THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
Exercises on Statistics

1 Expectation, Variance, Standard Deviation

1.1 Definitions for Discrete Random Variables

The *expected value (expectation)* of a random variable is the probability-weighted average of all possible outcomes of a probabilistic experiment. Variance quantifies the spread of the random variable's values around its expectation. Standard deviation is the positive square root of variance, with identical units as the random variable.

Let X be a discrete random variable with support $\{x_1, x_2, \ldots, x_n\}$ and probability mass function (PMF) $P(X = x_i) = p_i$, where $\sum_{i=1}^{n} p_i = 1,\ p_i \geq 0$.

(1) Expectation:

$$\mathbb{E}[X] = \mu_X = \sum_{i=1}^{n} x_i\, p_i$$

(2) Variance (two equivalent formulas):

$$\mathrm{Var}(X) = \sigma_X^2 = \mathbb{E}\big[(X - \mu_X)^2\big] = \sum_{i=1}^{n} (x_i - \mu_X)^2 p_i$$

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \big(\mathbb{E}[X]\big)^2, \quad \mathbb{E}[X^2] = \sum_{i=1}^{n} x_i^2 p_i$$

(3) Standard deviation:

$$\mathrm{SD}(X) = \sigma_X = \sqrt{\mathrm{Var}(X)}$$

1.2 Example 1: Discrete Random Variable X

Given PMF table for discrete X:

x	0	1	2	3	Sum
$P(X = x)$	0.010	0.840	0.145	0.005	1

1.2.1 Numerical Computation

1: Compute $\mathbb{E}[X]$

$$\begin{aligned}\mathbb{E}[X] &= (0)(0.010) + (1)(0.840) + (2)(0.145) + (3)(0.005) \\ &= 0 + 0.840 + 0.290 + 0.015 \\ &= 1.145\end{aligned}$$

2: Compute $\mathbb{E}[X^2]$ (for variance shortcut formula)

$$\begin{aligned}\mathbb{E}[X^2] &= (0^2)(0.010) + (1^2)(0.840) + (2^2)(0.145) + (3^2)(0.005)\\ &= 0 + (1)(0.840) + (4)(0.145) + (9)(0.005)\\ &= 0 + 0.840 + 0.580 + 0.045\\ &= 1.465\end{aligned}$$

3: Compute Variance $\mathrm{Var}(X)$

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}[X^2] - \big(\mathbb{E}[X]\big)^2\\ &= 1.465 - (1.145)^2\\ &= 1.465 - 1.311025\\ &= 0.153975\end{aligned}$$

Verification via direct squared deviation sum:

$$\begin{aligned}\sum (x_i - \mu_X)^2 p_i &= (0 - 1.145)^2(0.010) + (1 - 1.145)^2(0.840)\\ &\quad + (2 - 1.145)^2(0.145) + (3 - 1.145)^2(0.005)\\ &= (1.311025)(0.010) + (0.021025)(0.840)\\ &\quad + (0.731025)(0.145) + (3.441025)(0.005)\\ &= 0.01311025 + 0.017661 + 0.105998625 + 0.017205125\\ &= 0.153975 \quad \text{(matches shortcut result)}\end{aligned}$$

4: Compute Standard Deviation σ_X

$$\sigma_X = \sqrt{0.153975} \approx 0.39239648$$

1.3 Continuous Random Variable Definitions

Let X be a continuous random variable with probability density function (PDF) $f_X(t)$, supported on interval $[a, b]$, satisfying $\int_a^b f_X(t)dt = 1,\ f_X(t) \geq 0$.

