

Question 1: A discrete random vector

The vector PMF and the component joint PMF describe the same model:

$$P_{\mathbf{X}}([i \ j]') = P_{X_1, X_2}(i, j) = \frac{i+j}{36}, \quad i, j \in \{1, 2, 3\}.$$

- (a) The table below lists all nine support pairs. Each cell shows the sum value $s = i + j$ and the corresponding joint probability $p_{ij} = P_{X_1, X_2}(i, j)$. The column sums are the marginal probabilities $P_{X_2}(j)$, so this same table also displays the marginalisation used in (b):

	$j = 1$	$j = 2$	$j = 3$
$i = 1$	$s = 2$ $p = 2/36$	$s = 3$ $p = 3/36$	$s = 4$ $p = 4/36$
$i = 2$	$s = 3$ $p = 3/36$	$s = 4$ $p = 4/36$	$s = 5$ $p = 5/36$
$i = 3$	$s = 4$ $p = 4/36$	$s = 5$ $p = 5/36$	$s = 6$ $p = 6/36$
column sum $P_{X_2}(j)$	9/36	12/36	15/36

For the PMF of $S = X_1 + X_2$, group the cells on each diagonal $i + j = s$:

s	2	3	4	5	6
Number of pairs	1	2	3	2	1
$P_S(s) = \text{sum of cells}$	1/18	1/6	1/3	5/18	1/6

Thus the PMF is zero for all other s , and the cells (1, 2) and (2, 1) give $\boxed{P(S = 3) = 1/6}$.

- (b) Read the column sums in the table above, or sum the joint PMF over X_1 :

$$P_{X_2}(j) = \sum_{i=1}^3 P_{X_1, X_2}(i, j) = \frac{(1+j) + (2+j) + (3+j)}{36} = \frac{j+2}{12}, \quad j \in \{1, 2, 3\},$$

$$P_{X_1|X_2}(i | j) = \frac{P_{X_1, X_2}(i, j)}{P_{X_2}(j)} = \frac{(i+j)/36}{(j+2)/12} = \boxed{\frac{i+j}{3(j+2)}}, \quad i, j \in \{1, 2, 3\}.$$

For each admissible j , the conditional PMF is zero when $i \notin \{1, 2, 3\}$; it is not defined by this ratio for inadmissible j .

- (c) The conditional expectation uses the conditional PMF from part (b):

$$E[X_1 | X_2 = j] = \sum_{i=1}^3 i P_{X_1|X_2}(i | j) = \frac{\sum_{i=1}^3 i^2 + j \sum_{i=1}^3 i}{3(j+2)} = \boxed{\frac{14+6j}{3(j+2)}}, \quad j \in \{1, 2, 3\}.$$

Independence would require $E[X_1 | X_2 = j] = E[X_1]$ for every admissible j , so this mean would be constant. Instead, the result $\frac{14+6j}{3(j+2)} = 2 + \frac{2}{3(j+2)}$ varies with j : it equals 20/9 at $j = 1$ and 32/15 at $j = 3$. Knowing X_2 therefore changes the expected value of X_1 , so X_1 and X_2 are dependent.

- (d) For a discrete vector, calculate the means and second moments from its joint PMF. The model is symmetric in i, j , so both means and both diagonal second moments are equal:

$$\begin{aligned} E[X_1] &= \sum_{i=1}^3 \sum_{j=1}^3 i \frac{i+j}{36} = \frac{3(14) + 6(6)}{36} = \frac{13}{6}, \\ E[X_1^2] &= \sum_{i=1}^3 \sum_{j=1}^3 i^2 \frac{i+j}{36} = \frac{3(36) + 6(14)}{36} = \frac{16}{3}, \\ E[X_1 X_2] &= \sum_{i=1}^3 \sum_{j=1}^3 ij \frac{i+j}{36} = \frac{14(6) + 6(14)}{36} = \frac{14}{3}. \end{aligned}$$

Hence the mean vector and the correlation matrix (the matrix of second moments) are

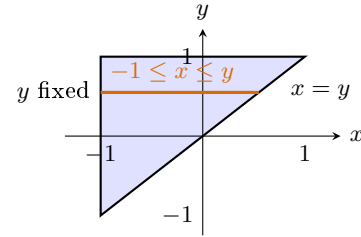
$$\boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} 13/6 \\ 13/6 \end{bmatrix}, \quad \mathbf{R}_{\mathbf{X}} = E[\mathbf{X}\mathbf{X}'] = \begin{bmatrix} 16/3 & 14/3 \\ 14/3 & 16/3 \end{bmatrix}.$$

