Probability and Statistics: MA6.101

Tutorial 9

Topics Covered: CLT, Random Vectors

Q1: Forty nine measurements are recorded to several decimal places. Each of these 49 numbers is rounded off to the nearest integer. The sum of the original 49 numbers is approximated by the sum of those integers. Assume that the errors made in rounding off are independent, identically distributed random variables with a uniform distribution over the interval $(-0.5, 0.5)$. Compute approximately the probability that the sum of the integers is within two units of the true sum.

A: Let $X_1, X_2, ..., X_n$, where $n = 49$, denote the 49 measurements, and $Y_1, Y_2, ..., Y_n$ be the corresponding nearest integers after rounding off. Then

$$X_i = Y_i + U_i \quad \text{for } i = 1, 2, ..., n,$$

where $U_1, U_2, ..., U_n$ are independent identically distributed uniform random variables on the interval $(-0.5, 0.5)$. We then have that for all i, $E(U_i) = 0 \quad$ for all i, and $\text{Var}(U_i) = \sigma^2 = \frac{1}{12}$.
The sum, S, of the original measurements is

$$S = \sum_{i=1}^{n} Y_i + \sum_{i=1}^{n} U_i,$$

where $\sum_{i=1}^{n} Y_i$ is the sum of the integer approximations. Let the difference

$$W = S - \sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} U_i.$$

We would like to estimate $\Pr(|W| \leq 2)$. We will do this by applying the Central Limit Theorem to $U_1, U_2, U_3, \ldots$.
Observe that $W = n\overline{U}_n$, where $\overline{U}_n$ is the sample mean of the rounding off values $U_1, U_2, \ldots, U_n$. By the Central Limit Theorem

$$\Pr\left(\frac{\overline{U}_n}{\sigma/\sqrt{n}} \leq z\right) \approx \Pr(Z \leq z),$$

for all $z \in \mathbb{R}$, where $Z \sim \text{Normal}(0, 1)$. We therefore get that, for $z > 0$,

$$\Pr\left(\frac{|\overline{U}_n|}{\sigma/\sqrt{n}} \leq z\right) \approx \Pr(|Z| \leq z),$$

It then follows that

$$
\begin{aligned}
\Pr(|W| \leq 2) &= \Pr(|\overline{U}_n| \leq 2/n) \\
&= \Pr\left(\frac{|\overline{U}_n|}{\sigma/\sqrt{n}} \leq \frac{2}{\sigma\sqrt{n}}\right) \approx \Pr\left(|Z| \leq \frac{2}{\sigma\sqrt{n}}\right) \\
&= 2F_Z\left(\frac{2}{\sigma\sqrt{n}}\right) - 1 = 2F_Z\left(\frac{4\sqrt{3}}{\sqrt{49}}\right) - 1 \\
&\approx 2F_Z(0.99) - 1 \approx 2(0.8389) - 1 \approx 0.6778,
\end{aligned}
$$

Q2: Let X_1 be a uniform random variable with support $R_{X_1} = [1, 2]$ and probability density function

$$
f_{X_1}(x_1) = \begin{cases} 1 & \text{if } x_1 \in R_{X_1} \\ 0 & \text{if } x_1 \notin R_{X_1} \end{cases}
$$

Let X_2 be a continuous random variable, independent of X_1, with support $R_{X_2} = [0, 2]$ and probability density function

$$
f_{X_2}(x_2) = \begin{cases} \frac{3}{8}x_2^2 & \text{if } x_2 \in R_{X_2} \\ 0 & \text{if } x_2 \notin R_{X_2} \end{cases}
$$

Let

$$
\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} X_1^2 \\ X_1 + X_2 \end{bmatrix}
$$

Find the joint probability density function of the random vector $\mathbf{Y}$.

