \frac{3J_i}{2} \quad (4)$$

This implies

$$E[J_{i-1}] = \frac{3}{2}E[J_i] \quad (5)$$

We can use this to calculate  $E[J_i]$  for all  $i$ . Since the jackpot starts at 1 million dollars,  $J_6 = 10^6$  and  $E[J_6] = 10^6$ . This implies

$$E[J_i] = (3/2)^{6-i}10^6 \quad (6)$$

Now we will find the second moment  $E[J_i^2]$ . Since  $J_{i-1}^2 = J_i^2 + N_iJ_i + N_i^2/4$ , we have

$$E[J_{i-1}^2|J_i] = E[J_i^2|J_i] + E[N_iJ_i|J_i] + E[N_i^2|J_i]/4 \quad (7)$$

$$= J_i^2 + J_iE[N_i|J_i] + (J_i + J_i^2)/4 \quad (8)$$

$$= (3/2)^2J_i^2 + J_i/4 \quad (9)$$

By taking the expectation over  $J_i$  we have

$$E[J_{i-1}^2] = (3/2)^2E[J_i^2] + E[J_i]/4 \quad (10)$$

This recursion allows us to calculate  $E[J_i^2]$  for  $i = 6, 5, \dots, 0$ . Since  $J_6 = 10^6$ ,  $E[J_6^2] = 10^{12}$ . From the recursion, we obtain

$$E[J_5^2] = (3/2)^2E[J_6^2] + E[J_6]/4 = (3/2)^210^{12} + \frac{1}{4}10^6 \quad (11)$$

$$E[J_4^2] = (3/2)^2E[J_5^2] + E[J_5]/4 = (3/2)^410^{12} + \frac{1}{4}[(3/2)^2 + (3/2)]10^6 \quad (12)$$

$$E[J_3^2] = (3/2)^2E[J_4^2] + E[J_4]/4 = (3/2)^610^{12} + \frac{1}{4}[(3/2)^4 + (3/2)^3 + (3/2)^2]10^6 \quad (13)$$

The same recursion will also allow us to show that

$$E[J_2^2] = (3/2)^810^{12} + \frac{1}{4}[(3/2)^6 + (3/2)^5 + (3/2)^4 + (3/2)^3]10^6 \quad (14)$$

$$E[J_1^2] = (3/2)^{10}10^{12} + \frac{1}{4}[(3/2)^8 + (3/2)^7 + (3/2)^6 + (3/2)^5 + (3/2)^4]10^6 \quad (15)$$

$$E[J_0^2] = (3/2)^{12}10^{12} + \frac{1}{4}[(3/2)^{10} + (3/2)^9 + \dots + (3/2)^5]10^6 \quad (16)$$

Finally, day 0 is the same as any other day in that  $J = J_0 + N_0/2$  where  $N_0$  is a Poisson random variable with mean  $J_0$ . By the same argument that we used to develop recursions for  $E[J_i]$  and  $E[J_i^2]$ , we can show

$$E[J] = (3/2)E[J_0] = (3/2)^710^6 \approx 17 \times 10^6 \quad (17)$$

and

$$E[J^2] = (3/2)^2E[J_0^2] + E[J_0]/4 \quad (18)$$

$$= (3/2)^{14}10^{12} + \frac{1}{4}[(3/2)^{12} + (3/2)^{11} + \dots + (3/2)^6]10^6 \quad (19)$$

$$= (3/2)^{14}10^{12} + \frac{10^6}{2}(3/2)^6[(3/2)^7 - 1] \quad (20)$$

Finally, the variance of  $J$  is

$$\text{Var}[J] = E[J^2] - (E[J])^2 = \frac{10^6}{2}(3/2)^6[(3/2)^7 - 1] \quad (21)$$

Since the variance is hard to interpret, we note that the standard deviation of  $J$  is  $\sigma_J \approx 9572$ . Although the expected jackpot grows rapidly, the standard deviation of the jackpot is fairly small.

### Problem 5.5.6 Solution

Let  $A$  denote the event  $X_n = \max(X_1, \dots, X_n)$ . We can find  $P[A]$  by conditioning on the value of  $X_n$ .

$$P[A] = P[X_1 \leq X_n, X_2 \leq X_n, \dots, X_{n-1} \leq X_n] \quad (1)$$

$$= \int_{-\infty}^{\infty} P[X_1 < X_n, X_2 < X_n, \dots, X_{n-1} < X_n | X_n = x] f_{X_n}(x) dx \quad (2)$$

$$= \int_{-\infty}^{\infty} P[X_1 < x, X_2 < x, \dots, X_{n-1} < x | X_n = x] f_X(x) dx \quad (3)$$

Since  $X_1, \dots, X_{n-1}$  are independent of  $X_n$ ,

$$P[A] = \int_{-\infty}^{\infty} P[X_1 < x, X_2 < x, \dots, X_{n-1} < x] f_X(x) dx. \quad (4)$$

Since  $X_1, \dots, X_{n-1}$  are iid,

$$P[A] = \int_{-\infty}^{\infty} P[X_1 \leq x] P[X_2 \leq x] \cdots P[X_{n-1} \leq x] f_X(x) dx \quad (5)$$

$$= \int_{-\infty}^{\infty} [F_X(x)]^{n-1} f_X(x) dx = \frac{1}{n} [F_X(x)]^n \Big|_{-\infty}^{\infty} = \frac{1}{n} (1 - 0) \quad (6)$$

Not surprisingly, since the  $X_i$  are identical, symmetry would suggest that  $X_n$  is as likely as any of the other  $X_i$  to be the largest. Hence  $P[A] = 1/n$  should not be surprising.

### Problem 5.6.1 Solution

(a) The covariance matrix of  $\mathbf{X} = [X_1 \ X_2]'$  is

$$\mathbf{C}_X = \begin{bmatrix} \text{Var}[X_1] & \text{Cov}[X_1, X_2] \\ \text{Cov}[X_1, X_2] & \text{Var}[X_2] \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 3 & 9 \end{bmatrix}. \quad (1)$$

(b) From the problem statement,

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} \mathbf{X} = \mathbf{A}\mathbf{X}. \quad (2)$$

By Theorem 5.13,  $\mathbf{Y}$  has covariance matrix

$$\mathbf{C}_Y = \mathbf{A}\mathbf{C}_X\mathbf{A}' = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 28 & -66 \\ -66 & 252 \end{bmatrix}. \quad (3)$$

### Problem 5.6.2 Solution

The mean value of a sum of random variables is always the sum of their individual means.

$$E[Y] = \sum_{i=1}^n E[X_i] = 0 \quad (1)$$

The variance of any sum of random variables can be expressed in terms of the individual variances and co-variances. Since the  $E[Y]$  is zero,  $\text{Var}[Y] = E[Y^2]$ . Thus,

$$\text{Var}[Y] = E \left[ \left( \sum_{i=1}^n X_i \right)^2 \right] = E \left[ \sum_{i=1}^n \sum_{j=1}^n X_i X_j \right] = \sum_{i=1}^n E[X_i^2] + \sum_{i=1}^n \sum_{j \neq i} E[X_i X_j] \quad (2)$$

Since  $E[X_i] = 0$ ,  $E[X_i^2] = \text{Var}[X_i] = 1$  and for  $i \neq j$ ,

$$E[X_i X_j] = \text{Cov}[X_i, X_j] = \rho \quad (3)$$

Thus,  $\text{Var}[Y] = n + n(n-1)\rho$ .

### Problem 5.6.3 Solution

Since  $\mathbf{X}$  and  $\mathbf{Y}$  are independent and  $E[Y_j] = 0$  for all components  $Y_j$ , we observe that  $E[X_i Y_j] = E[X_i]E[Y_j] = 0$ . This implies that the cross-covariance matrix is

$$E[\mathbf{X}\mathbf{Y}'] = E[\mathbf{X}]E[\mathbf{Y}'] = \mathbf{0}. \quad (1)$$

### Problem 5.6.4 Solution

Inspection of the vector PDF  $f_{\mathbf{X}}(\mathbf{x})$  will show that  $X_1, X_2, X_3$ , and  $X_4$  are iid uniform  $(0, 1)$  random variables. That is,

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x_1) f_{X_2}(x_2) f_{X_3}(x_3) f_{X_4}(x_4) \quad (1)$$

where each  $X_i$  has the uniform  $(0, 1)$  PDF

$$f_{X_i}(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

It follows that for each  $i$ ,  $E[X_i] = 1/2$ ,  $E[X_i^2] = 1/3$  and  $\text{Var}[X_i] = 1/12$ . In addition,  $X_i$  and  $X_j$  have correlation

$$E[X_i X_j] = E[X_i]E[X_j] = 1/4. \quad (3)$$

and covariance  $\text{Cov}[X_i, X_j] = 0$  for  $i \neq j$  since independent random variables always have zero covariance.