Finally,

$$\mathbf{C}_{\mathbf{X}} = \mathbf{R}_{\mathbf{X}} - \boldsymbol{\mu}_{\mathbf{X}} \boldsymbol{\mu}_{\mathbf{X}}' = \frac{1}{36} \begin{bmatrix} 192 & 168 \\ 168 & 192 \end{bmatrix} - \frac{1}{36} \begin{bmatrix} 169 & 169 \\ 169 & 169 \end{bmatrix} = \boxed{\frac{1}{36} \begin{bmatrix} 23 & -1 \\ -1 & 23 \end{bmatrix}}.$$

Question 2: Conditioning on a value of a continuous random variable

The joint density is $1/2$ on the shaded triangle. For a fixed admissible y , the horizontal slice runs from $x = -1$ to $x = y$. Conditioning on $Y = y$ normalises the density along this slice, rather than selecting the two-dimensional region $Y > 0$ used in Tutorial 7.



- (a) Integrate out X over the slice at $Y = y$:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx = \int_{-1}^y \frac{1}{2} dx = \frac{y+1}{2}, \quad -1 \leq y \leq 1.$$

Thus

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

- (b) For an admissible conditioning value, divide the joint PDF by the marginal PDF:

$$f_{X|Y}(x | y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{1/2}{(y+1)/2} = \frac{1}{y+1}.$$

Since $f_Y(y) > 0$ is required, the admissible values are $-1 < y \leq 1$. For each such fixed y ,

$$f_{X|Y}(x | y) = \begin{cases} 1/(y+1), & -1 \leq x \leq y, \\ 0, & \text{otherwise.} \end{cases}$$

This is a uniform density on $[-1, y]$, with total area $(y+1)/(y+1) = 1$.

- (c) Use the conditional-expectation formula before substituting the density:

$$E[X | Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x | y) dx = \frac{1}{y+1} \int_{-1}^y x dx = \frac{y^2 - 1}{2(y+1)} = \boxed{\frac{y-1}{2}}.$$

Given $Y = y$, Y has the fixed value y ; only X remains random. Split the integral and take the constant y outside:

$$\begin{aligned} E[Y - X | Y = y] &= \int_{-\infty}^{\infty} (y - x) f_{X|Y}(x | y) dx \\ &= y \underbrace{\int_{-\infty}^{\infty} f_{X|Y}(x | y) dx}_{=1} - \underbrace{\int_{-\infty}^{\infty} x f_{X|Y}(x | y) dx}_{=E[X|Y=y]} \\ &= y - E[X | Y = y] \\ &= y - \frac{y-1}{2} = \frac{2y - y + 1}{2} = \boxed{\frac{y+1}{2}}. \end{aligned}$$

The first integral is 1 because it integrates a conditional PDF over its whole support. Both formulas apply for $-1 < y \leq 1$.

- (d) Iterated expectation averages the conditional mean over the distribution of Y :

$$\begin{aligned} E[X] &= E[E[X | Y]] = \int_{-\infty}^{\infty} E[X | Y = y] f_Y(y) dy \\ &= \int_{-1}^1 \frac{y-1}{2} \frac{y+1}{2} dy = \frac{1}{4} \left[\frac{y^3}{3} - y \right]_{-1}^1 = \boxed{-\frac{1}{3}}. \end{aligned}$$

The undefined conditional value at the single endpoint $y = -1$ does not affect this integral.

Question 3: Gaussian conditioning in an industrial monitor

The alarm is triggered only when *both* thresholds are exceeded: $I = \{T > 38, W > 10\}$. The given standard deviations are $\sigma_W = 2$ volts and $\sigma_T = 1$ degree Celsius, so the unconditional variances are $\text{Var}[W] = 4$ and $\text{Var}[T] = 1$. Throughout, Φ denotes the standard Gaussian CDF.