(1) Expectation:

$$\mathbb{E}[X] = \mu_X = \int_{-\infty}^{\infty} t\, f_X(t)dt$$

(2) Variance:

$$\mathrm{Var}(X) = \mathbb{E}\big[(X - \mu_X)^2\big] = \int_{-\infty}^{\infty} (t - \mu_X)^2 f_X(t)dt = \mathbb{E}[X^2] - (\mathbb{E}[X])^2, \quad \mathbb{E}[X^2] = \int_{-\infty}^{\infty} t^2 f_X(t)dt$$

(3) Standard deviation:

$$\sigma_X = \sqrt{\mathrm{Var}(X)}$$

1.4 Expectation & Variance Properties

Let $a, b \in \mathbb{R}$ be constants, X, Y random variables:

1. $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$; $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
2. $\text{Var}(aX + b) = a^2 \text{Var}(X)$; $\text{Var}(b) = 0$ (constant has zero variance)
3. Standardized random variable: define $Z = \dfrac{X - \mu_X}{\sigma_X}$. By properties 1 and 2:

$$\mathbb{E}[Z] = \mathbb{E}\left[\frac{X - \mu_X}{\sigma_X}\right] = \frac{1}{\sigma_X}\big(\mathbb{E}[X] - \mu_X\big) = 0$$

$$\text{Var}(Z) = \text{Var}\left(\frac{X - \mu_X}{\sigma_X}\right) = \frac{1}{\sigma_X^2}\text{Var}(X) = \frac{\sigma_X^2}{\sigma_X^2} = 1$$

Z always has mean 0, variance 1.

1.5 Example 2: Continuous Random Variable X, PDF $f(t) = 3t^2,\ 0 \le t \le 1$

Define $Y = 4X + 2$. Compute $\mathbb{E}[X], \text{Var}(X), \mathbb{E}[Y], \text{Var}(Y)$.

1.5.1 Analytical Integration

1: Compute $\mathbb{E}[X]$

$$\mathbb{E}[X] = \int_0^1 t \cdot f(t)dt = \int_0^1 t \cdot 3t^2 dt = 3\int_0^1 t^3 dt = 3\left.\frac{t^4}{4}\right|_0^1 = \frac{3}{4} = 0.75$$

2: Compute $\mathbb{E}[X^2]$

$$\mathbb{E}[X^2] = \int_0^1 t^2 \cdot 3t^2 dt = 3\int_0^1 t^4 dt = 3\left.\frac{t^5}{5}\right|_0^1 = \frac{3}{5} = 0.6$$

3: Variance of X

$$\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = \frac{3}{5} - \frac{9}{16} = \frac{48 - 45}{80} = \frac{3}{80} = 0.0375$$

4: Expectation of $Y = 4X + 2$ (linear property)

$$\mathbb{E}[Y] = \mathbb{E}[4X + 2] = 4\mathbb{E}[X] + 2 = 4 \cdot \frac{3}{4} + 2 = 3 + 2 = 5$$

5: Variance of $Y = 4X + 2$ (variance scaling property)

$$\text{Var}(Y) = 4^2 \text{Var}(X) = 16 \cdot \frac{3}{80} = \frac{48}{80} = \frac{3}{5} = 0.6$$

Integral verification for $\mathbb{E}[Y^2]$ (cross-check):

$$\begin{aligned}
\mathbb{E}[Y^2] &= \int_0^1 (4t + 2)^2 \cdot 3t^2 dt = 3\int_0^1 (16t^2 + 16t + 4)t^2 dt \\
&= 3\int_0^1 (16t^4 + 16t^3 + 4t^2)dt = 3\left.\left(\frac{16t^5}{5} + \frac{16t^4}{4} + \frac{4t^3}{3}\right)\right|_0^1 \\
&= 3\left(\frac{16}{5} + 4 + \frac{4}{3}\right) = 3\left(\frac{48 + 60 + 20}{15}\right) = 3 \cdot \frac{128}{15} = \frac{128}{5} = 25.6
\end{aligned}$$

$$\mathrm{Var}(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = 25.6 - 25 = 0.6 \quad \text{(matches property result)}$$

2 Joint Probability Distribution

For two discrete r.v.s X, Y with joint PMF $p(x, y) = P(X = x, Y = y)$:

$$\sum_x \sum_y p(x, y) = 1, \quad p(x, y) \geq 0$$

Marginal PMFs:

$$p_X(x) = \sum_y p(x, y), \quad p_Y(y) = \sum_x p(x, y)$$

For continuous r.v.s with joint PDF $f(x, y)$:

$$\iint_{\mathbb{R}^2} f(x, y) dx dy = 1, \ f(x, y) \geq 0$$

Marginals:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Joint expectation: $\mathbb{E}[g(X, Y)] = \sum_x \sum_y g(x, y) p(x, y)$ (discrete), $\iint g(x, y) f(x, y) dx dy$ (continuous).