(a) The expected value vector is

$$E[\mathbf{X}] = [E[X_1] \ E[X_2] \ E[X_3] \ E[X_4]]' = [1/2 \ 1/2 \ 1/2 \ 1/2]'. \quad (4)$$

(b) The correlation matrix is

$$\mathbf{R}_X = E[\mathbf{X}\mathbf{X}'] = \begin{bmatrix} E[X_1^2] & E[X_1X_2] & E[X_1X_3] & E[X_1X_4] \\ E[X_2X_1] & E[X_2^2] & E[X_2X_3] & E[X_2X_4] \\ E[X_3X_1] & E[X_3X_2] & E[X_3^2] & E[X_3X_4] \\ E[X_4X_1] & E[X_4X_2] & E[X_4X_3] & E[X_4^2] \end{bmatrix} \quad (5)$$

$$= \begin{bmatrix} 1/3 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/3 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/3 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/3 \end{bmatrix} \quad (6)$$

(c) The covariance matrix for  $\mathbf{X}$  is the diagonal matrix

$$\mathbf{C}_X = \begin{bmatrix} \text{Var}[X_1] & \text{Cov}[X_1, X_2] & \text{Cov}[X_1, X_3] & \text{Cov}[X_1, X_4] \\ \text{Cov}[X_2, X_1] & \text{Var}[X_2] & \text{Cov}[X_2, X_3] & \text{Cov}[X_2, X_4] \\ \text{Cov}[X_3, X_1] & \text{Cov}[X_3, X_2] & \text{Var}[X_3] & \text{Cov}[X_3, X_4] \\ \text{Cov}[X_4, X_1] & \text{Cov}[X_4, X_2] & \text{Cov}[X_4, X_3] & \text{Var}[X_4] \end{bmatrix} \quad (7)$$

$$= \begin{bmatrix} 1/12 & 0 & 0 & 0 \\ 0 & 1/12 & 0 & 0 \\ 0 & 0 & 1/12 & 0 \\ 0 & 0 & 0 & 1/12 \end{bmatrix} \quad (8)$$

Note that its easy to verify that  $\mathbf{C}_X = \mathbf{R}_X - \boldsymbol{\mu}_X \boldsymbol{\mu}_X'$ .

### Problem 5.6.5 Solution

The random variable  $J_m$  is the number of times that message  $m$  is transmitted. Since each transmission is a success with probability  $p$ , independent of any other transmission,  $J_1$ ,  $J_2$  and  $J_3$  are iid geometric ( $p$ ) random variables with

$$E[J_m] = \frac{1}{p}, \quad \text{Var}[J_m] = \frac{1-p}{p^2}. \quad (1)$$

Thus the vector  $\mathbf{J} = [J_1 \ J_2 \ J_3]'$  has expected value

$$E[\mathbf{J}] = [E[J_1] \ E[J_2] \ E[J_3]]' = [1/p \ 1/p \ 1/p]'. \quad (2)$$

For  $m \neq n$ , the correlation matrix  $\mathbf{R}_J$  has  $m, n$ th entry

$$R_J(m, n) = E[J_m J_n] = E[J_m] J_n = 1/p^2 \quad (3)$$

For  $m = n$ ,

$$R_J(m, m) = E[J_m^2] = \text{Var}[J_m] + (E[J_m])^2 = \frac{1-p}{p^2} + \frac{1}{p^2} = \frac{2-p}{p^2}. \quad (4)$$

Thus

$$\mathbf{R}_J = \frac{1}{p^2} \begin{bmatrix} 2-p & 1 & 1 \\ 1 & 2-p & 1 \\ 1 & 1 & 2-p \end{bmatrix}. \quad (5)$$

Because  $J_m$  and  $J_n$  are independent, off-diagonal terms in the covariance matrix are

$$C_J(m, n) = \text{Cov}[J_m, J_n] = 0 \quad (6)$$

Since  $C_{\mathbf{J}}(m, m) = \text{Var}[J_m]$ , we have that

$$\mathbf{C}_{\mathbf{J}} = \frac{1-p}{p^2} \mathbf{I} = \frac{1-p}{p^2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (7)$$

### Problem 5.6.6 Solution

This problem is quite difficult unless one uses the observation that the vector  $\mathbf{K}$  can be expressed in terms of the vector  $\mathbf{J} = [J_1 \ J_2 \ J_3]'$  where  $J_i$  is the number of transmissions of message  $i$ . Note that we can write

$$\mathbf{K} = \mathbf{A}\mathbf{J} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \mathbf{J} \quad (1)$$

We also observe that since each transmission is an independent Bernoulli trial with success probability  $p$ , the components of  $\mathbf{J}$  are iid geometric ( $p$ ) random variables. Thus  $E[J_i] = 1/p$  and  $\text{Var}[J_i] = (1-p)/p^2$ . Thus  $\mathbf{J}$  has expected value

$$E[\mathbf{J}] = \boldsymbol{\mu}_{\mathbf{J}} = [E[J_1] \ E[J_2] \ E[J_3]]' = [1/p \ 1/p \ 1/p]'. \quad (2)$$

Since the components of  $\mathbf{J}$  are independent, it has the diagonal covariance matrix

$$\mathbf{C}_{\mathbf{J}} = \begin{bmatrix} \text{Var}[J_1] & 0 & 0 \\ 0 & \text{Var}[J_2] & 0 \\ 0 & 0 & \text{Var}[J_3] \end{bmatrix} = \frac{1-p}{p^2} \mathbf{I} \quad (3)$$

Given these properties of  $\mathbf{J}$ , finding the same properties of  $\mathbf{K} = \mathbf{A}\mathbf{J}$  is simple.

(a) The expected value of  $\mathbf{K}$  is

$$E[\mathbf{K}] = \mathbf{A}\boldsymbol{\mu}_{\mathbf{J}} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1/p \\ 1/p \\ 1/p \end{bmatrix} = \begin{bmatrix} 1/p \\ 2/p \\ 3/p \end{bmatrix} \quad (4)$$

(b) From Theorem 5.13, the covariance matrix of  $\mathbf{K}$  is

$$\mathbf{C}_{\mathbf{K}} = \mathbf{A}\mathbf{C}_{\mathbf{J}}\mathbf{A}' \quad (5)$$

$$= \frac{1-p}{p^2} \mathbf{A}\mathbf{I}\mathbf{A}' \quad (6)$$

$$= \frac{1-p}{p^2} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \frac{1-p}{p^2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \quad (7)$$

(c) Given the expected value vector  $\boldsymbol{\mu}_{\mathbf{K}}$  and the covariance matrix  $\mathbf{C}_{\mathbf{K}}$ , we can use Theorem 5.12

to find the correlation matrix

$$\mathbf{R}_K = \mathbf{C}_K + \boldsymbol{\mu}_K \boldsymbol{\mu}'_K \quad (8)$$

$$= \frac{1-p}{p^2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 1/p \\ 2/p \\ 3/p \end{bmatrix} \begin{bmatrix} 1/p & 2/p & 3/p \end{bmatrix} \quad (9)$$

$$= \frac{1-p}{p^2} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} + \frac{1}{p^2} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} \quad (10)$$

$$= \frac{1}{p^2} \begin{bmatrix} 2-p & 3-p & 4-p \\ 3-p & 6-2p & 8-2p \\ 4-p & 8-2p & 12-3p \end{bmatrix} \quad (11)$$

### Problem 5.6.7 Solution

The preliminary work for this problem appears in a few different places. In Example 5.5, we found the marginal PDF of  $Y_3$  and in Example 5.6, we found the marginal PDFs of  $Y_1$ ,  $Y_2$ , and  $Y_4$ . We summarize these results here:

$$f_{Y_1}(y) = f_{Y_3}(y) = \begin{cases} 2(1-y) & 0 \leq y \leq 1, \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

$$f_{Y_2}(y) = f_{Y_4}(y) = \begin{cases} 2y & 0 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