- (a) If $\rho_{W,T} = 0$, the variables are independent *because they are jointly Gaussian*. The events $\{T > 38\}$ and $\{W > 10\}$ are therefore independent: the probability that both occur is the product of their individual probabilities. Zero correlation does not imply independence for arbitrary random variables. For a Gaussian variable, standardising gives

$$P(W > w) = 1 - \Phi\left(\frac{w - \mu_W}{\sigma_W}\right), \quad P(T > t) = 1 - \Phi\left(\frac{t - \mu_T}{\sigma_T}\right).$$

Therefore,

$$\begin{aligned} P(I) &= P(T > 38, W > 10) = P(T > 38) P(W > 10) \\ &= \left[1 - \Phi\left(\frac{38 - 37}{1}\right)\right] \left[1 - \Phi\left(\frac{10 - 7}{2}\right)\right] \\ &= [1 - \Phi(1)] [1 - \Phi(1.5)] \\ &\approx (1 - 0.8413)(1 - 0.9332) = 0.01060116 \approx \boxed{0.0106}. \end{aligned}$$

The CDF values used in the numerical approximation are rounded to four decimal places.

- (b) Now let $\rho_{W,T} = 1/\sqrt{2}$. The conditional distribution of W given $T = t$ is Gaussian. Its mean and variance follow the Lecture 20 formulas:

$$\begin{aligned} \tilde{\mu}_W(t) &= E[W | T = t] = \mu_W + \rho_{W,T} \frac{\sigma_W}{\sigma_T} (t - \mu_T) \\ &= 7 + \frac{1}{\sqrt{2}} \frac{2}{1} (t - 37) = \boxed{7 + \sqrt{2}(t - 37)}, \\ \tilde{\sigma}_W^2 &= \text{Var}[W | T = t] = \sigma_W^2 (1 - \rho_{W,T}^2) = 4 \left(1 - \frac{1}{2}\right) = \boxed{2}. \end{aligned}$$

Substitute these parameters into the scalar Gaussian PDF:

$$\begin{aligned} f_{W|T}(w | t) &= \frac{1}{\tilde{\sigma}_W \sqrt{2\pi}} \exp\left[-\frac{(w - \tilde{\mu}_W(t))^2}{2\tilde{\sigma}_W^2}\right] \\ &= \boxed{\frac{1}{\sqrt{4\pi}} \exp\left[-\frac{[w - 7 - \sqrt{2}(t - 37)]^2}{4}\right]}, \quad w, t \in \mathbb{R}. \end{aligned}$$

- (c) When $t \leq 38$, the temperature threshold is not exceeded, so the alarm cannot trigger, whatever W is. When $t > 38$, substitute the definition $I = \{T > 38, W > 10\}$ into the conditional probability. Since $T = t$ is given and $t > 38$, the condition $T > 38$ is certain and therefore drops out:

$$\begin{aligned} P(I | T = t) &= P(\{T > 38, W > 10\} | T = t) \\ &= P(W > 10 | T = t), \quad t > 38 \\ &= 1 - \Phi\left(\frac{10 - \tilde{\mu}_W(t)}{\tilde{\sigma}_W}\right) = 1 - \Phi\left(\frac{10 - [7 + \sqrt{2}(t - 37)]}{\sqrt{2}}\right), \quad t > 38. \end{aligned}$$

The complete answer, including the temperature gate, is

$$P(I | T = t) = \begin{cases} 0, & t \leq 38, \\ 1 - \Phi\left(\frac{3}{\sqrt{2}} - (t - 37)\right), & t > 38. \end{cases}$$

- (d) The unconditional variance is $\text{Var}[W] = \sigma_W^2 = 4$, whereas $\text{Var}[W | T = t] = 2$: observing T halves the uncertainty measured by variance. The conditional standard deviation decreases from 2 to $\sqrt{2}$ volts; it is not halved. In this jointly Gaussian model the conditional mean changes with t , but the conditional variance is the same for every observed t .

Question 4: A Gaussian random vector

- (a) The module uses the correlation matrix $\mathbf{R}_\mathbf{X} = \mathbb{E}[\mathbf{X}\mathbf{X}']$; its entries are second moments, not correlation coefficients.

$$\begin{aligned}\mathbf{R}_\mathbf{X} &= \mathbf{C}_\mathbf{X} + \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}_\mathbf{X}' \\ &= \begin{bmatrix} 4 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 4 \end{bmatrix} + \begin{bmatrix} 16 & 32 & 24 \\ 32 & 64 & 48 \\ 24 & 48 & 36 \end{bmatrix} = \boxed{\begin{bmatrix} 20 & 30 & 25 \\ 30 & 68 & 46 \\ 25 & 46 & 40 \end{bmatrix}}.\end{aligned}$$

