STAT 420

1. A large class took two exams. Suppose the exam scores X (Exam 1) and Y (Exam 2) follow a bivariate normal distribution with

$\mu_1 = 70$, $\sigma_1 = 10$,

$\mu_2 = 60$, $\sigma_2 = 15$, $\rho = 0.6$.

a) A students is selected at random. What is the probability that his/her score on Exam 2 is over 75?

$$P(Y > 75) = P\left(Z > \frac{75-60}{15}\right) = P(Z > 1.00) = \mathbf{0.1587}.$$

b) Suppose you're told that a student got a 80 on Exam 1. What is the probability that his/her score on Exam 2 is over 75?

Given $X = 80$, Y has Normal distribution

with mean $60 + 0.6 \cdot \frac{15}{10} \cdot (80 - 70) = 69$

and variance $\left(1 - 0.6^2\right) \cdot 15^2 = 144$ (standard deviation 12).

$$P(Y > 75 \mid X = 80) = P\left(Z > \frac{75-69}{12}\right) = P(Z > 0.50) = \mathbf{0.3085}.$$

c) Suppose you're told that a student got a 66 on Exam 1. What is the probability that his/her score on Exam 2 is over 75?

Given $X = 66$, Y has Normal distribution

with mean $60 + 0.6 \cdot \frac{15}{10} \cdot (66 - 70) = 56.4$

and variance $\left(1 - 0.6^2\right) \cdot 15^2 = 144$ (standard deviation 12).

$$P(Y > 75 \mid X = 66) = P\left(Z > \frac{75-56.4}{12}\right) = P(Z > 1.55) = \mathbf{0.0606}.$$

d) Suppose you're told that a student got a 70 on Exam 2. What is the probability that his/her score on Exam 1 is over 80?

Given Y = 70, X has Normal distribution

with mean $70 + 0.6 \cdot \frac{10}{15} \cdot (70 - 60) = 74$

and variance $\left(1 - 0.6^2\right) \cdot 10^2 = 64$ (standard deviation 8).

$P(X > 80 \mid Y = 70) = P(Z > \frac{80 - 74}{8}) = P(Z > 0.75) = \mathbf{0.2266}$.

e) A students is selected at random. What is the probability that the sum of his/her Exam 1 and Exam 2 scores is over 150?

X + Y has Normal distribution,

$E(X + Y) = \mu_X + \mu_Y = 70 + 60 = 130$,

$Var(X + Y) = \sigma_X^2 + 2\sigma_{XY} + \sigma_Y^2 = \sigma_X^2 + 2\rho\sigma_X\sigma_Y + \sigma_Y^2$

$= 10^2 + 2 \cdot 0.6 \cdot 10 \cdot 15 + 15^2 = 505$ (standard deviation $\approx$ 22.4722).

$P(X + Y > 150) = P(Z > \frac{150 - 130}{22.4722}) = P(Z > 0.89) = \mathbf{0.1867}$.

f) What proportion of students did better on Exam 1 than on Exam 2?

Want $P(X > Y) = P(X - Y > 0) = ?$

X – Y has Normal distribution,

$E(X - Y) = \mu_X - \mu_Y = 70 - 60 = 10$,

$Var(X - Y) = \sigma_X^2 - 2\sigma_{XY} + \sigma_Y^2 = \sigma_X^2 - 2\rho\sigma_X\sigma_Y + \sigma_Y^2$

$= 10^2 - 2 \cdot 0.6 \cdot 10 \cdot 15 + 15^2 = 145$ (standard deviation $\approx$ 12.0416).

$P(X - Y > 0) = P(Z > \frac{0 - 10}{12.0416}) = P(Z > -0.83) = \mathbf{0.7967}$.

g) Find $P(2X + 3Y > 350)$.

$2X + 3Y$ has Normal distribution,

$E(2X + 3Y) = 2\mu_X + 3\mu_Y = 2 \times 70 + 3 \times 60 = 320$,

$\text{Var}(2X + 3Y) = 4\sigma_X^2 + 12\sigma_{XY} + 9\sigma_Y^2 = 4\sigma_X^2 + 12\rho\sigma_X\sigma_Y + 9\sigma_Y^2$

$= 4 \times 10^2 + 12 \cdot 0.6 \cdot 10 \cdot 15 + 9 \times 15^2 = 3505$

(standard deviation ≈ 59.203).

$$P(2X + 3Y > 350) = P\left(Z > \frac{350 - 320}{59.203}\right) = P(Z > 0.5067) \approx \mathbf{0.3050}.$$

h) Find $P(5X + 3Y < 570)$.