This implies

$$E[Y_1] = E[Y_3] = \int_0^1 2y(1-y) dy = 1/3 \quad (3)$$

$$E[Y_2] = E[Y_4] = \int_0^1 2y^2 dy = 2/3 \quad (4)$$

Thus  $\mathbf{Y}$  has expected value  $E[\mathbf{Y}] = [1/3 \ 2/3 \ 1/3 \ 2/3]'$ . The second part of the problem is to find the correlation matrix  $\mathbf{R}_Y$ . In fact, we need to find  $R_Y(i, j) = E[Y_i Y_j]$  for each  $i, j$  pair. We will see that these are seriously tedious calculations. For  $i = j$ , the second moments are

$$E[Y_1^2] = E[Y_3^2] = \int_0^1 2y^2(1-y) dy = 1/6, \quad (5)$$

$$E[Y_2^2] = E[Y_4^2] = \int_0^1 2y^3 dy = 1/2. \quad (6)$$

In terms of the correlation matrix,

$$R_Y(1, 1) = R_Y(3, 3) = 1/6, \quad R_Y(2, 2) = R_Y(4, 4) = 1/2. \quad (7)$$

To find the off diagonal terms  $R_Y(i, j) = E[Y_i Y_j]$ , we need to find the marginal PDFs  $f_{Y_i, Y_j}(y_i, y_j)$ . Example 5.5 showed that

$$f_{Y_1, Y_4}(y_1, y_4) = \begin{cases} 4(1-y_1)y_4 & 0 \leq y_1 \leq 1, 0 \leq y_4 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

$$f_{Y_2, Y_3}(y_2, y_3) = \begin{cases} 4y_2(1-y_3) & 0 \leq y_2 \leq 1, 0 \leq y_3 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

Inspection will show that  $Y_1$  and  $Y_4$  are independent since  $f_{Y_1, Y_4}(y_1, y_4) = f_{Y_1}(y_1)f_{Y_4}(y_4)$ . Similarly,  $Y_2$  and  $Y_3$  are independent since  $f_{Y_2, Y_3}(y_2, y_3) = f_{Y_2}(y_2)f_{Y_3}(y_3)$ . This implies

$$R_{\mathbf{Y}}(1, 4) = E[Y_1 Y_4] = E[Y_1] E[Y_4] = 2/9 \quad (10)$$

$$R_{\mathbf{Y}}(2, 3) = E[Y_2 Y_3] = E[Y_2] E[Y_3] = 2/9 \quad (11)$$

We also need to calculate  $f_{Y_1, Y_2}(y_1, y_2)$ ,  $f_{Y_3, Y_4}(y_3, y_4)$ ,  $f_{Y_1, Y_3}(y_1, y_3)$  and  $f_{Y_2, Y_4}(y_2, y_4)$ . To start, for  $0 \leq y_1 \leq y_2 \leq 1$ ,

$$f_{Y_1, Y_2}(y_1, y_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{Y_1, Y_2, Y_3, Y_4}(y_1, y_2, y_3, y_4) dy_3 dy_4 \quad (12)$$

$$= \int_0^1 \int_0^{y_4} 4 dy_3 dy_4 = \int_0^1 4y_4 dy_4 = 2. \quad (13)$$

Similarly, for  $0 \leq y_3 \leq y_4 \leq 1$ ,

$$f_{Y_3, Y_4}(y_3, y_4) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{Y_1, Y_2, Y_3, Y_4}(y_1, y_2, y_3, y_4) dy_1 dy_2 \quad (14)$$

$$= \int_0^1 \int_0^{y_2} 4 dy_1 dy_2 = \int_0^1 4y_2 dy_2 = 2. \quad (15)$$

In fact, these PDFs are the same in that

$$f_{Y_1, Y_2}(x, y) = f_{Y_3, Y_4}(x, y) = \begin{cases} 2 & 0 \leq x \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (16)$$

This implies  $R_{\mathbf{Y}}(1, 2) = R_{\mathbf{Y}}(3, 4) = E[Y_3 Y_4]$  and that

$$E[Y_3 Y_4] = \int_0^1 \int_0^y 2xy dx dy = \int_0^1 (yx^2|_0^y) dy = \int_0^1 y^3 dy = \frac{1}{4}. \quad (17)$$

Continuing in the same way, we see for  $0 \leq y_1 \leq 1$  and  $0 \leq y_3 \leq 1$  that

$$f_{Y_1, Y_3}(y_1, y_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{Y_1, Y_2, Y_3, Y_4}(y_1, y_2, y_3, y_4) dy_2 dy_4 \quad (18)$$

$$= 4 \left( \int_{y_1}^1 dy_2 \right) \left( \int_{y_3}^1 dy_4 \right) \quad (19)$$

$$= 4(1 - y_1)(1 - y_3). \quad (20)$$

We observe that  $Y_1$  and  $Y_3$  are independent since  $f_{Y_1, Y_3}(y_1, y_3) = f_{Y_1}(y_1)f_{Y_3}(y_3)$ . It follows that

$$R_{\mathbf{Y}}(1, 3) = E[Y_1 Y_3] = E[Y_1] E[Y_3] = 1/9. \quad (21)$$

Finally, we need to calculate

$$f_{Y_2, Y_4}(y_2, y_4) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{Y_1, Y_2, Y_3, Y_4}(y_1, y_2, y_3, y_4) dy_1 dy_3 \quad (22)$$

$$= 4 \left( \int_0^{y_2} dy_1 \right) \left( \int_0^{y_4} dy_3 \right) \quad (23)$$

$$= 4y_2 y_4. \quad (24)$$

We observe that  $Y_2$  and  $Y_4$  are independent since  $f_{Y_2, Y_4}(y_2, y_4) = f_{Y_2}(y_2)f_{Y_4}(y_4)$ . It follows that  $R_{\mathbf{Y}}(2, 4) = E[Y_2 Y_4] = E[Y_2]E[Y_4] = 4/9$ . The above results give  $R_{\mathbf{Y}}(i, j)$  for  $i \leq j$ . Since  $\mathbf{R}_{\mathbf{Y}}$  is a symmetric matrix,

$$\mathbf{R}_{\mathbf{Y}} = \begin{bmatrix} 1/6 & 1/4 & 1/9 & 2/9 \\ 1/4 & 1/2 & 2/9 & 4/9 \\ 1/9 & 2/9 & 1/6 & 1/4 \\ 2/9 & 4/9 & 1/4 & 1/2 \end{bmatrix}. \quad (25)$$

Since  $\boldsymbol{\mu}_{\mathbf{X}} = [1/3 \ 2/3 \ 1/3 \ 2/3]'$ , the covariance matrix is

$$\mathbf{C}_{\mathbf{Y}} = \mathbf{R}_{\mathbf{Y}} - \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}_{\mathbf{X}}' \quad (26)$$

$$= \begin{bmatrix} 1/6 & 1/4 & 1/9 & 2/9 \\ 1/4 & 1/2 & 2/9 & 4/9 \\ 2/9 & 4/9 & 1/4 & 1/2 \end{bmatrix} - \begin{bmatrix} 1/3 \\ 2/3 \\ 1/3 \\ 2/3 \end{bmatrix} [1/3 \ 2/3 \ 1/3 \ 2/3] \quad (27)$$

$$= \begin{bmatrix} 1/18 & 1/36 & 0 & 0 \\ 1/36 & 1/18 & 0 & 0 \\ 0 & 0 & 1/18 & 1/36 \\ 0 & 0 & 1/36 & 1/18 \end{bmatrix}. \quad (28)$$

The off-diagonal zero blocks are a consequence of  $[Y_1 \ Y_2]'$  being independent of  $[Y_3 \ Y_4]'$ . Along the diagonal, the two identical sub-blocks occur because  $f_{Y_1, Y_2}(x, y) = f_{Y_3, Y_4}(x, y)$ . In short, the matrix structure is the result of  $[Y_1 \ Y_2]'$  and  $[Y_3 \ Y_4]'$  being iid random vectors.

### Problem 5.6.8 Solution

The 2-dimensional random vector  $\mathbf{Y}$  has PDF

$$f_{\mathbf{Y}}(\mathbf{y}) = \begin{cases} 2 & \mathbf{y} \geq \mathbf{0}, [1 \ 1] \mathbf{y} \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Rewritten in terms of the variables  $y_1$  and  $y_2$ ,

$$f_{Y_1, Y_2}(y_1, y_2) = \begin{cases} 2 & y_1 \geq 0, y_2 \geq 0, y_1 + y_2 \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

In this problem, the PDF is simple enough that we can compute  $E[Y_i^n]$  for arbitrary integers  $n \geq 0$ .

$$E[Y_1^n] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1^n f_{Y_1, Y_2}(y_1, y_2) dy_1 dy_2 = \int_0^1 \int_0^{1-y_2} 2y_1^n dy_1 dy_2. \quad (3)$$

A little calculus yields

$$E[Y_1^n] = \int_0^1 \left( \frac{2}{n+1} y_1^{n+1} \Big|_0^{1-y_2} \right) dy_2 = \frac{2}{n+1} \int_0^1 (1-y_2)^{n+1} dy_2 = \frac{2}{(n+1)(n+2)}. \quad (4)$$

Symmetry of the joint PDF  $f_{Y_1, 2}(y_1, 2)$  implies that  $E[Y_2^n] = E[Y_1^n]$ . Thus,  $E[Y_1] = E[Y_2] = 1/3$  and

$$E[\mathbf{Y}] = \boldsymbol{\mu}_{\mathbf{Y}} = [1/3 \ 1/3]'. \quad (5)$$

In addition,

$$R_{\mathbf{Y}}(1, 1) = E[Y_1^2] = 1/6, \quad R_{\mathbf{Y}}(2, 2) = E[Y_2^2] = 1/6. \quad (6)$$