- (b) A subvector of a Gaussian vector is also Gaussian. For $\mathbf{Z} = [X_1 \ X_2]'$, select the corresponding entries of the mean vector and covariance matrix:

$$\boldsymbol{\mu}_\mathbf{Z} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}, \quad \mathbf{C}_\mathbf{Z} = \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix}, \quad \det(\mathbf{C}_\mathbf{Z}) = 12, \quad \mathbf{C}_\mathbf{Z}^{-1} = \frac{1}{12} \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}.$$

The two-dimensional Gaussian density formula is

$$f_\mathbf{Z}(\mathbf{z}) = \frac{1}{2\pi\sqrt{\det(\mathbf{C}_\mathbf{Z})}} \exp\left[-\frac{1}{2}(\mathbf{z} - \boldsymbol{\mu}_\mathbf{Z})' \mathbf{C}_\mathbf{Z}^{-1} (\mathbf{z} - \boldsymbol{\mu}_\mathbf{Z})\right].$$

Here $\mathbf{z} = [x_1 \ x_2]'$, and, writing $u = x_1 - 4$ and $v = x_2 - 8$, the matrix multiplication inside the quadratic form is

$$\mathbf{C}_\mathbf{Z}^{-1}(\mathbf{z} - \boldsymbol{\mu}_\mathbf{Z}) = \frac{1}{12} \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 2u + v \\ u + 2v \end{bmatrix}.$$

Thus,

$$(\mathbf{z} - \boldsymbol{\mu}_\mathbf{Z})' \mathbf{C}_\mathbf{Z}^{-1} (\mathbf{z} - \boldsymbol{\mu}_\mathbf{Z}) = \frac{1}{6} \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} 2u + v \\ u + 2v \end{bmatrix} = \frac{u^2 + uv + v^2}{3} = \frac{(x_1 - 4)^2 + (x_1 - 4)(x_2 - 8) + (x_2 - 8)^2}{3}.$$

Therefore,

$$f_{X_1, X_2}(x_1, x_2) = \frac{1}{4\pi\sqrt{3}} \exp\left[-\frac{(x_1 - 4)^2 + (x_1 - 4)(x_2 - 8) + (x_2 - 8)^2}{6}\right],$$

for all $(x_1, x_2) \in \mathbb{R}^2$. This is called the *centred form* because it uses the deviations $x_1 - 4$ and $x_2 - 8$ from the respective means. Multiplying out the squares and product gives

$$\begin{aligned}(x_1 - 4)^2 + (x_1 - 4)(x_2 - 8) + (x_2 - 8)^2 \\ &= (x_1^2 - 8x_1 + 16) + (x_1x_2 - 8x_1 - 4x_2 + 32) + (x_2^2 - 16x_2 + 64) \\ &= x_1^2 + x_1x_2 + x_2^2 - 16x_1 - 20x_2 + 112.\end{aligned}$$

Thus an equivalent answer, with the exponent fully expanded, is

$$f_{X_1, X_2}(x_1, x_2) = \frac{1}{4\pi\sqrt{3}} \exp\left[-\frac{x_1^2}{6} - \frac{x_1x_2}{6} - \frac{x_2^2}{6} + \frac{8}{3}x_1 + \frac{10}{3}x_2 - \frac{56}{3}\right].$$

Both expressions give exactly the same PDF on $\mathbb{R}^2$; only the algebraic form differs.

- (c) The marginal X_1 is Gaussian with mean $\mu_{X_1} = 4$, variance $\sigma_{X_1}^2 = 4$ and standard deviation $\sigma_{X_1} = 2$. Writing Φ for the CDF of a Gaussian variable with mean 0 and variance 1,

$$\begin{aligned}P(X_1 > a) &= 1 - \Phi\left(\frac{a - \mu_{X_1}}{\sigma_{X_1}}\right), \\ P(X_1 > 8) &= 1 - P(X_1 \leq 8) = 1 - \Phi\left(\frac{8 - 4}{2}\right) = 1 - \Phi(2) \approx \boxed{0.0228}.\end{aligned}$$

- (d) Components of a jointly Gaussian vector are independent if and only if their covariance matrix is diagonal. Here, for example, $\text{Cov}[X_1, X_2] = -2 \neq 0$, so the components are **not independent**. In fact, every pair has a nonzero covariance. Compare Further practice, Question 8(c): outside the jointly Gaussian case, even a diagonal covariance matrix does not ensure independence.