A:
Since X_1 and X_2 are independent, their joint probability density function is equal to the product of their marginal density functions:

$$
f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2) = \begin{cases} \frac{3}{8}x_2^2 & \text{if } x_1 \in [1, 2] \text{ and } x_2 \in [0, 2] \\ 0 & \text{otherwise} \end{cases}
$$

The support of Y_1 is $R_{Y_1} = [1, 4]$ and the support of Y_2 is $R_{Y_2} = [1, 4]$. The support of $\mathbf{Y}$ is

$$
R_{\mathbf{Y}} = \{(y_1, y_2) : y_1 \in [1, 4], y_2 \in [\sqrt{y_1}, 4]\}
$$

The function $y = g(x)$ is one-to-one and its inverse $g^{-1}(y)$ is defined by

$$
x_1 = \sqrt{y_1} \qquad x_2 = y_2 - \sqrt{y_1}
$$

with Jacobian matrix

$$
J_{g^{-1}}(y) = \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}y_1^{-\frac{1}{2}} & 0 \\ -\frac{1}{2}y_1^{-\frac{1}{2}} & 1 \end{bmatrix}
$$

The determinant of the Jacobian matrix is

$$\det(J_{g^{-1}}(y)) = \frac{1}{2\sqrt{y_1}} \cdot 1 - 0 \cdot (-\frac{1}{2\sqrt{y_1}}) = \frac{1}{2\sqrt{y_1}}$$

which is greater than zero for any y belonging to R_Y. The formula for the joint probability density function of $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(y) = \begin{cases} f_{\mathbf{X}}(g^{-1}(y))|\det(J_{g^{-1}}(y))| & \text{if } y \in R_Y \\ 0 & \text{if } y \notin R_Y \end{cases}$$

and

$$\begin{aligned} f_{\mathbf{X}}(g^{-1}(y))|\det(J_{g^{-1}}(y))| &= f_{X_1}(x_1) f_{X_2}(x_2) \cdot \frac{1}{2\sqrt{y_1}} \\ &= \frac{3}{8} x_2^2 \cdot \frac{1}{2\sqrt{y_1}} = \frac{3}{16}(y_2 - \sqrt{y_1})^2 \frac{1}{\sqrt{y_1}} \end{aligned}$$

$$= \begin{cases} \frac{3}{16}(y_2 - \sqrt{y_1})^2 \frac{1}{\sqrt{y_1}} & \text{if } y_1 \in [1,4] \text{ and } y_2 \in [\sqrt{y_1}, 4] \\ 0 & \text{otherwise} \end{cases}$$

.

Q3: Find the Kullback–Leibler divergence between two multivariate gaussian distributions, $P \sim N(\mu_1, \Sigma_1)$ and $Q \sim N(\mu_2, \Sigma_2)$, both n dimensional. KL divergence between two distributions P and Q of a continuous random variable is given by

$$\mathrm{KL}[P\,||\,Q] = \int_{\mathcal{X}} p(x)\, \ln \frac{p(x)}{q(x)}\, \mathrm{d}x$$

A:

$$\begin{aligned} \mathrm{KL}[P\,||\,Q] &= \int_{\mathbb{R}^n} \mathcal{N}(x; \mu_1, \Sigma_1)\, \ln \frac{\mathcal{N}(x; \mu_1, \Sigma_1)}{\mathcal{N}(x; \mu_2, \Sigma_2)}\, \mathrm{d}x \\ &= \left\langle \ln \frac{\mathcal{N}(x; \mu_1, \Sigma_1)}{\mathcal{N}(x; \mu_2, \Sigma_2)} \right\rangle_{p(x)} . \end{aligned}$$

Substituting the multivariate gaussian pdf

$$\begin{aligned} \mathrm{KL}[P\,||\,Q] &= \left\langle \ln \frac{\frac{1}{\sqrt{(2\pi)^n|\Sigma_1|}} \cdot \exp\left[-\frac{1}{2}(x-\mu_1)^{\mathrm{T}}\Sigma_1^{-1}(x-\mu_1)\right]}{\frac{1}{\sqrt{(2\pi)^n|\Sigma_2|}} \cdot \exp\left[-\frac{1}{2}(x-\mu_2)^{\mathrm{T}}\Sigma_2^{-1}(x-\mu_2)\right]} \right\rangle_{p(x)} \\ &= \left\langle \frac{1}{2}\ln\frac{|\Sigma_2|}{|\Sigma_1|} - \frac{1}{2}(x-\mu_1)^{\mathrm{T}}\Sigma_1^{-1}(x-\mu_1) + \frac{1}{2}(x-\mu_2)^{\mathrm{T}}\Sigma_2^{-1}(x-\mu_2) \right\rangle_{p(x)} \\ &= \frac{1}{2}\left\langle \ln\frac{|\Sigma_2|}{|\Sigma_1|} - (x-\mu_1)^{\mathrm{T}}\Sigma_1^{-1}(x-\mu_1) + (x-\mu_2)^{\mathrm{T}}\Sigma_2^{-1}(x-\mu_2) \right\rangle_{p(x)} \end{aligned}$$