3 Covariance, Correlation Coefficient

3.1 Covariance Definition

$$\mathrm{Cov}(X, Y) = \mathbb{E}\big[(X - \mu_X)(Y - \mu_Y)\big] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Properties:

1. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
2. $\mathrm{Cov}(aX + b, cY + d) = ac\,\mathrm{Cov}(X, Y)$
3. $\mathrm{Cov}(X_1 + X_2, Y) = \mathrm{Cov}(X_1, Y) + \mathrm{Cov}(X_2, Y)$

3.2 Pearson Correlation Coefficient

$$\rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}, \quad -1 \leq \rho_{XY} \leq 1$$

- $\rho = 1$: perfect positive linear relation

- $\rho = -1$: perfect negative linear relation

- $\rho = 0$: uncorrelated (no linear association; independent $\implies$ uncorrelated, converse false)

4 Covariance Matrix

For random vector $\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix}$, mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{bmatrix}$, $\mu_i = \mathbb{E}[X_i]$. Covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{k\times k}$:

$$\mathbf{\Sigma} = \begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) & \dots & \mathrm{Cov}(X_1, X_k) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) & \dots & \mathrm{Cov}(X_2, X_k) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_k, X_1) & \mathrm{Cov}(X_k, X_2) & \dots & \mathrm{Var}(X_k) \end{bmatrix}$$

Properties: symmetric ($\mathbf{\Sigma} = \mathbf{\Sigma}^T$), positive semi-definite.

4.1 Exercise 1

Select a continuous PDF $f(x)$ from a statistics textbook (e.g., uniform, exponential, normal, beta). Repeat steps:

1. Verify $\int f(x)dx = 1$ (valid PDF check)
2. Compute $\mathbb{E}[X] = \int xf(x)dx$
3. Compute $\mathbb{E}[X^2] = \int x^2 f(x)dx$
4. $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$
5. $\sigma_X = \sqrt{\mathrm{Var}(X)}$

5 Joint Probability Distribution

When analyzing two or more random variables simultaneously, we study their *joint probability distribution*, which describes the probability that multiple events occur at the same time. Given a joint distribution and one marginal distribution, we can solve for the remaining marginal distribution of a single random variable.

5.1 Discrete Random Variables: Joint Probability Mass Function (PMF)

Definition of Joint PMF

Let X, Y be discrete random variables with support values $x_1, x_2, \dots, x_s$ and $y_1, y_2, \dots, y_t$ respectively. The **joint probability mass function** is defined:

$$p_{ij} = p(x_i, y_j) = \mathbb{P}(X = x_i,\ Y = y_j), \quad x_i, y_j \in \mathbb{R},\ i = 1, \dots, s,\ j = 1, \dots, t$$

5.2 Joint Probability Table Layout

X	Y	y_1 p_{i1}	y_2 p_{i2}	$\cdots$ $\cdots$	**Row Sum (Marginal for X)** $\mathbb{P}(X = x_i)$
x_1		p_{11}	p_{12}	$\cdots$	$\mathbb{P}(X = x_1)$
x_2		p_{21}	p_{22}	$\cdots$	$\mathbb{P}(X = x_2)$
$\vdots$		$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_s		p_{s1}	p_{s2}	$\cdots$	$\mathbb{P}(X = x_s)$
Column Sum (Marginal for Y)		$\mathbb{P}(Y = y_1)$	$\mathbb{P}(Y = y_2)$	$\cdots$	**1**

5.3 Marginal PMF Formulas

From the joint PMF, marginal distributions for single variables are obtained by summing out the other variable:

$$\mathbb{P}(X = x_i) = \sum_{j=1}^{t} \mathbb{P}(X = x_i, Y = y_j) = \sum_{j=1}^{t} p_{ij}$$

$$\mathbb{P}(Y = y_j) = \sum_{i=1}^{s} \mathbb{P}(X = x_i, Y = y_j) = \sum_{i=1}^{s} p_{ij}$$

The marginal distribution is the univariate probability distribution of one variable alone, extracted from the joint table.