Question 5: An affine Gaussian transformation

An affine transformation of a Gaussian vector is Gaussian. Here $\mathbf{A}$ has full row rank, so $\mathbf{Y}$ has a nondegenerate two-dimensional Gaussian distribution. The constant offset $\mathbf{b}$ changes the mean but not the covariance.

- (a) The mean vector is

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

- (b) The covariance formula and the intermediate product give

$$\mathbf{C}_{\mathbf{Y}} = \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}', \quad \mathbf{A}\mathbf{C}_{\mathbf{X}} = \frac{1}{3} \begin{bmatrix} 11 & -4 & 8 \\ 17 & -16 & 14 \end{bmatrix}.$$

$$\mathbf{C}_{\mathbf{Y}} = \frac{1}{3} \begin{bmatrix} 11 & -4 & 8 \\ 17 & -16 & 14 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1/2 & -1/2 \\ 2/3 & 2/3 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 43 & 55 \\ 55 & 103 \end{bmatrix}.$$

- (c) The correlation matrix is

$$\mathbf{R}_{\mathbf{Y}} = \mathbf{C}_{\mathbf{Y}} + \boldsymbol{\mu}_{\mathbf{Y}}\boldsymbol{\mu}_{\mathbf{Y}}' = \frac{1}{9} \begin{bmatrix} 43 & 55 \\ 55 & 103 \end{bmatrix} + \begin{bmatrix} 64 & 0 \\ 0 & 0 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 619 & 55 \\ 55 & 103 \end{bmatrix}.$$

Equivalently, the lecture's direct affine correlation formula gives the same result:

$$\mathbf{R}_{\mathbf{Y}} = \mathbf{A}\mathbf{R}_{\mathbf{X}}\mathbf{A}' + (\mathbf{A}\boldsymbol{\mu}_{\mathbf{X}})\mathbf{b}' + \mathbf{b}(\mathbf{A}\boldsymbol{\mu}_{\mathbf{X}})' + \mathbf{b}\mathbf{b}'.$$

- (d) Since $\mathbf{Y}$ is an affine transformation of a Gaussian vector, it is also Gaussian; its component Y_2 is therefore Gaussian with mean 0 and variance $103/9$, hence $\sigma_{Y_2} = \sqrt{103}/3$. For a Gaussian random variable,

$$P(a \leq Y_2 \leq b) = \Phi\left(\frac{b - \mu_{Y_2}}{\sigma_{Y_2}}\right) - \Phi\left(\frac{a - \mu_{Y_2}}{\sigma_{Y_2}}\right),$$

$$P(-1 \leq Y_2 \leq 1) = \Phi\left(\frac{3}{\sqrt{103}}\right) - \Phi\left(-\frac{3}{\sqrt{103}}\right) = 2\Phi\left(\frac{3}{\sqrt{103}}\right) - 1 \approx \boxed{0.2325}.$$

- (e) The cross-covariance matrix has rows indexed by X_i and columns by Y_j :

$$\mathbf{C}_{\mathbf{X}\mathbf{Y}} = E[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})'] = \mathbf{C}_{\mathbf{X}}\mathbf{A}'$$

$$= (\mathbf{A}\mathbf{C}_{\mathbf{X}})' = \frac{1}{3} \begin{bmatrix} 11 & 17 \\ -4 & -16 \\ 8 & 14 \end{bmatrix}.$$

- (f) The cross-correlation matrix is

$$\mathbf{R}_{\mathbf{X}\mathbf{Y}} = E[\mathbf{X}\mathbf{Y}'] = \mathbf{C}_{\mathbf{X}\mathbf{Y}} + \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}_{\mathbf{Y}}'$$

$$= \frac{1}{3} \begin{bmatrix} 11 & 17 \\ -4 & -16 \\ 8 & 14 \end{bmatrix} + \begin{bmatrix} 4 \\ 8 \\ 6 \end{bmatrix} \begin{bmatrix} 8 & 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 107 & 17 \\ 188 & -16 \\ 152 & 14 \end{bmatrix}.$$

An equivalent direct route is $\mathbf{R}_{\mathbf{X}\mathbf{Y}} = \mathbf{R}_{\mathbf{X}}\mathbf{A}' + \boldsymbol{\mu}_{\mathbf{X}}\mathbf{b}'$.