To complete the correlation matrix, we find

$$R_{\mathbf{Y}}(1, 2) = E[Y_1 Y_2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 y_2 f_{Y_1, Y_2}(y_1, y_2) dy_1 dy_2 = \int_0^1 \int_0^{1-y_2} 2y_1 y_2 dy_1 dy_2. \quad (7)$$

Following through on the calculus, we obtain

$$R_{\mathbf{Y}}(1, 2) = \int_0^1 \left( y_1^2 \Big|_0^{1-y_2} \right) y_2 dy_2 = \int_0^1 y_2 (1 - y_2)^2 dy_2 = \frac{1}{2} y_2^2 - \frac{2}{3} y_2^3 + \frac{1}{4} y_2^4 \Big|_0^1 = \frac{1}{12}. \quad (8)$$

Thus we have found that

$$\mathbf{R}_{\mathbf{Y}} = \begin{bmatrix} E[Y_1^2] & E[Y_1 Y_2] \\ E[Y_2 Y_1] & E[Y_2^2] \end{bmatrix} = \begin{bmatrix} 1/6 & 1/12 \\ 1/12 & 1/6 \end{bmatrix}. \quad (9)$$

Lastly,  $\mathbf{Y}$  has covariance matrix

$$\mathbf{C}_{\mathbf{Y}} = \mathbf{R}_{\mathbf{Y}} - \boldsymbol{\mu}_{\mathbf{Y}} \boldsymbol{\mu}'_{\mathbf{Y}} = \begin{bmatrix} 1/6 & 1/12 \\ 1/12 & 1/6 \end{bmatrix} - \begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix} \begin{bmatrix} 1/3 & 1/3 \end{bmatrix} \quad (10)$$

$$= \begin{bmatrix} 1/9 & -1/36 \\ -1/36 & 1/9 \end{bmatrix}. \quad (11)$$

### Problem 5.6.9 Solution

Given an arbitrary random vector  $\mathbf{X}$ , we can define  $\mathbf{Y} = \mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}$  so that

$$\mathbf{C}_{\mathbf{X}} = E[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})'] = E[\mathbf{Y}\mathbf{Y}'] = \mathbf{R}_{\mathbf{Y}}. \quad (1)$$

It follows that the covariance matrix  $\mathbf{C}_{\mathbf{X}}$  is positive semi-definite if and only if the correlation matrix  $\mathbf{R}_{\mathbf{Y}}$  is positive semi-definite. Thus, it is sufficient to show that every correlation matrix, whether it is denoted  $\mathbf{R}_{\mathbf{Y}}$  or  $\mathbf{R}_{\mathbf{X}}$ , is positive semi-definite.

To show a correlation matrix  $\mathbf{R}_{\mathbf{X}}$  is positive semi-definite, we write

$$\mathbf{a}'\mathbf{R}_{\mathbf{X}}\mathbf{a} = \mathbf{a}'E[\mathbf{X}\mathbf{X}']\mathbf{a} = E[\mathbf{a}'\mathbf{X}\mathbf{X}'\mathbf{a}] = E[(\mathbf{a}'\mathbf{X})(\mathbf{X}'\mathbf{a})] = E[(\mathbf{a}'\mathbf{X})^2]. \quad (2)$$

We note that  $W = \mathbf{a}'\mathbf{X}$  is a random variable. Since  $E[W^2] \geq 0$  for any random variable  $W$ ,

$$\mathbf{a}'\mathbf{R}_{\mathbf{X}}\mathbf{a} = E[W^2] \geq 0. \quad (3)$$

### Problem 5.7.1 Solution

(a) From Theorem 5.12, the correlation matrix of  $\mathbf{X}$  is

$$\mathbf{R}_{\mathbf{X}} = \mathbf{C}_{\mathbf{X}} + \boldsymbol{\mu}_{\mathbf{X}} \boldsymbol{\mu}'_{\mathbf{X}} \quad (1)$$

$$= \begin{bmatrix} 4 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 4 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \\ 6 \end{bmatrix} \begin{bmatrix} 4 & 8 & 6 \end{bmatrix} \quad (2)$$

$$= \begin{bmatrix} 4 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 4 \end{bmatrix} + \begin{bmatrix} 16 & 32 & 24 \\ 32 & 64 & 48 \\ 24 & 48 & 36 \end{bmatrix} = \begin{bmatrix} 20 & 30 & 25 \\ 30 & 68 & 46 \\ 25 & 46 & 40 \end{bmatrix} \quad (3)$$

- (b) Let  $\mathbf{Y} = [X_1 \ X_2]'$ . Since  $\mathbf{Y}$  is a subset of the components of  $\mathbf{X}$ , it is a Gaussian random vector with expected value vector

$$\boldsymbol{\mu}_Y = [E[X_1] \ E[X_2]]' = [4 \ 8]'. \quad (4)$$

and covariance matrix

$$\mathbf{C}_Y = \begin{bmatrix} \text{Var}[X_1] & \text{Cov}[X_1, X_2] \\ \mathbf{C}_{X_1 X_2} & \text{Var}[X_2] \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix} \quad (5)$$

We note that  $\det(\mathbf{C}_Y) = 12$  and that

$$\mathbf{C}_Y^{-1} = \frac{1}{12} \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/6 \\ 1/6 & 1/3 \end{bmatrix}. \quad (6)$$

This implies that

$$(\mathbf{y} - \boldsymbol{\mu}_Y)' \mathbf{C}_Y^{-1} (\mathbf{y} - \boldsymbol{\mu}_Y) = [y_1 - 4 \ y_2 - 8] \begin{bmatrix} 1/3 & 1/6 \\ 1/6 & 1/3 \end{bmatrix} \begin{bmatrix} y_1 - 4 \\ y_2 - 8 \end{bmatrix} \quad (7)$$

$$= [y_1 - 4 \ y_2 - 8] \begin{bmatrix} y_1/3 + y_2/6 - 8/3 \\ y_1/6 + y_2/3 - 10/3 \end{bmatrix} \quad (8)$$

$$= \frac{y_1^2}{3} + \frac{y_1 y_2}{3} - \frac{16y_1}{3} - \frac{20y_2}{3} + \frac{y_2^2}{3} + \frac{112}{3} \quad (9)$$

The PDF of  $\mathbf{Y}$  is

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{2\pi\sqrt{12}} e^{-(\mathbf{y} - \boldsymbol{\mu}_Y)' \mathbf{C}_Y^{-1} (\mathbf{y} - \boldsymbol{\mu}_Y)/2} \quad (10)$$

$$= \frac{1}{\sqrt{48\pi^2}} e^{-(y_1^2 + y_1 y_2 - 16y_1 - 20y_2 + y_2^2 + 112)/6} \quad (11)$$

Since  $\mathbf{Y} = [X_1, X_2]'$ , the PDF of  $X_1$  and  $X_2$  is simply

$$f_{X_1, X_2}(x_1, x_2) = f_{Y_1, Y_2}(x_1, x_2) = \frac{1}{\sqrt{48\pi^2}} e^{-(x_1^2 + x_1 x_2 - 16x_1 - 20x_2 + x_2^2 + 112)/6} \quad (12)$$

- (c) We can observe directly from  $\boldsymbol{\mu}_X$  and  $\mathbf{C}_X$  that  $X_1$  is a Gaussian (4, 2) random variable. Thus,

$$P[X_1 > 8] = P\left[\frac{X_1 - 4}{2} > \frac{8 - 4}{2}\right] = Q(2) = 0.0228 \quad (13)$$

### Problem 5.7.2 Solution

We are given that  $\mathbf{X}$  is a Gaussian random vector with

$$\boldsymbol{\mu}_X = \begin{bmatrix} 4 \\ 8 \\ 6 \end{bmatrix} \quad \mathbf{C}_X = \begin{bmatrix} 4 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 4 \end{bmatrix}. \quad (1)$$

We are also given that  $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$  where

$$\mathbf{A} = \begin{bmatrix} 1 & 1/2 & 2/3 \\ 1 & -1/2 & 2/3 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} -4 \\ -4 \end{bmatrix}. \quad (2)$$

Since the two rows of  $\mathbf{A}$  are linearly independent row vectors,  $\mathbf{A}$  has rank 2. By Theorem 5.16,  $\mathbf{Y}$  is a Gaussian random vector. Given these facts, the various parts of this problem are just straightforward calculations using Theorem 5.16.