$$(x-\mu_1)^{\mathrm{T}}\Sigma_1^{-1}(x-\mu_1) = \mathrm{tr}\left[(x-\mu_1)\Sigma_1^{-1}(x-\mu_1)^{\mathrm{T}}\right] \text{ because it is a scalar}$$

$$\begin{aligned}
\mathrm{KL}[P\,||\,Q] &= \frac{1}{2}\left\langle \ln\frac{|\Sigma_2|}{|\Sigma_1|} - \mathrm{tr}\left[\Sigma_1^{-1}(x-\mu_1)(x-\mu_1)^{\mathrm{T}}\right] + \mathrm{tr}\left[\Sigma_2^{-1}(x-\mu_2)(x-\mu_2)^{\mathrm{T}}\right]\right\rangle_{p(x)} \\
&= \frac{1}{2}\left\langle \ln\frac{|\Sigma_2|}{|\Sigma_1|} - \mathrm{tr}\left[\Sigma_1^{-1}(x-\mu_1)(x-\mu_1)^{\mathrm{T}}\right] + \mathrm{tr}\left[\Sigma_2^{-1}\left(xx^{\mathrm{T}} - 2\mu_2x^{\mathrm{T}} + \mu_2\mu_2^{\mathrm{T}}\right)\right]\right\rangle_{p(x)} \\
&= \frac{1}{2}\Bigg(\ln\frac{|\Sigma_2|}{|\Sigma_1|} - \mathrm{tr}\left[\Sigma_1^{-1}\left\langle(x-\mu_1)(x-\mu_1)^{\mathrm{T}}\right\rangle_{p(x)}\right] + \\
&\quad \mathrm{tr}\left[\Sigma_2^{-1}\left\langle xx^{\mathrm{T}} - 2\mu_2x^{\mathrm{T}} + \mu_2\mu_2^{\mathrm{T}}\right\rangle_{p(x)}\right]\Bigg) \\
&= \frac{1}{2}\Bigg(\ln\frac{|\Sigma_2|}{|\Sigma_1|} - \mathrm{tr}\left[\Sigma_1^{-1}\left\langle(x-\mu_1)(x-\mu_1)^{\mathrm{T}}\right\rangle_{p(x)}\right] + \\
&\quad \mathrm{tr}\left[\Sigma_2^{-1}\left(\left\langle xx^{\mathrm{T}}\right\rangle_{p(x)} - \left\langle 2\mu_2x^{\mathrm{T}}\right\rangle_{p(x)} + \left\langle\mu_2\mu_2^{\mathrm{T}}\right\rangle_{p(x)}\right)\right]\Bigg) .
\end{aligned}$$

Using the following property:

$$x \sim \mathcal{N}(\mu,\Sigma) \quad \Rightarrow \quad \left\langle(x-\mu)(x-\mu)^T\right\rangle = \Sigma$$

We can simplify our expression

$$\begin{aligned}
\mathrm{KL}[P\,||\,Q] &= \frac{1}{2}\left(\ln\frac{|\Sigma_2|}{|\Sigma_1|} - \mathrm{tr}\left[\Sigma_1^{-1}\Sigma_1\right] + \mathrm{tr}\left[\Sigma_2^{-1}\left(\Sigma_1 + \mu_1\mu_1^{\mathrm{T}} - 2\mu_2\mu_1^{\mathrm{T}} + \mu_2\mu_2^{\mathrm{T}}\right)\right]\right) \\
&= \frac{1}{2}\left(\ln\frac{|\Sigma_2|}{|\Sigma_1|} - \mathrm{tr}\left[I_n\right] + \mathrm{tr}\left[\Sigma_2^{-1}\Sigma_1\right] + \mathrm{tr}\left[\Sigma_2^{-1}\left(\mu_1\mu_1^{\mathrm{T}} - 2\mu_2\mu_1^{\mathrm{T}} + \mu_2\mu_2^{\mathrm{T}}\right)\right]\right) \\
&= \frac{1}{2}\left(\ln\frac{|\Sigma_2|}{|\Sigma_1|} - n + \mathrm{tr}\left[\Sigma_2^{-1}\Sigma_1\right] + \mathrm{tr}\left[\mu_1^{\mathrm{T}}\Sigma_2^{-1}\mu_1 - 2\mu_1^{\mathrm{T}}\Sigma_2^{-1}\mu_2 + \mu_2^{\mathrm{T}}\Sigma_2^{-1}\mu_2\right]\right) \\
&= \frac{1}{2}\left[\ln\frac{|\Sigma_2|}{|\Sigma_1|} - n + \mathrm{tr}\left[\Sigma_2^{-1}\Sigma_1\right] + (\mu_2-\mu_1)^{\mathrm{T}}\Sigma_2^{-1}(\mu_2-\mu_1)\right] .
\end{aligned}$$