- (g) Use entry (1,2) of $\mathbf{C}_{\mathbf{X}\mathbf{Y}}$, entry (1,1) of $\mathbf{C}_{\mathbf{X}}$ and entry (2,2) of $\mathbf{C}_{\mathbf{Y}}$:

$$\rho_{X_1, Y_2} = \frac{\text{Cov}[X_1, Y_2]}{\sqrt{\text{Var}[X_1]}\sqrt{\text{Var}[Y_2]}} = \frac{17/3}{\sqrt{4(103/9)}} = \frac{17}{2\sqrt{103}} \approx 0.83753.$$

- (h) The inverse-covariance rule requires a square, invertible transformation. Here $\mathbf{A}$ is 2×3 , so $\mathbf{A}^{-1}$ does not exist. The three-dimensional input has been mapped to two dimensions; its full covariance cannot be recovered from $\mathbf{C}_{\mathbf{Y}}$ by that rule. This differs from the invertible 2×2 transformation in Further practice, Question 8.

Further practice

Verdere oefening

Question 6

- (a) To obtain a marginal density, integrate out the components that are not retained:

$$\begin{aligned} f_{X_1, X_2}(x_1, x_2) &= \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2, x_3) dx_3 \\ &= \frac{1}{3} \int_0^1 (x_1 + 2x_2 + 3x_3) dx_3 = \boxed{\frac{x_1 + 2x_2 + 3/2}{3}}, \quad 0 \leq x_1, x_2 \leq 1, \\ f_{X_3}(x_3) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2, x_3) dx_1 dx_2 \\ &= \frac{1}{3} \int_0^1 \int_0^1 (x_1 + 2x_2 + 3x_3) dx_1 dx_2 = \boxed{x_3 + \frac{1}{2}}, \quad 0 \leq x_3 \leq 1. \end{aligned}$$

Independence would require the joint density to factorise as $f_{\mathbf{X}} = f_{X_1, X_2} f_{X_3}$. However, on the unit cube,

$$f_{X_1, X_2}(x_1, x_2) f_{X_3}(x_3) = \frac{(x_1 + 2x_2 + 3/2)(x_3 + 1/2)}{3} \neq \frac{x_1 + 2x_2 + 3x_3}{3}$$

as functions of (x_1, x_2, x_3) . The product contains cross terms such as $x_1 x_3$ that are absent from the joint density. Therefore, **the vector and X_3 are not independent.**

- (b) The vector CDF requires all three inequalities to hold together:

$$\begin{aligned} F_{\mathbf{X}}(x_1, x_2, x_3) &= P(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3) \\ &= \int_0^{x_3} \int_0^{x_2} \int_0^{x_1} \frac{t_1 + 2t_2 + 3t_3}{3} dt_1 dt_2 dt_3 \\ &= \frac{1}{3} \int_0^{x_3} \int_0^{x_2} \left(\frac{x_1^2}{2} + 2x_1 t_2 + 3x_1 t_3 \right) dt_2 dt_3 \\ &= \frac{1}{3} \int_0^{x_3} \left(\frac{x_1^2 x_2}{2} + x_1 x_2^2 + 3x_1 x_2 t_3 \right) dt_3 \\ &= \frac{1}{3} \left(\frac{x_1^2 x_2 x_3}{2} + x_1 x_2^2 x_3 + \frac{3x_1 x_2 x_3^2}{2} \right) \\ &= \boxed{\frac{x_1 x_2 x_3 (x_1 + 2x_2 + 3x_3)}{6}}, \quad 0 \leq x_1, x_2, x_3 \leq 1. \end{aligned}$$

Thus,

$$F_{\mathbf{X}}(0.3, 0.5, 1) = \frac{(0.3)(0.5)(1)(0.3 + 2(0.5) + 3(1))}{6} = \boxed{\frac{43}{400} = 0.1075}.$$