5.4 Properties of Discrete Joint PMF

1. Non-negativity: $p(x_i, y_j) \geq 0$ for all valid x_i, y_j.
2. Total probability sums to 1: $\sum_{i=1}^{s} \sum_{j=1}^{t} p(x_i, y_j) = 1$.
3. Rectangle probability: $\mathbb{P}(a < X < b,\ c < Y < d) = \sum_{\substack{a<x_i<b \\ c<y_j<d}} p(x_i, y_j)$.

5.5 Example 3: Hypergeometric Draw of Colored Balls

Problem Statement

Pocket contents:

- 3 blue balls, 2 red balls, 3 green balls; total $3 + 2 + 3 = 8$ balls. Draw 2 balls without replacement.

Define:

- X = number of blue balls drawn

- Y = number of red balls drawn

Tasks:

(1) Derive formula for joint PMF $p(x, y) = \mathbb{P}(X = x, Y = y)$.

(2) Construct full joint probability table.

(3) Compute $\mathbb{P}(X+Y \leq 1)$.

(4) Find marginal distribution of X.

(5) Find marginal distribution of Y.

1: Joint PMF Formula (Hypergeometric) Total ways to draw 2 balls: $\binom{8}{2}$. For x blue, y red, the remaining $2-x-y$ balls are green. Valid ranges: $x = 0, 1, 2$; $y = 0, 1, 2$; $x + y \leq 2$.

$$p(x, y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}}$$

Compute denominator first:

$$\binom{8}{2} = \frac{8!}{2!6!} = \frac{8 \cdot 7}{2} = 28$$

Individual Joint Probability Calculations

$$p(0,0) = \frac{\binom{3}{0}\binom{2}{0}\binom{3}{2}}{28} = \frac{1 \cdot 1 \cdot 3}{28} = \frac{3}{28}$$
$$p(0,1) = \frac{\binom{3}{0}\binom{2}{1}\binom{3}{1}}{28} = \frac{1 \cdot 2 \cdot 3}{28} = \frac{6}{28} = \frac{3}{14}$$
$$p(0,2) = \frac{\binom{3}{0}\binom{2}{2}\binom{3}{0}}{28} = \frac{1 \cdot 1 \cdot 1}{28} = \frac{1}{28}$$
$$p(1,0) = \frac{\binom{3}{1}\binom{2}{0}\binom{3}{1}}{28} = \frac{3 \cdot 1 \cdot 3}{28} = \frac{9}{28}$$
$$p(1,1) = \frac{\binom{3}{1}\binom{2}{1}\binom{3}{0}}{28} = \frac{3 \cdot 2 \cdot 1}{28} = \frac{6}{28} = \frac{3}{14}$$
$$p(1,2) = \binom{3}{1}\binom{2}{2}\binom{3}{-1} \quad \text{(invalid, zero)} = 0$$
$$p(2,0) = \frac{\binom{3}{2}\binom{2}{0}\binom{3}{0}}{28} = \frac{3 \cdot 1 \cdot 1}{28} = \frac{3}{28}$$
$$p(2,1) = p(2,2) = 0 \quad \text{(insufficient balls, negative green count)}$$

2: Joint Probability Table

$X \setminus Y$	$y=0$	$y=1$	$y=2$	Row Sum $\mathbb{P}(X=x)$
$x=0$	$\frac{3}{28}$	$\frac{3}{14}$	$\frac{1}{28}$	$\frac{3}{28}+\frac{6}{28}+\frac{1}{28}=\frac{10}{28}=\frac{5}{14}$
$x=1$	$\frac{9}{28}$	$\frac{3}{14}$	0	$\frac{9}{28}+\frac{6}{28}+0=\frac{15}{28}$
$x=2$	$\frac{3}{28}$	0	0	$\frac{3}{28}+0+0=\frac{3}{28}$
Column Sum $\mathbb{P}(Y=y)$	$\frac{3+9+3}{28}=\frac{15}{28}$	$\frac{6+6+0}{28}=\frac{12}{28}=\frac{3}{7}$	$\frac{1+0+0}{28}=\frac{1}{28}$	$\frac{28}{28}=1$