(a) The expected value of  $\mathbf{Y}$  is

$$\boldsymbol{\mu}_{\mathbf{Y}} = \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b} = \begin{bmatrix} 1 & 1/2 & 2/3 \\ 1 & -1/2 & 2/3 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \\ 6 \end{bmatrix} + \begin{bmatrix} -4 \\ -4 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \end{bmatrix}. \quad (3)$$

(b) The covariance matrix of  $\mathbf{Y}$  is

$$\mathbf{C}_{\mathbf{Y}} = \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A} \quad (4)$$

$$= \begin{bmatrix} 1 & 1/2 & 2/3 \\ 1 & -1/2 & 2/3 \end{bmatrix} \begin{bmatrix} 4 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1/2 & -1/2 \\ 2/3 & 2/3 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 43 & 55 \\ 55 & 103 \end{bmatrix}. \quad (5)$$

(c)  $\mathbf{Y}$  has correlation matrix

$$\mathbf{R}_{\mathbf{Y}} = \mathbf{C}_{\mathbf{Y}} + \boldsymbol{\mu}_{\mathbf{Y}}\boldsymbol{\mu}'_{\mathbf{Y}} = \frac{1}{9} \begin{bmatrix} 43 & 55 \\ 55 & 103 \end{bmatrix} + \begin{bmatrix} 8 \\ 0 \end{bmatrix} \begin{bmatrix} 8 & 0 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 619 & 55 \\ 55 & 103 \end{bmatrix} \quad (6)$$

(d) From  $\boldsymbol{\mu}_{\mathbf{Y}}$ , we see that  $E[Y_2] = 0$ . From the covariance matrix  $\mathbf{C}_{\mathbf{Y}}$ , we learn that  $Y_2$  has variance  $\sigma_2^2 = C_{\mathbf{Y}}(2, 2) = 103/9$ . Since  $Y_2$  is a Gaussian random variable,

$$P[-1 \leq Y_2 \leq 1] = P\left[-\frac{1}{\sigma_2} \leq \frac{Y_2}{\sigma_2} \leq \frac{1}{\sigma_2}\right] \quad (7)$$

$$= \Phi\left(\frac{1}{\sigma_2}\right) - \Phi\left(\frac{-1}{\sigma_2}\right) \quad (8)$$

$$= 2\Phi\left(\frac{1}{\sigma_2}\right) - 1 \quad (9)$$

$$= 2\Phi\left(\frac{3}{\sqrt{103}}\right) - 1 = 0.2325. \quad (10)$$

### Problem 5.7.3 Solution

This problem is just a special case of Theorem 5.16 with the matrix  $\mathbf{A}$  replaced by the row vector  $\mathbf{a}'$  and a 1 element vector  $\mathbf{b} = b = 0$ . In this case, the vector  $\mathbf{Y}$  becomes the scalar  $Y$ . The expected value vector  $\boldsymbol{\mu}_{\mathbf{Y}} = [\mu_Y]$  and the covariance “matrix” of  $Y$  is just the  $1 \times 1$  matrix  $[\sigma_Y^2]$ . Directly from Theorem 5.16, we can conclude that  $Y$  is a length 1 Gaussian random vector, which is just a Gaussian random variable. In addition,  $\mu_Y = \mathbf{a}'\boldsymbol{\mu}_{\mathbf{X}}$  and

$$\text{Var}[Y] = \mathbf{C}_Y = \mathbf{a}'\mathbf{C}_{\mathbf{X}}\mathbf{a}. \quad (1)$$

### Problem 5.7.4 Solution

From Definition 5.17, the  $n = 2$  dimensional Gaussian vector  $\mathbf{X}$  has PDF

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{2\pi[\det(\mathbf{C}_{\mathbf{X}})]^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})'\mathbf{C}_{\mathbf{X}}^{-1}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})\right) \quad (1)$$

where  $\mathbf{C}_{\mathbf{X}}$  has determinant

$$\det(\mathbf{C}_{\mathbf{X}}) = \sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2 = \sigma_1^2\sigma_2^2(1 - \rho^2). \quad (2)$$

Thus,

$$\frac{1}{2\pi[\det(\mathbf{C}_\mathbf{X})]^{1/2}} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}. \quad (3)$$

Using the  $2 \times 2$  matrix inverse formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \quad (4)$$

we obtain

$$\mathbf{C}_\mathbf{X}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} = \frac{1}{1-\rho^2} \begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix}. \quad (5)$$

Thus

$$-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_\mathbf{X})' \mathbf{C}_\mathbf{X}^{-1} (\mathbf{x} - \boldsymbol{\mu}_\mathbf{X}) = -\frac{\begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}}{2(1-\rho^2)} \quad (6)$$

$$= -\frac{\begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \begin{bmatrix} \frac{x_1 - \mu_1}{\sigma_1^2} - \frac{\rho(x_2 - \mu_2)}{\sigma_1\sigma_2} \\ -\frac{\rho(x_1 - \mu_1)}{\sigma_1\sigma_2} + \frac{x_2 - \mu_2}{\sigma_2^2} \end{bmatrix}}{2(1-\rho^2)} \quad (7)$$

$$= -\frac{\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2}}{2(1-\rho^2)}. \quad (8)$$

Combining Equations (1), (3), and (8), we see that

$$f_\mathbf{X}(\mathbf{x}) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left[ -\frac{\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2}}{2(1-\rho^2)} \right], \quad (9)$$

which is the bivariate Gaussian PDF in Definition 4.17.

### Problem 5.7.5 Solution

Since

$$\mathbf{W} = \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ \mathbf{A} \end{bmatrix} \mathbf{X} = \mathbf{D}\mathbf{X} \quad (1)$$

Suppose that  $\mathbf{X}$  Gaussian  $(\mathbf{0}, \mathbf{I})$  random vector. By Theorem 5.13,  $\boldsymbol{\mu}_\mathbf{W} = \mathbf{0}$  and  $\mathbf{C}_\mathbf{W} = \mathbf{D}\mathbf{D}'$ . The matrix  $\mathbf{D}$  is  $(m+n) \times n$  and has rank  $n$ . That is, the rows of  $\mathbf{D}$  are dependent and there exists a vector  $\mathbf{y}$  such that  $\mathbf{y}'\mathbf{D} = \mathbf{0}$ . This implies  $\mathbf{y}'\mathbf{D}\mathbf{D}'\mathbf{y} = 0$ . Hence  $\det(\mathbf{C}_\mathbf{W}) = 0$  and  $\mathbf{C}_\mathbf{W}^{-1}$  does not exist. Hence  $\mathbf{W}$  is **not** a Gaussian random vector.

The point to keep in mind is that the definition of a Gaussian random vector does not permit a component random variable to be a deterministic linear combination of other components.

### Problem 5.7.6 Solution

(a) From Theorem 5.13,  $\mathbf{Y}$  has covariance matrix

$$\mathbf{C}_Y = \mathbf{Q}\mathbf{C}_X\mathbf{Q}' \quad (1)$$

$$= \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (2)$$

$$= \begin{bmatrix} \sigma_1^2 \cos^2 \theta + \sigma_2^2 \sin^2 \theta & (\sigma_1^2 - \sigma_2^2) \sin \theta \cos \theta \\ (\sigma_1^2 - \sigma_2^2) \sin \theta \cos \theta & \sigma_1^2 \sin^2 \theta + \sigma_2^2 \cos^2 \theta \end{bmatrix}. \quad (3)$$

We conclude that  $Y_1$  and  $Y_2$  have covariance

$$\text{Cov}[Y_1, Y_2] = C_Y(1, 2) = (\sigma_1^2 - \sigma_2^2) \sin \theta \cos \theta. \quad (4)$$

Since  $Y_1$  and  $Y_2$  are jointly Gaussian, they are independent if and only if  $\text{Cov}[Y_1, Y_2] = 0$ . Thus,  $Y_1$  and  $Y_2$  are independent for all  $\theta$  if and only if  $\sigma_1^2 = \sigma_2^2$ . In this case, when the joint PDF  $f_{\mathbf{X}}(\mathbf{x})$  is symmetric in  $x_1$  and  $x_2$ . In terms of polar coordinates, the PDF  $f_{\mathbf{X}}(\mathbf{x}) = f_{X_1, X_2}(x_1, x_2)$  depends on  $r = \sqrt{x_1^2 + x_2^2}$  but for a given  $r$ , is constant for all  $\phi = \tan^{-1}(x_2/x_1)$ . The transformation of  $\mathbf{X}$  to  $\mathbf{Y}$  is just a rotation of the coordinate system by  $\theta$  preserves this circular symmetry.