Q4: Let $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ belong to a bivariate normal distribution $\mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$. Show that $x_1|x_2 \sim \mathcal{N}(\mu_{1|2}, \Sigma_{1|2})$ where

$$\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)$$
$$\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

A: The marginal distribution of x_2 is given by $\mathcal{N}(\mu_2, \Sigma_2)$. Using conditional probability

$$\begin{aligned}
p(x_1|x_2) &= \frac{\mathcal{N}(x;\mu,\Sigma)}{\mathcal{N}(x_2;\mu_2,\Sigma_{22})} \\
&= p(x_1|x_2) = \frac{\frac{1}{2\pi\sqrt{|\Sigma|}}\exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]}{\frac{1}{\sqrt{(2\pi)\Sigma_{22}}}\exp\left[-\frac{1}{2}\frac{(x_2-\mu_2)^2}{\Sigma_{22}}\right]} \\
&= \frac{1}{\sqrt{(2\pi)|\Sigma_{11}|}} \cdot \frac{\sqrt{|\Sigma_{22}|}}{\sqrt{|\Sigma|}}\exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) + \frac{1}{2}\frac{(x_2-\mu_2)^2}{\Sigma_{22}}\right] \\
&= \frac{1}{\sqrt{(2\pi)|\Sigma_{11}|}} \cdot \frac{\sqrt{|\Sigma_{22}|}}{\sqrt{|\Sigma|}}\exp\left[-\frac{1}{2}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}^T\Sigma^{-1}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix} + \frac{1}{2}\frac{(x_2-\mu_2)^2}{\Sigma_{22}}\right]
\end{aligned}$$

For a 2x2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ the inverse is given by

$$A^{-1} = \frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Using this and expanding

$$\begin{aligned}
p(x_1|x_2) &= \frac{1}{\sqrt{(2\pi)|\Sigma_{11}|}} \cdot \frac{\sqrt{|\Sigma_{22}|}}{\sqrt{|\Sigma|}} \exp\Bigg[-\frac{1}{2|\Sigma|}\big(\Sigma_{11}(x_2-\mu_2)^2 \\
&\quad + 2\Sigma_{12}(x_2-\mu_2)(x_1-\mu_1) + \Sigma_{22}(x_1-\mu_1)^2\big) + \frac{1}{2}\frac{(x_2-\mu_2)^2}{\Sigma_{22}}\Bigg] \\
&= \frac{1}{\sqrt{(2\pi)|\Sigma_{11}|}} \cdot \frac{\sqrt{|\Sigma_{22}|}}{\sqrt{|\Sigma|}} \exp\Bigg[-\frac{1}{2|\Sigma|\Sigma_{22}}\big(\Sigma_{12}^2(x_2-\mu_2)^2 \\
&\quad - 2\Sigma_{12}\Sigma_{22}(x_2-\mu_2)(x_1-\mu_1) + \Sigma_{22}^2(x_1-\mu_1)^2\big)\Bigg] \\
&= \frac{1}{\sqrt{(2\pi)|\Sigma_{11}|}} \cdot \frac{\sqrt{|\Sigma_{22}|}}{\sqrt{|\Sigma|}} \exp\Bigg[-\frac{1}{2|\Sigma|\Sigma_{22}}\big(\Sigma_{22}(x_1-\mu_1) - \Sigma_{12}(x_2-\mu_2)\big)^2\Bigg] \\
&= \frac{1}{\sqrt{(2\pi)|\Sigma_{11}|}} \cdot \frac{\sqrt{|\Sigma_{22}|}}{\sqrt{|\Sigma|}} \exp\Bigg[-\frac{\Sigma_{22}}{2|\Sigma|}\Big(x_1-\mu_1-\frac{\Sigma_{12}}{\Sigma_{22}}(x_2-\mu_2)\Big)^2\Bigg] \\
&= \frac{1}{\sqrt{(2\pi)|\Sigma_{11}|}} \cdot \frac{\sqrt{|\Sigma_{22}|}}{\sqrt{|\Sigma|}} \exp\Bigg[-\frac{1}{2}\Bigg(\frac{x_1-(\mu_1-\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2))}{\sqrt{\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}}}\Bigg)^2\Bigg]
\end{aligned}$$