3: Compute $\mathbb{P}(X+Y\le 1)$ Valid (x,y) pairs satisfying $x+y\le 1$: $(0,0),\ (0,1),\ (1,0)$

$$\begin{aligned}\mathbb{P}(X+Y\le 1) &= p(0,0)+p(0,1)+p(1,0)\\ &= \frac{3}{28}+\frac{3}{14}+\frac{9}{28}\\ &= \frac{3}{28}+\frac{6}{28}+\frac{9}{28}\\ &= \frac{18}{28}=\frac{9}{14}\end{aligned}$$

4: Marginal Distribution of X

$$\begin{aligned}\mathbb{P}(X=0) &= p(0,0)+p(0,1)+p(0,2)=\frac{3}{28}+\frac{6}{28}+\frac{1}{28}=\frac{10}{28}=\frac{5}{14}\\ \mathbb{P}(X=1) &= p(1,0)+p(1,1)+p(1,2)=\frac{9}{28}+\frac{6}{28}+0=\frac{15}{28}\\ \mathbb{P}(X=2) &= p(2,0)+p(2,1)+p(2,2)=\frac{3}{28}+0+0=\frac{3}{28}\end{aligned}$$

x	0	1	2	Sum
$\mathbb{P}(X=x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

5: Marginal Distribution of Y

$$\begin{aligned}\mathbb{P}(Y=0) &= p(0,0)+p(1,0)+p(2,0)=\frac{3}{28}+\frac{9}{28}+\frac{3}{28}=\frac{15}{28}\\ \mathbb{P}(Y=1) &= p(0,1)+p(1,1)+p(2,1)=\frac{6}{28}+\frac{6}{28}+0=\frac{12}{28}=\frac{3}{7}\\ \mathbb{P}(Y=2) &= p(0,2)+p(1,2)+p(2,2)=\frac{1}{28}+0+0=\frac{1}{28}\end{aligned}$$

y	0	1	2	Sum
$\mathbb{P}(Y=y)$	$\frac{15}{28}$	$\frac{3}{7}$	$\frac{1}{28}$	1

5.6 Continuous Random Variables: Joint Probability Density Function (PDF)

For continuous random variables X, Y, we use a **joint density function** $f(x,y)$ with double integrals for probability calculations.

1. Non-negativity: $f(x,y)\ge 0,\ \forall x,y\in\mathbb{R}$.
2. Total integral equals 1: $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy=1$.
3. Rectangle probability: $\mathbb{P}(a<X<b,\ c<Y<d)=\displaystyle\int_c^d\int_a^b f(x,y)\,dxdy$.
4. General region probability: For an area $A\subset\mathbb{R}^2$,

$$\mathbb{P}\big((X,Y)\in A\big)=\iint_A f(x,y)\,dxdy$$

5. Marginal PDF formulas:

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy,\qquad f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx$$

5.7 Example 4: Joint PDF $f(x, y) = x + y,\ 0 < x < 1,\ 0 < y < 1$

Problem Statement

The joint PDF of continuous random variables X, Y is

$$f(x,y) = \begin{cases} x + y & 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Calculate the marginal density functions $f_X(x)$ and $f_Y(y)$ over the support region $(0, 1) \times (0, 1)$.