(b) If  $\sigma_2^2 > \sigma_1^2$ , then  $Y_1$  and  $Y_2$  are independent if and only if  $\sin \theta \cos \theta = 0$ . This occurs in the following cases:

- $\theta = 0$ :  $Y_1 = X_1$  and  $Y_2 = X_2$
- $\theta = \pi/2$ :  $Y_1 = -X_2$  and  $Y_2 = -X_1$
- $\theta = \pi$ :  $Y_1 = -X_1$  and  $Y_2 = -X_2$
- $\theta = -\pi/2$ :  $Y_1 = X_2$  and  $Y_2 = X_1$

In all four cases,  $Y_1$  and  $Y_2$  are just relabeled versions, possibly with sign changes, of  $X_1$  and  $X_2$ . In these cases,  $Y_1$  and  $Y_2$  are independent because  $X_1$  and  $X_2$  are independent. For other values of  $\theta$ , each  $Y_i$  is a linear combination of both  $X_1$  and  $X_2$ . This mixing results in correlation between  $Y_1$  and  $Y_2$ .

### Problem 5.7.7 Solution

The difficulty of this problem is overrated since its a pretty simple application of Problem 5.7.6. In particular,

$$\mathbf{Q} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \Big|_{\theta=45^\circ} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \quad (1)$$

Since  $\mathbf{X} = \mathbf{Q}\mathbf{Y}$ , we know from Theorem 5.16 that  $\mathbf{X}$  is Gaussian with covariance matrix

$$\mathbf{C}_X = \mathbf{Q}\mathbf{C}_Y\mathbf{Q}' \quad (2)$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1+\rho & 0 \\ 0 & 1-\rho \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad (3)$$

$$= \frac{1}{2} \begin{bmatrix} 1+\rho & -(1-\rho) \\ 1+\rho & 1-\rho \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad (4)$$

$$= \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \quad (5)$$

### Problem 5.7.8 Solution

As given in the problem statement, we define the  $m$ -dimensional vector  $\mathbf{X}$ , the  $n$ -dimensional vector  $\mathbf{Y}$  and  $\mathbf{W} = \begin{bmatrix} \mathbf{X}' \\ \mathbf{Y}' \end{bmatrix}'$ . Note that  $\mathbf{W}$  has expected value

$$\boldsymbol{\mu}_{\mathbf{W}} = E[\mathbf{W}] = E \left[ \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \right] = \begin{bmatrix} E[\mathbf{X}] \\ E[\mathbf{Y}] \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{X}} \\ \boldsymbol{\mu}_{\mathbf{Y}} \end{bmatrix}. \quad (1)$$

The covariance matrix of  $\mathbf{W}$  is

$$\mathbf{C}_{\mathbf{W}} = E[(\mathbf{W} - \boldsymbol{\mu}_{\mathbf{W}})(\mathbf{W} - \boldsymbol{\mu}_{\mathbf{W}})'] \quad (2)$$

$$= E \left[ \begin{bmatrix} \mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}} \\ \mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}} \end{bmatrix} \begin{bmatrix} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})' & (\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})' \end{bmatrix} \right] \quad (3)$$

$$= \begin{bmatrix} E[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})'] & E[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})'] \\ E[(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})'] & E[(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})'] \end{bmatrix} \quad (4)$$

$$= \begin{bmatrix} \mathbf{C}_{\mathbf{X}} & \mathbf{C}_{\mathbf{XY}} \\ \mathbf{C}_{\mathbf{YX}} & \mathbf{C}_{\mathbf{Y}} \end{bmatrix}. \quad (5)$$

The assumption that  $\mathbf{X}$  and  $\mathbf{Y}$  are independent implies that

$$\mathbf{C}_{\mathbf{XY}} = E[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{Y}' - \boldsymbol{\mu}'_{\mathbf{Y}})] = (E[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})] E[(\mathbf{Y}' - \boldsymbol{\mu}'_{\mathbf{Y}})]) = \mathbf{0}. \quad (6)$$

This also implies  $\mathbf{C}_{\mathbf{YX}} = \mathbf{C}'_{\mathbf{XY}} = \mathbf{0}'$ . Thus

$$\mathbf{C}_{\mathbf{W}} = \begin{bmatrix} \mathbf{C}_{\mathbf{X}} & \mathbf{0} \\ \mathbf{0}' & \mathbf{C}_{\mathbf{Y}} \end{bmatrix}. \quad (7)$$

### Problem 5.7.9 Solution

- (a) If you are familiar with the Gram-Schmidt procedure, the argument is that applying Gram-Schmidt to the rows of  $\mathbf{A}$  yields  $m$  orthogonal row vectors. It is then possible to augment those vectors with an additional  $n - m$  orthogonal vectors. Those orthogonal vectors would be the rows of  $\tilde{\mathbf{A}}$ .

An alternate argument is that since  $\mathbf{A}$  has rank  $m$  the nullspace of  $\mathbf{A}$ , i.e., the set of all vectors  $\mathbf{y}$  such that  $\mathbf{A}\mathbf{y} = \mathbf{0}$  has dimension  $n - m$ . We can choose any  $n - m$  linearly independent vectors  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{n-m}$  in the nullspace of  $\mathbf{A}$ . We then define  $\tilde{\mathbf{A}}$  to have columns  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{n-m}$ . It follows that  $\mathbf{A}\tilde{\mathbf{A}}' = \mathbf{0}$ .

- (b) To use Theorem 5.16 for the case  $m = n$  to show

$$\bar{\mathbf{Y}} = \begin{bmatrix} \mathbf{Y} \\ \hat{\mathbf{Y}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \hat{\mathbf{A}} \end{bmatrix} \mathbf{X}. \quad (1)$$

is a Gaussian random vector requires us to show that

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A} \\ \hat{\mathbf{A}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1} \end{bmatrix} \quad (2)$$

is a rank  $n$  matrix. To prove this fact, we will suppose there exists  $\mathbf{w}$  such that  $\bar{\mathbf{A}}\mathbf{w} = \mathbf{0}$ , and then show that  $\mathbf{w}$  is a zero vector. Since  $\mathbf{A}$  and  $\tilde{\mathbf{A}}$  together have  $n$  linearly independent rows, we can write the row vector  $\mathbf{w}'$  as a linear combination of the rows of  $\mathbf{A}$  and  $\tilde{\mathbf{A}}$ . That is, for some  $\mathbf{v}$  and  $\tilde{\mathbf{v}}$ ,

$$\mathbf{w}' = \mathbf{v}\mathbf{t}'\mathbf{A} + \tilde{\mathbf{v}}'\tilde{\mathbf{A}}. \quad (3)$$

The condition  $\bar{\mathbf{A}}\mathbf{w} = \mathbf{0}$  implies

$$\begin{bmatrix} \mathbf{A} \\ \tilde{\mathbf{A}}\mathbf{C}_X^{-1} \end{bmatrix} (\mathbf{A}'\mathbf{v} + \tilde{\mathbf{A}}'\tilde{\mathbf{v}}') = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \quad (4)$$

This implies

$$\mathbf{A}\mathbf{A}'\mathbf{v} + \mathbf{A}\tilde{\mathbf{A}}'\tilde{\mathbf{v}} = \mathbf{0} \quad (5)$$

$$\tilde{\mathbf{A}}\mathbf{C}_X^{-1}\mathbf{A}\mathbf{v} + \tilde{\mathbf{A}}\mathbf{C}_X^{-1}\tilde{\mathbf{A}}'\tilde{\mathbf{v}} = \mathbf{0} \quad (6)$$

Since  $\mathbf{A}\tilde{\mathbf{A}}' = \mathbf{0}$ , Equation (5) implies that  $\mathbf{A}\mathbf{A}'\mathbf{v} = \mathbf{0}$ . Since  $\mathbf{A}$  is rank  $m$ ,  $\mathbf{A}\mathbf{A}'$  is an  $m \times m$  rank  $m$  matrix. It follows that  $\mathbf{v} = \mathbf{0}$ . We can then conclude from Equation (6) that

$$\tilde{\mathbf{A}}\mathbf{C}_X^{-1}\tilde{\mathbf{A}}'\tilde{\mathbf{v}} = \mathbf{0}. \quad (7)$$

This would imply that  $\tilde{\mathbf{v}}'\tilde{\mathbf{A}}\mathbf{C}_X^{-1}\tilde{\mathbf{A}}'\tilde{\mathbf{v}} = 0$ . Since  $\mathbf{C}_X^{-1}$  is invertible, this would imply that  $\tilde{\mathbf{A}}'\tilde{\mathbf{v}} = \mathbf{0}$ . Since the rows of  $\tilde{\mathbf{A}}$  are linearly independent, it must be that  $\tilde{\mathbf{v}} = \mathbf{0}$ . Thus  $\bar{\mathbf{A}}$  is full rank and  $\bar{\mathbf{Y}}$  is a Gaussian random vector.