Thus the mean and variance are

$$\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)$$

$$\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Q5: X and Y are said to be bivariate normal if $aX + bY$ is normal for all a and b. If X and Y are bivariate normal with 0 mean, variance of 1, and ρ correlation, then their joint pdf is:

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(x^2 - 2\rho xy + y^2\right)\right)$$

Find the joint pdf of $X + Y$ and $X - Y$.

A: Let $U = X + Y$ and $V = X - Y$.
Since X and Y are bivariate normal with means 0, variances 1, and correlation ρ, the joint distribution of (X, Y) is:

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(x^2 - 2\rho xy + y^2\right)\right)$$

Step 1: Define the Transformation We define the transformation:

$$U = X + Y \quad \text{and} \quad V = X - Y$$

Step 2: Find the Inverse Transformation Solving for X and Y in terms of U and V:

$$X = \frac{U+V}{2} \quad \text{and} \quad Y = \frac{U-V}{2}$$

Step 3: Calculate the Jacobian The Jacobian of the transformation from (X, Y) to (U, V) is:

$$J = \begin{vmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

Thus, $|J| = \frac{1}{2}$.

Step 4: Distribution of U and V Since X and Y are jointly normal with variances 1 and correlation ρ, we can derive the variances and covariance of U and V:

$$\mathrm{Var}(U) = \mathrm{Var}(X+Y) = \mathrm{Var}(X)+\mathrm{Var}(Y)+2\,\mathrm{Cov}(X,Y) = 1+1+2\rho = 2(1+\rho)$$

$$\mathrm{Var}(V) = \mathrm{Var}(X-Y) = \mathrm{Var}(X)+\mathrm{Var}(Y)-2\,\mathrm{Cov}(X,Y) = 1+1-2\rho = 2(1-\rho)$$

$$\mathrm{Cov}(U,V) = \mathrm{Cov}(X+Y, X-Y) = \mathrm{Var}(X) - \mathrm{Var}(Y) = 0$$

Therefore, U and V are independent with variances $2(1+\rho)$ and $2(1-\rho)$, respectively.

Step 5: Joint pdf of U and V Since U and V are independent, their joint pdf is the product of their marginal pdfs:

$$f_{U,V}(u,v) = f_U(u) f_V(v)$$

where

$$f_U(u) = \frac{1}{\sqrt{4\pi(1+\rho)}} \exp\left(-\frac{u^2}{4(1+\rho)}\right)$$

and

$$f_V(v) = \frac{1}{\sqrt{4\pi(1-\rho)}} \exp\left(-\frac{v^2}{4(1-\rho)}\right)$$

Thus, the joint pdf of $U = X+Y$ and $V = X-Y$ is:

$$f_{U,V}(u,v) = \frac{1}{4\pi\sqrt{(1+\rho)(1-\rho)}} \exp\left(-\frac{u^2}{4(1+\rho)} - \frac{v^2}{4(1-\rho)}\right)$$

Q6: Let $\mathbf{Z} = [Z_1, Z_2]^T$ be a normal random vector with the following mean and covariance matrices:

$$\mathbf{m} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix}.$$

Let also:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 4 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} W_1 \\ W_2 \\ W_3 \end{bmatrix} = \mathbf{AZ} + \mathbf{b}.$$

Answer the following:

1. Find $P(Z_2 > 1)$. 2. Find the expected value vector of $\mathbf{W}$, denoted as $\mathbf{m_W} = \mathbb{E}[\mathbf{W}]$. 3. Find the covariance matrix of $\mathbf{W}$, denoted as $\mathbf{C_W}$. 4. Find $P(W_3 \leq 4)$.

