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

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

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, $N_1, \ldots, N_4$ are mutually independent and have joint PMF

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

(a) For $n = 3$, $P_{N_1,N_2,N_3,N_4}(n_1,n_2,n_3,n_4) = \prod_{i=1}^{4} P_{N_i}(n_i)$, where

$P_{N_i}(n_i) = \binom{3}{n_i} p_i^{n_i}(1-p_i)^{3-n_i}$ for $i = 1, 2, 3, 4$.

(b) For $n = 1$, $P_{no\ additional\ feature} = P_{N_1,N_2,N_3,N_4}(0,0,0,0) = \dfrac{315}{1024}$.

(c) For $n = 1$, $P_{at\ least\ 3\ additional\ features} = \sum_{\substack{o+p+q+r\geq 3 \\ o,p,q,r\in\{0,1\}}} P_{N_1,N_2,N_3,N_4}(o,p,q,r) = \dfrac{27}{1024}$.

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 5.3.2 Solution

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

Problem 5.3.6 Solution

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

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

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:

$$\{N_1 = n_1, N_2 = n_2\} = \left\{N_1 = n_1, N_2 = n_2, \hat{N} = n - n_1 - n_2\right\} \tag{3}$$

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

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

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

(b) We could find the PMF of T_i by summing the joint PMF $P_{N_1,\ldots,N_r}(n_1, \ldots, n_r)$. 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 $\{s_1, s_2, \ldots, s_i\}$ and otherwise a failure occurs. In this case, the success probability is $q_i = p_1 + \cdots + p_i$ and T_i is the number of successes in n trials. Thus, T_i has the binomial PMF

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

(c) The joint PMF of T_1 and T_2 satisfies

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

$$= P[N_1 = t_1, N_2 = t_2 - t_1] \tag{8}$$

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

By the result of part (a),

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

Problem 5.4.3 Solution

We will use the PDF

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

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

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

$$= \left(\int_0^1 dx_2\right)\left(\int_0^1 dx_3\right)\left(\int_0^1 dx_4\right) = 1. \tag{3}$$

Thus,

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

Following similar steps, one can show that

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

Thus

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

We conclude that X_1, X_2, X_3 and X_4 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

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

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 5.5.1 Solution

For discrete random vectors, it is true in general that

$$P_{\mathbf{Y}}(\mathbf{y}) = P[\mathbf{Y} = \mathbf{y}] = P[\mathbf{AX} + \mathbf{b} = \mathbf{y}] = P[\mathbf{AX} = \mathbf{y} - \mathbf{b}]. \tag{1}$$

For an arbitrary matrix $\mathbf{A}$, the system of equations $\mathbf{Ax} = \mathbf{y} - \mathbf{b}$ 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,

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

As an aside, we note that when $\mathbf{Ax} = \mathbf{y} - \mathbf{b}$ 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 $\mathbf{Ax} = \mathbf{y} - \mathbf{b}$. This can get disagreeably complicated.

Problem 5.5.6 Solution

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

$$\begin{aligned}
P[A] &= P[X_1 \leq X_n, X_2 \leq X_n, \cdots, X_{n-1} \leq X_n] & (1)\\
&= \int_{-\infty}^{\infty} P[X_1 < X_n, X_2 < X_n, \cdots, X_{n-1} < X_n | X_n = x] f_{X_n}(x)\, dx & (2)\\
&= \int_{-\infty}^{\infty} P[X_1 < x, X_2 < x, \cdots, X_{n-1} < x | X_n = x] f_X(x)\, dx & (3)
\end{aligned}$$

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

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

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

$$\begin{aligned}
P[A] &= \int_{-\infty}^{\infty} P[X_1 \leq x] P[X_2 \leq x] \cdots P[X_{n-1} \leq x] f_X(x)\, dx & (5)\\
&= \int_{-\infty}^{\infty} [F_X(x)]^{n-1} f_X(x)\, dx = \frac{1}{n} [F_X(x)]^n \Big|_{-\infty}^{\infty} = \frac{1}{n}(1 - 0) & (6)
\end{aligned}$$

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 5.6.4 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,

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

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

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

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

$$E[X_i X_j] = E[X_i] E[X_j] = 1/4. \tag{3}$$

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

$$E[\mathbf{X}] = \begin{bmatrix} E[X_1] & E[X_2] & E[X_3] & E[X_4] \end{bmatrix}' = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \end{bmatrix}'. \tag{4}$$

(b) The correlation matrix is

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

$$= \begin{bmatrix} 1/3 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/3 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/3 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/3 \end{bmatrix} \tag{6}$$

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

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

$$= \begin{bmatrix} 1/12 & 0 & 0 & 0 \\ 0 & 1/12 & 0 & 0 \\ 0 & 0 & 1/12 & 0 \\ 0 & 0 & 0 & 1/12 \end{bmatrix} \tag{8}$$

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

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

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

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

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

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

$$\mathbf{a}'\mathbf{R}_{\mathbf{X}}\mathbf{a} = E[W^2] \ge 0. \tag{3}$$

Problem 5.7.6 Solution

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

$$\mathbf{C_Y} = \mathbf{QC_XQ}' \tag{1}$$

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

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

We conclude that Y_1 and Y_2 have covariance

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

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 θ 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 θ 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 θ, 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.