- (c) For a function of a continuous random vector, integrate the function against its joint PDF:

$$\begin{aligned} E[g(\mathbf{X})] &= \iiint g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \\ E[X_1 X_2 X_3] &= \frac{1}{3} \int_0^1 \int_0^1 \int_0^1 x_1 x_2 x_3 (x_1 + 2x_2 + 3x_3) dx_1 dx_2 dx_3 \\ &= \frac{1}{3} \left[\left(\int_0^1 x_1^2 dx_1 \right) \left(\int_0^1 x_2 dx_2 \right) \left(\int_0^1 x_3 dx_3 \right) \right. \\ &\quad + 2 \left(\int_0^1 x_1 dx_1 \right) \left(\int_0^1 x_2^2 dx_2 \right) \left(\int_0^1 x_3 dx_3 \right) \\ &\quad \left. + 3 \left(\int_0^1 x_1 dx_1 \right) \left(\int_0^1 x_2 dx_2 \right) \left(\int_0^1 x_3^2 dx_3 \right) \right] \\ &= \frac{1}{3} \left(\frac{1}{12} + \frac{1}{6} + \frac{1}{4} \right) = \boxed{\frac{1}{6}}. \end{aligned}$$

Question 7

Each completion time has

$$F_T(t) = \begin{cases} 0, & t < 0, \\ t/2, & 0 \leq t \leq 2, \\ 1, & t > 2, \end{cases} \quad f_T(t) = \begin{cases} 1/2, & 0 < t < 2, \\ 0, & \text{otherwise.} \end{cases}$$

- (a) The batch is complete by time m only if *all three* jobs are complete by m . Independence therefore gives

$$\begin{aligned} F_M(m) &= P(\max\{T_1, T_2, T_3\} \leq m) \\ &= P(T_1 \leq m, T_2 \leq m, T_3 \leq m) = \prod_{i=1}^3 P(T_i \leq m) = [F_T(m)]^3. \end{aligned}$$

Differentiate the CDF to obtain the PDF:

$$F_M(m) = \begin{cases} 0, & m < 0, \\ m^3/8, & 0 \leq m \leq 2, \\ 1, & m > 2, \end{cases} \quad f_M(m) = \begin{cases} 3m^2/8, & 0 < m < 2, \\ 0, & \text{otherwise.} \end{cases}$$

- (b) No job has finished by time ℓ exactly when *all three* completion times exceed ℓ . It is therefore convenient to use the complementary event:

$$\begin{aligned} F_L(\ell) &= 1 - P(L > \ell) \\ &= 1 - P(T_1 > \ell, T_2 > \ell, T_3 > \ell) = 1 - \prod_{i=1}^3 P(T_i > \ell) \\ &= 1 - [1 - F_T(\ell)]^3. \end{aligned}$$

Hence,

$$F_L(\ell) = \begin{cases} 0, & \ell < 0, \\ 1 - (1 - \ell/2)^3, & 0 \leq \ell \leq 2, \\ 1, & \ell > 2, \end{cases} \quad f_L(\ell) = \begin{cases} \frac{3}{2}(1 - \ell/2)^2, & 0 < \ell < 2, \\ 0, & \text{otherwise.} \end{cases}$$

The first-completion density is largest near 0, while the batch-completion density is largest near 2. This agrees with the meanings of a minimum and maximum.

- (c) For all three jobs to be complete by 1.5, the last completion time must be at most 1.5. This is the event $M = \max\{T_1, T_2, T_3\} \leq 1.5$, so use $F_M(1.5)$. For at least one job to be complete by 0.5, the first completion time must be at most 0.5. This is the event $L = \min\{T_1, T_2, T_3\} \leq 0.5$, so use $F_L(0.5)$:

$$\begin{aligned} P(M \leq 1.5) &= F_M(1.5) = \left(\frac{1.5}{2}\right)^3 = \boxed{\frac{27}{64} = 0.421875}, \\ P(L \leq 0.5) &= F_L(0.5) = 1 - \left(1 - \frac{0.5}{2}\right)^3 = \boxed{\frac{37}{64} = 0.578125}. \end{aligned}$$

- (d) For independent random variables, the expectation of their product factorises:

$$E[T_1 T_2 T_3] = E[T_1] E[T_2] E[T_3].$$

Here each mean is

$$E[T_i] = \int_{-\infty}^{\infty} t f_T(t) dt = \int_0^2 \frac{t}{2} dt = \left[\frac{t^2}{4}\right]_0^2 = 1 \text{ s.}$$

Therefore,

$$\boxed{E[T_1 T_2 T_3] = 1 \cdot 1 \cdot 1 = 1 \text{ s}^3}.$$

Unlike Question 6(c), this shortcut is justified because the three components are independent.