Analytical Derivation for $f_X(x)$

By definition of marginal PDF:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

Outside $0 < y < 1$, $f(x, y) = 0$, so we restrict integration bounds to $y = 0$ to $y = 1$:

$$f_X(x) = \int_0^1 (x + y)\, dy, \quad 0 < x < 1$$

Integrate term-by-term with respect to y (treat x as constant):

$$\begin{aligned} \int_0^1 (x + y)\, dy &= \int_0^1 x\, dy + \int_0^1 y\, dy \\ &= xy\big|_0^1 + \frac{1}{2}y^2\bigg|_0^1 \\ &= x(1 - 0) + \left(\frac{1}{2}(1)^2 - \frac{1}{2}(0)^2\right) \\ &= x + \frac{1}{2} \end{aligned}$$

Final marginal for X:

$$f_X(x) = \begin{cases} x + \frac{1}{2} & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Analytical Derivation for $f_Y(y)$

Symmetric calculation integrating out x:

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_0^1 (x + y)\, dx, \quad 0 < y < 1$$

Integrate term-by-term with respect to x (treat y as constant):

$$\begin{aligned} \int_0^1 (x + y)\, dx &= \int_0^1 x\, dx + \int_0^1 y\, dx \\ &= \frac{1}{2}x^2\bigg|_0^1 + y\,(x)\,\big|_0^1 \\ &= \left(\frac{1}{2}(1)^2 - \frac{1}{2}(0)^2\right) + y(1 - 0) \\ &= \frac{1}{2} + y \end{aligned}$$

Final marginal for Y:

$$f_Y(y) = \begin{cases} y + \dfrac{1}{2} & 0 < y < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Consistency Check (Verify Marginal Integrals Equal 1)

Check $\displaystyle\int_0^1 f_X(x)dx = 1$:

$$\int_0^1 \left(x + \frac{1}{2}\right) dx = \left[\frac{1}{2}x^2 + \frac{1}{2}x\right]_0^1 = \left(\frac{1}{2} + \frac{1}{2}\right) - 0 = 1$$

Check $\displaystyle\int_0^1 f_Y(y)dy = 1$:

$$\int_0^1 \left(y + \frac{1}{2}\right) dy = \left[\frac{1}{2}y^2 + \frac{1}{2}y\right]_0^1 = \left(\frac{1}{2} + \frac{1}{2}\right) - 0 = 1$$

Both marginal PDFs satisfy the normalization rule for valid density functions.

6 Covariance Matrix

Matrix notation provides a compact way to quantify pairwise relationships between a collection of random variables. The **covariance matrix** assembles variances (self-spread) and covariances (cross-linear association) for every pair of random variables in the set.

Definition of Population Covariance Matrix

Let $\{X_1, X_2, \ldots, X_p\}$ be a set of p random variables. The covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{p\times p}$ is defined element-wise:

$$\mathbf{\Sigma}_{ij} = \begin{cases} \mathrm{Var}(X_i) & i = j \quad \text{(main diagonal: variance of } X_i\text{)} \\ \mathrm{Cov}(X_i, X_j) & i \neq j \quad \text{(off-diagonal: covariance between } X_i, X_j\text{)} \end{cases}$$

Matrix expansion:

$$\mathbf{\Sigma} = \begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) & \ldots & \mathrm{Cov}(X_1, X_p) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) & \ldots & \mathrm{Cov}(X_2, X_p) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_p, X_1) & \mathrm{Cov}(X_p, X_2) & \ldots & \mathrm{Var}(X_p) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \ldots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \ldots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \ldots & \sigma_p^2 \end{bmatrix}$$

Property: Covariance is symmetric $\mathrm{Cov}(X_i, X_j) = \mathrm{Cov}(X_j, X_i)$, so $\mathbf{\Sigma} = \mathbf{\Sigma}^T$ (symmetric square matrix). Simple interpretation: Diagonal entries measure individual variable spread; off-diagonals measure linear co-movement between two distinct variables.

6.1 Example 5: Sample Covariance Matrix Calculation

Given Sample Dataset

5 variables (A, B, C, D, E), $n = 6$ sample observations per variable:

Variable	Obs1	Obs2	Obs3	Obs4	Obs5	Obs6
A	1	2	3	4	5	6
B	2	3	5	6	1	9
C	3	5	5	5	10	8
D	10	20	30	40	50	55
E	7	8	9	4	6	10

Sample covariance formula (unbiased, denominator $n-1$):

$$\widehat{\mathrm{Cov}}(X_i, X_j) = \frac{1}{n-1}\sum_{k=1}^{n}\left(X_{i,k} - \bar{X}_i\right)\left(X_{j,k} - \bar{X}_j\right), \quad \widehat{\mathrm{Var}}(X_i) = \widehat{\mathrm{Cov}}(X_i, X_i)$$

$n = 6$, so $n - 1 = 5$.