(c) We note that By Theorem 5.16, the Gaussian vector  $\bar{\mathbf{Y}} = \bar{\mathbf{A}}\mathbf{X}$  has covariance matrix

$$\bar{\mathbf{C}} = \bar{\mathbf{A}}\mathbf{C}_X\bar{\mathbf{A}}'. \quad (8)$$

Since  $(\mathbf{C}_X^{-1})' = \mathbf{C}_X^{-1}$ ,

$$\bar{\mathbf{A}}' = [\mathbf{A}' \quad (\tilde{\mathbf{A}}\mathbf{C}_X^{-1})'] = [\mathbf{A}' \quad \mathbf{C}_X^{-1}\tilde{\mathbf{A}}']. \quad (9)$$

Applying this result to Equation (8) yields

$$\bar{\mathbf{C}} = \begin{bmatrix} \mathbf{A} \\ \tilde{\mathbf{A}}\mathbf{C}_X^{-1} \end{bmatrix} \mathbf{C}_X [\mathbf{A}' \quad \mathbf{C}_X^{-1}\tilde{\mathbf{A}}'] = \begin{bmatrix} \mathbf{A}\mathbf{C}_X \\ \tilde{\mathbf{A}} \end{bmatrix} [\mathbf{A}' \quad \mathbf{C}_X^{-1}\tilde{\mathbf{A}}'] = \begin{bmatrix} \mathbf{A}\mathbf{C}_X\mathbf{A}' & \mathbf{A}\tilde{\mathbf{A}}' \\ \tilde{\mathbf{A}}\mathbf{A}' & \tilde{\mathbf{A}}\mathbf{C}_X^{-1}\tilde{\mathbf{A}}' \end{bmatrix}. \quad (10)$$

Since  $\tilde{\mathbf{A}}\mathbf{A}' = \mathbf{0}$ ,

$$\bar{\mathbf{C}} = \begin{bmatrix} \mathbf{A}\mathbf{C}_X\mathbf{A}' & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{A}}\mathbf{C}_X^{-1}\tilde{\mathbf{A}}' \end{bmatrix} = \begin{bmatrix} \mathbf{C}_Y & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\hat{\mathbf{Y}}} \end{bmatrix}. \quad (11)$$

We see that  $\bar{\mathbf{C}}$  is block diagonal covariance matrix. From the claim of Problem 5.7.8, we can conclude that  $\mathbf{Y}$  and  $\hat{\mathbf{Y}}$  are independent Gaussian random vectors.

### Problem 5.8.1 Solution

We can use Theorem 5.16 since the scalar  $Y$  is also a 1-dimensional vector. To do so, we write

$$Y = [1/3 \quad 1/3 \quad 1/3] \mathbf{X} = \mathbf{A}\mathbf{X}. \quad (1)$$

By Theorem 5.16,  $Y$  is a Gaussian vector with expected value

$$E[Y] = \mathbf{A}\boldsymbol{\mu}_X = (E[X_1] + E[X_2] + E[X_3])/3 = (4 + 8 + 6)/3 = 6 \quad (2)$$

and covariance matrix

$$\mathbf{C}_Y = \text{Var}[Y] = \mathbf{A}\mathbf{C}_X\mathbf{A}' \quad (3)$$

$$= \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} 4 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 4 \end{bmatrix} \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \frac{2}{3} \quad (4)$$

Thus  $Y$  is a Gaussian  $(6, \sqrt{2/3})$  random variable, implying

$$P[Y > 4] = P\left[\frac{Y-6}{\sqrt{2/3}} > \frac{4-6}{\sqrt{2/3}}\right] = 1 - \Phi(-\sqrt{6}) = \Phi(\sqrt{6}) = 0.9928 \quad (5)$$

### Problem 5.8.2 Solution

- (a) The covariance matrix  $\mathbf{C}_X$  has  $\text{Var}[X_i] = 25$  for each diagonal entry. For  $i \neq j$ , the  $i, j$ th entry of  $\mathbf{C}_X$  is

$$[\mathbf{C}_X]_{ij} = \rho_{X_i X_j} \sqrt{\text{Var}[X_i] \text{Var}[X_j]} = (0.8)(25) = 20 \quad (1)$$

The covariance matrix of  $X$  is a  $10 \times 10$  matrix of the form

$$\mathbf{C}_X = \begin{bmatrix} 25 & 20 & \cdots & 20 \\ 20 & 25 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 20 \\ 20 & \cdots & 20 & 25 \end{bmatrix}. \quad (2)$$

- (b) We observe that

$$Y = [1/10 \quad 1/10 \quad \cdots \quad 1/10] \mathbf{X} = \mathbf{A}\mathbf{X} \quad (3)$$

Since  $Y$  is the average of 10 iid random variables,  $E[Y] = E[X_i] = 5$ . Since  $Y$  is a scalar, the  $1 \times 1$  covariance matrix  $\mathbf{C}_Y = \text{Var}[Y]$ . By Theorem 5.13, the variance of  $Y$  is

$$\text{Var}[Y] = \mathbf{C}_Y = \mathbf{A}\mathbf{C}_X\mathbf{A}' = 20.5 \quad (4)$$

Since  $Y$  is Gaussian,

$$P[Y \leq 25] = P\left[\frac{Y-5}{\sqrt{20.5}} \leq \frac{25-20.5}{\sqrt{20.5}}\right] = \Phi(0.9939) = 0.8399. \quad (5)$$

### Problem 5.8.3 Solution

Under the model of Quiz 5.8, the temperature on day  $i$  and on day  $j$  have covariance

$$\text{Cov}[T_i, T_j] = C_T[i-j] = \frac{36}{1+|i-j|} \quad (1)$$

From this model, the vector  $\mathbf{T} = [T_1 \quad \cdots \quad T_{31}]'$  has covariance matrix

$$\mathbf{C}_T = \begin{bmatrix} C_T[0] & C_T[1] & \cdots & C_T[30] \\ C_T[1] & C_T[0] & \ddots & \vdots \\ \vdots & \ddots & \ddots & C_T[1] \\ C_T[30] & \cdots & C_T[1] & C_T[0] \end{bmatrix}. \quad (2)$$



If you have read the solution to Quiz 5.8, you know that  $\mathbf{C}_T$  is a symmetric Toeplitz matrix and that MATLAB has a `toeplitz` function to generate Toeplitz matrices. Using the `toeplitz` function to generate the covariance matrix, it is easy to use `gaussvector` to generate samples of the random vector  $\mathbf{T}$ . Here is the code for estimating  $P[A]$  using  $m$  samples.

<pre>function p=julytemp583(m); c=36./(1+(0:30)); CT=toeplitz(c); mu=80*ones(31,1); T=gaussvector(mu,CT,m); Y=sum(T)/31; Tmin=min(T); p=sum((Tmin&gt;=72) &amp; (Y &lt;= 82))/m;</pre>	<pre>julytemp583(100000) ans =     0.0684 &gt;&gt; julytemp583(100000) ans =     0.0706 &gt;&gt; julytemp583(100000) ans =     0.0714 &gt;&gt; julytemp583(100000) ans =     0.0701</pre>
--	---

We see from repeated experiments with  $m = 100,000$  trials that  $P[A] \approx 0.07$ .

#### Problem 5.8.4 Solution

The covariance matrix  $\mathbf{C}_X$  has  $\text{Var}[X_i] = 25$  for each diagonal entry. For  $i \neq j$ , the  $i, j$ th entry of  $\mathbf{C}_X$  is

$$[\mathbf{C}_X]_{ij} = \rho_{X_i X_j} \sqrt{\text{Var}[X_i] \text{Var}[X_j]} = (0.8)(25) = 20 \quad (1)$$

The covariance matrix of  $X$  is a  $10 \times 10$  matrix of the form

$$\mathbf{C}_X = \begin{bmatrix} 25 & 20 & \cdots & 20 \\ 20 & 25 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 20 \\ 20 & \cdots & 20 & 25 \end{bmatrix}. \quad (2)$$

A program to estimate  $P[W \leq 25]$  uses `gaussvector` to generate  $m$  sample vector of race times  $\mathbf{X}$ . In the program `sailboats.m`,  $\mathbf{X}$  is an  $10 \times m$  matrix such that each column of  $\mathbf{X}$  is a vector of race times. In addition `min(X)` is a row vector indicating the fastest time in each race.

<pre>function p=sailboats(w,m) %Usage: p=sailboats(f,m) %In Problem 5.8.4, W is the %winning time in a 10 boat race. %We use m trials to estimate %P[W&lt;=w] CX=(5*eye(10))+(20*ones(10,10)); mu=35*ones(10,1); X=gaussvector(mu,CX,m); W=min(X); p=sum(W&lt;=w)/m;</pre>	<pre>&gt;&gt; sailboats(25,10000) ans =     0.0827 &gt;&gt; sailboats(25,100000) ans =     0.0801 &gt;&gt; sailboats(25,100000) ans =     0.0803 &gt;&gt; sailboats(25,100000) ans =     0.0798</pre>
--	---

We see from repeated experiments with  $m = 100,000$  trials that  $P[W \leq 25] \approx 0.08$ .