Solution

1. Finding $P(Z_2 > 1)$

We are given that Z_2 is part of a bivariate normal distribution. The marginal distribution of Z_2 is:

$$Z_2 \sim \mathcal{N}(5, 3)$$

Thus, we need to find $P(Z_2 > 1)$. This is equivalent to calculating:

$$P(Z_2 > 1) = 1 - P(Z_2 \leq 1)$$

We standardize Z_2 by using the formula:

$$P(Z_2 \leq 1) = P\left(\frac{Z_2 - 5}{\sqrt{3}} \leq \frac{1 - 5}{\sqrt{3}}\right) = P\left(Z' \leq \frac{-4}{\sqrt{3}}\right)$$

Using the standard normal table:

$$P\left(Z' \leq \frac{-4}{\sqrt{3}}\right) \approx P(Z' \leq -2.309) \approx 0.0105$$

Thus,

$$P(Z_2 > 1) = 1 - 0.0105 = 0.9895$$

2. Expected Value of W

The transformation of $\mathbf{Z}$ to $\mathbf{W}$ involves a linear transformation:

$$\mathbf{W} = \mathbf{AZ} + \mathbf{b}$$

The expected value of $\mathbf{W}$ is given by:

$$\mathbf{m_W} = \mathbb{E}[\mathbf{W}] = \mathbf{A}\mathbb{E}[\mathbf{Z}] + \mathbf{b} = \mathbf{Am} + \mathbf{b}$$

Substituting the values of $\mathbf{A}$, $\mathbf{m}$, and $\mathbf{b}$:

$$\mathbf{m_W} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$\mathbf{m_W} = \begin{bmatrix} 1 \cdot 3 + 0 \cdot 5 \\ 2 \cdot 3 + 3 \cdot 5 \\ 4 \cdot 3 + 1 \cdot 5 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 21 \\ 17 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$\mathbf{m_W} = \begin{bmatrix} 4 \\ 23 \\ 16 \end{bmatrix}$$

Thus, the expected value vector of $\mathbf{W}$ is:

$$\mathbf{m_W} = \begin{bmatrix} 4 \\ 23 \\ 16 \end{bmatrix}$$

3. Covariance Matrix of W

The covariance matrix of $\mathbf{W}$ is given by:

$$\mathbf{C_W} = \mathbf{A}\mathbf{C_Z}\mathbf{A}^T$$

where $\mathbf{C_Z}$ is the covariance matrix of $\mathbf{Z}$. Substituting the values of $\mathbf{A}$ and $\mathbf{C_Z}$:

$$\mathbf{C_W} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 1 \end{bmatrix}$$

First, compute $\mathbf{A}\mathbf{C_Z}$:

$$\mathbf{A}\mathbf{C_Z} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 16 & 13 \\ 22 & 11 \end{bmatrix}$$

Now compute the product with $\mathbf{A}^T$:

$$\mathbf{C_W} = \begin{bmatrix} 5 & 2 \\ 16 & 13 \\ 22 & 11 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 11 & 21 \\ 16 & 67 & 89 \\ 22 & 85 & 129 \end{bmatrix}$$

Thus, the covariance matrix of $\mathbf{W}$ is:

$$\mathbf{C_W} = \begin{bmatrix} 5 & 11 & 21 \\ 16 & 67 & 89 \\ 22 & 85 & 129 \end{bmatrix}$$

4. Finding $P(W_3 \leq 4)$

Since W_3 is a linear combination of normal variables, it is normally distributed with mean $\mu_{W_3} = 16$ and variance $\sigma^2_{W_3} = 129$. We need to find $P(W_3 \leq 4)$.

Standardize W_3:

$$P(W_3 \leq 4) = P\left(\frac{W_3 - 16}{\sqrt{129}} \leq \frac{4 - 16}{\sqrt{129}}\right) = P\left(Z \leq \frac{-12}{\sqrt{129}}\right)$$

Using the standard normal table:

$$P\left(Z \leq -1.056\right) \approx 0.1451$$

Thus, $P(W_3 \leq 4) \approx 0.1451$.