1: Compute Sample Means for Each Variable

$$\begin{aligned}
\bar{A} &= \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5 \\
\bar{B} &= \frac{2+3+5+6+1+9}{6} = \frac{26}{6} \approx 4.3333333 \\
\bar{C} &= \frac{3+5+5+5+10+8}{6} = \frac{36}{6} = 6 \\
\bar{D} &= \frac{10+20+30+40+50+55}{6} = \frac{205}{6} \approx 34.1666667 \\
\bar{E} &= \frac{7+8+9+4+6+10}{6} = \frac{44}{6} \approx 7.3333333
\end{aligned}$$

2: Centered Data Matrix (Subtract Row Mean from Each Entry) Centered row = Original row − row mean:

$$\begin{aligned}
\tilde{A} &= [1-3.5,\ 2-3.5,\ 3-3.5,\ 4-3.5,\ 5-3.5,\ 6-3.5] = [-2.5,\ -1.5,\ -0.5,\ 0.5,\ 1.5,\ 2.5] \\
\tilde{B} &= \left[2-\tfrac{26}{6},\ 3-\tfrac{26}{6},\ 5-\tfrac{26}{6},\ 6-\tfrac{26}{6},\ 1-\tfrac{26}{6},\ 9-\tfrac{26}{6}\right] = \left[-\tfrac{14}{6},\ -\tfrac{8}{6},\ \tfrac{4}{6},\ \tfrac{10}{6},\ -\tfrac{20}{6},\ \tfrac{28}{6}\right] \\
\tilde{C} &= [3-6,\ 5-6,\ 5-6,\ 5-6,\ 10-6,\ 8-6] = [-3,\ -1,\ -1,\ -1,\ 4,\ 2] \\
\tilde{D} &= \left[10-\tfrac{205}{6},\ 20-\tfrac{205}{6},\ 30-\tfrac{205}{6},\ 40-\tfrac{205}{6},\ 50-\tfrac{205}{6},\ 55-\tfrac{205}{6}\right] = \left[-\tfrac{145}{6},\ -\tfrac{85}{6},\ -\tfrac{25}{6},\ \tfrac{35}{6},\ \tfrac{95}{6},\ \tfrac{125}{6}\right] \\
\tilde{E} &= \left[7-\tfrac{44}{6},\ 8-\tfrac{44}{6},\ 9-\tfrac{44}{6},\ 4-\tfrac{44}{6},\ 6-\tfrac{44}{6},\ 10-\tfrac{44}{6}\right] = \left[\tfrac{2}{6},\ \tfrac{8}{6},\ \tfrac{14}{6},\ -\tfrac{20}{6},\ -\tfrac{8}{6},\ \tfrac{16}{6}\right]
\end{aligned}$$

3: Sample Covariance Matrix Formula via Centered Matrix Let $\widetilde{\mathbf{X}}$ be the 5×6 centered data matrix. The unbiased sample covariance matrix:

$$\widehat{\mathbf{\Sigma}} = \frac{1}{n-1}\widetilde{\mathbf{X}}\,\widetilde{\mathbf{X}}^T$$

$n - 1 = 5$, so scale the outer product by $\frac{1}{5}$.

Verified Numerical Sample Covariance Matrix (Matching Sage Output)

$$\widehat{\boldsymbol{\Sigma}} = \begin{bmatrix} 3.5000000 & 3.0000000 & 0.4000000 & 32.5000000 & 0.4000000 \\ 3.0000000 & 8.6666667 & 0.4000000 & 25.3333333 & 2.4666667 \\ 0.4000000 & 0.4000000 & 38.0000000 & 0.4000000 & 0.4000000 \\ 32.5000000 & 25.3333333 & 0.4000000 & 304.1666667 & 1.3333333 \\ 0.4000000 & 2.4666667 & 0.4000000 & 1.3333333 & 4.6666667 \end{bmatrix}$$