Question 8

- (a) From the two component equations,

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}, \quad \mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \det(\mathbf{A}) = -2.$$

Since the determinant is nonzero, the transformation is invertible:

$$\mathbf{X} = \mathbf{A}^{-1}(\mathbf{Y} - \mathbf{b}), \quad \mathbf{A}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad x_1 = \frac{y_1 - 1 + y_2}{2}, \quad x_2 = \frac{y_1 - 1 - y_2}{2}.$$

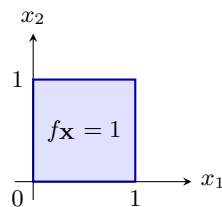
- (b) Independence and the uniform input densities give $f_{\mathbf{X}}(\mathbf{x}) = 1$ on the unit square and 0 elsewhere. The support transforms with the variables:

$$S_{\mathbf{Y}} = \{(y_1, y_2) : 0 \leq y_1 - 1 + y_2 \leq 2, \quad 0 \leq y_1 - 1 - y_2 \leq 2\}.$$

This is the parallelogram with vertices $(1, 0), (2, 1), (3, 0), (2, -1)$. The change-of-variables formula is

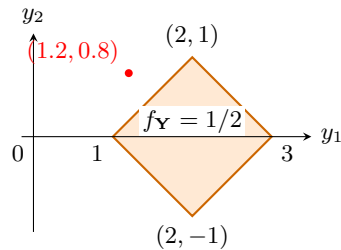
$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{|\det(\mathbf{A})|} f_{\mathbf{X}}(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})) = \begin{cases} 1/2, & \mathbf{y} \in S_{\mathbf{Y}}, \\ 0, & \text{otherwise.} \end{cases}$$

Input: area 1



$\mathbf{A}\mathbf{X} + \mathbf{b}$

Output: area 2



The output density integrates to $(1/2) \times 2 = 1$, as required.

- (c) Here $\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is the 2×2 identity matrix. Constant offsets do not affect covariance. Applying the forward and inverse rules,

$$\mathbf{C}_{\mathbf{Y}} = \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}' = \frac{1}{12} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{6} \end{bmatrix},$$

$$\mathbf{C}_{\mathbf{X}} = \mathbf{A}^{-1}\mathbf{C}_{\mathbf{Y}}(\mathbf{A}^{-1})' = \frac{1}{24} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{12} & 0 \\ 0 & \frac{1}{12} \end{bmatrix}.$$

Thus $\text{Cov}[Y_1, Y_2] = 0$, but independence still needs to be checked. The parallelogram in (b) makes the marginal step explicit: integrate its constant height $1/2$ over a vertical or horizontal slice. At the chosen values,

$$f_{Y_1}(1.2) = \frac{1}{2} 2(1.2 - 1) = 0.2 > 0, \quad f_{Y_2}(0.8) = \frac{1}{2} 2(1 - |0.8|) = 0.2 > 0.$$

Hence $f_{Y_1}(1.2)f_{Y_2}(0.8) = 0.04 > 0$. However, the marked point $(1.2, 0.8)$ lies outside the parallelogram because $y_1 - 1 - y_2 = -0.6 < 0$, so $f_{\mathbf{Y}}(1.2, 0.8) = 0$. Therefore $f_{\mathbf{Y}} \neq f_{Y_1}f_{Y_2}$: uncorrelated components need not be independent, and these variables are not jointly Gaussian.

- (d) Here $\mathbf{V} = 2\mathbf{X} + [1 \ 1]'$. Because the scale factor is positive, both inequalities keep their direction:

$$\begin{aligned} F_{\mathbf{V}}(v_1, v_2) &= P(2X_1 + 1 \leq v_1, \quad 2X_2 + 1 \leq v_2) \\ &= F_{\mathbf{X}}\left(\frac{v_1 - 1}{2}, \frac{v_2 - 1}{2}\right). \end{aligned}$$

For example, inside $[1, 3]^2$ this is $(v_1 - 1)(v_2 - 1)/4$. The determinant is $\det(2\mathbf{I}_2) = 2^2 = 4$, so

$$f_{\mathbf{V}}(v_1, v_2) = \frac{1}{4} f_{\mathbf{X}}\left(\frac{v_1 - 1}{2}, \frac{v_2 - 1}{2}\right) = \begin{cases} 1/4, & 1 \leq v_1 \leq 3, \quad 1 \leq v_2 \leq 3, \\ 0, & \text{otherwise.} \end{cases}$$

The factor is $1/2^2$, not $1/2$, because *both* coordinates have been scaled by 2.