#### Problem 5.8.5 Solution

When we built `poissonrv.m`, we went to some trouble to be able to generate  $m$  iid samples at once. In this problem, each Poisson random variable that we generate has an expected value that

is different from that of any other Poisson random variables. Thus, we must generate the daily jackpots sequentially. Here is a simple program for this purpose.

```
function jackpot=lottery1(jstart,M,D)
%Usage: function j=lottery1(jstart,M,D)
%Perform M trials of the D day lottery
%of Problem 5.5.5 and initial jackpot jstart
jackpot=zeros(M,1);
for m=1:M,
    disp('trm')
    jackpot(m)=jstart;
    for d=1:D,
        jackpot(m)=jackpot(m)+(0.5*poissonrv(jackpot(m),1));
    end
end
```

The main problem with `lottery1` is that it will run *very* slowly. Each call to `poissonrv` generates an entire Poisson PMF  $P_X(x)$  for  $x = 0, 1, \dots, x_{\max}$  where  $x_{\max} \geq 2 \cdot 10^6$ . This is slow in several ways. First, we repeating the calculation of  $\sum_{j=1}^{x_{\max}} \log j$  with each call to `poissonrv`. Second, each call to `poissonrv` asks for a Poisson sample value with expected value  $\alpha > 1 \cdot 10^6$ . In these cases, for small values of  $x$ ,  $P_X(x) = \alpha^x e^{-\alpha} / x!$  is so small that it is less than the smallest nonzero number that MATLAB can store!

To speed up the simulation, we have written a program `bigpoissonrv` which generates Poisson ( $\alpha$ ) samples for large  $\alpha$ . The program makes an approximation that for a Poisson ( $\alpha$ ) random variable  $X$ ,  $P_X(x) \approx 0$  for  $|x - \alpha| > 6\sqrt{\alpha}$ . Since  $X$  has standard deviation  $\sqrt{\alpha}$ , we are assuming that  $X$  cannot be more than six standard deviations away from its mean value. The error in this approximation is very small. In fact, for a Poisson ( $a$ ) random variable, the program `poissonsigma(a,k)` calculates the error  $P[|X - a| > k\sqrt{a}]$ . Here is `poissonsigma.m` and some simple calculations:

```
function err=poissonsigma(a,k);
xmin=max(0,floor(a-k*sqrt(a)));
xmax=a+ceil(k*sqrt(a));
sx=xmin:xmax;
logfacts =cumsum([0,log(1:xmax)]);
%logfacts includes 0 in case xmin=0
%Now we extract needed values:
logfacts=logfacts(sx+1);
%pmf(i,:) is a Poisson a(i) PMF
% from xmin to xmax
pmf=exp(-a+ (log(a)*sx)-(logfacts));
err=1-sum(pmf);
```

```
>> poissonsigma(1,6)
ans =
    1.0249e-005
>> poissonsigma(10,6)
ans =
    2.5100e-007
>> poissonsigma(100,6)
ans =
    1.2620e-008
>> poissonsigma(1000,6)
ans =
    2.6777e-009
>> poissonsigma(10000,6)
ans =
    1.8081e-009
>> poissonsigma(100000,6)
ans =
   -1.6383e-010
```

The error reported by `poissonsigma(a,k)` should always be positive. In fact, we observe negative errors for very large  $a$ . For large  $\alpha$  and  $x$ , numerical calculation of  $P_X(x) = \alpha^x e^{-\alpha} / x!$  is tricky because we are taking ratios of very large numbers. In fact, for  $\alpha = x = 1,000,000$ , MATLAB calculation of  $\alpha^x$  and  $x!$  will report infinity while  $e^{-\alpha}$  will evaluate as zero. Our method

of calculating the Poisson ( $\alpha$ ) PMF is to use the fact that  $\ln x! = \sum_{j=1}^x \ln j$  to calculate

$$\exp(\ln P_X(x)) = \exp\left(x \ln \alpha - \alpha - \sum_{j=1}^x \ln j\right). \quad (1)$$

This method works reasonably well except that the calculation of the logarithm has finite precision. The consequence is that the calculated sum over the PMF can vary from 1 by a very small amount, on the order of  $10^{-7}$  in our experiments. In our problem, the error is inconsequential, however, one should keep in mind that this may not be the case in other other experiments using large Poisson random variables. In any case, we can conclude that within the accuracy of MATLAB's simulated experiments, the approximations to be used by `bigpoissonrv` are not significant.

The other feature of `bigpoissonrv` is that for a vector `alpha` corresponding to expected values  $[\alpha_1 \ \cdots \ \alpha_m]'$ , `bigpoissonrv` returns a vector `X` such that `X(i)` is a Poisson `alpha(i)` sample. The work of calculating the sum of logarithms is done only once for all calculated samples. The result is a significant savings in cpu time as long as the values of `alpha` are reasonably close to each other.

```
function x=bigpoissonrv(alpha)
%for vector alpha, returns a vector x such that
% x(i) is a Poisson (alpha(i)) rv
%set up Poisson CDF from xmin to xmax for each alpha(i)
alpha=alpha(:);
amin=min(alpha(:));
amax=max(alpha(:));
%Assume Poisson PMF is negligible +-6 sigma from the average
xmin=max(0,floor(amin-6*sqrt(amax)));
xmax=amax+ceil(6*sqrt(amax));%set max range
sx=xmin:xmax;
%Now we include the basic code of poissonpmf (but starting at xmin)
logfacts =cumsum([0,log(1:xmax)]); %include 0 in case xmin=0
logfacts=logfacts(sx+1); %extract needed values
%pmf(i,:) is a Poisson alpha(i) PMF from xmin to xmax
pmf=exp(-alpha*ones(size(sx))+ ...
        (log(alpha)*sx)-(ones(size(alpha))*logfacts));
cdf=cumsum(pmf,2); %each row is a cdf
x=(xmin-1)+sum((rand(size(alpha))*ones(size(sx)))<=cdf,2);
```

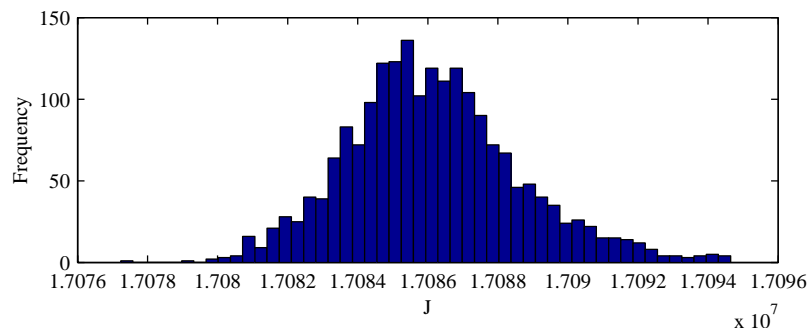
Finally, given `bigpoissonrv`, we can write a short program `lottery` that simulates trials of the jackpot experiment. Ideally, we would like to use `lottery` to perform  $m = 1,000$  trials in a single pass. In general, MATLAB is more efficient when calculations are executed in parallel using vectors. However, in `bigpoissonrv`, the matrix `pmf` will have  $m$  rows and at least  $12\sqrt{\alpha} = 12,000$  columns. For  $m$  more than several hundred, MATLAB running on my laptop reported an “Out of Memory” error. Thus, we wrote the program `lottery` to perform  $M$  trials at once and to repeat that  $N$  times. The output is an  $M \times N$  matrix where each  $i, j$  entry is a sample jackpot after seven days.

```

function jackpot=lottery(jstart,M,N,D)
%Usage: function j=lottery(jstart,M,N,D)
%Perform M trials of the D day lottery
%of Problem 5.5.5 and initial jackpot jstart
jackpot=zeros(M,N);
for n=1:N,
jackpot(:,n)=jstart*ones(M,1);
for d=1:D,
disp(d);
jackpot(:,n)=jackpot(:,n)+(0.5*bigpoissonrv(jackpot(:,n)));
end
end
end

```

Executing `J=lottery(1e6,200,10,7)` generates a matrix `J` of 2,000 sample jackpots. The command `hist(J(:),50)` generates a histogram of the values with 50 bins. An example is shown here:



If you go back and solve Problem 5.5.5, you will see that the jackpot  $J$  has expected value  $E[J] = (3/2)^7 \times 10^6 = 1.70859 \times 10^7$  dollars. Thus it is not surprising that the histogram is centered around a jackpot of  $1.708 \times 10^7$  dollars. If we did more trials, and used more histogram bins, the histogram would appear to converge to the shape of a Gaussian PDF. This fact is explored in Chapter 6.