$5X + 3Y$ has Normal distribution,

$E(5X + 3Y) = 5\mu_X + 3\mu_Y = 5 \times 70 + 3 \times 60 = 530$,

$\text{Var}(5X + 3Y) = 25\sigma_X^2 + 30\sigma_{XY} + 9\sigma_Y^2 = 25\sigma_X^2 + 30\rho\sigma_X\sigma_Y + 9\sigma_Y^2$

$= 25 \times 10^2 + 30 \cdot 0.6 \cdot 10 \cdot 15 + 9 \times 15^2 = 7225$

(standard deviation $= 85$).

$$P(5X + 3Y < 570) = P\left(Z < \frac{570 - 530}{85}\right) = P(Z < 0.47) = \mathbf{0.6808}.$$

i) Find $P(5X - 4Y > 150)$.

$5X - 4Y$ has Normal distribution,

$E(5X - 4Y) = 5\mu_X - 4\mu_Y = 5 \times 70 - 4 \times 60 = 110$,

$\text{Var}(5X - 4Y) = 25\sigma_X^2 - 40\sigma_{XY} + 16\sigma_Y^2 = 25\sigma_X^2 - 40\rho\sigma_X\sigma_Y + 16\sigma_Y^2$

$= 25 \times 10^2 - 40 \cdot 0.6 \cdot 10 \cdot 15 + 16 \times 15^2 = 2500$

(standard deviation $= 50$).

$$P(5X - 4Y > 150) = P\left(Z > \frac{150 - 110}{50}\right) = P(Z > 0.80) = \mathbf{0.2119}.$$

2. Suppose that company A and company B are in the same industry sector, and the prices of their stocks, \$X per share for company A and \$Y per share for company B, vary from day to day randomly according to a bivariate normal distribution with parameters $\mu_X = 45$, $\sigma_X = 5.6$, $\mu_Y = 25$, $\sigma_Y = 5$, $\rho = 0.8$.

a) What is the probability that on a given day the price of stock for company B (Y) exceeds \$33?

Y has Normal distribution with mean $\mu_Y = 25$

and standard deviation $\sigma_Y = 5$.

$$P(Y > 33) = P\left(Z > \frac{33-25}{5}\right) = P(Z > 1.60)$$

$$= 1 - \Phi(1.60) = 1 - 0.9452 = \mathbf{0.0548}.$$

b) Suppose that on a given day the price of stock for company A (X) is \$52. What is the probability that the price of stock for company B (Y) exceeds \$33?

Given $X = 52$, Y has Normal distribution

with mean $\mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) = 25 + 0.8 \cdot \frac{5}{5.6} \cdot (52-45) = 30$

and variance $\left(1-\rho^2\right) \cdot \sigma_Y^2 = \left(1-0.8^2\right) \cdot 5^2 = 9$

(standard deviation = 3).

$$P(Y > 33 \mid X = 52) = P\left(Z > \frac{33-30}{3}\right) = P(Z > 1.00) = 1 - \Phi(1.00)$$

$$= 1 - \Phi(1.00) = 1 - 0.8413 = \mathbf{0.1587}.$$

c) Alex bought 5 shares of company A stock and 3 shares of company B stock. What is the probability that on a given day the value of his portfolio $(5 X + 3 Y)$ is below \$250?

Portfolio = $5 X + 3 Y$.

Portfolio has Normal distribution

with mean $5 \mu_X + 3 \mu_Y = 5 \cdot 45 + 3 \cdot 25 = 300$

and variance

$$\mathrm{Var}(5 X + 3 Y) = \mathrm{Cov}(5 X + 3 Y, 5 X + 3 Y)$$

$$= \mathrm{Cov}(5 X, 5 X) + \mathrm{Cov}(5 X, 3 Y) + \mathrm{Cov}(3 Y, 5 X) + \mathrm{Cov}(3 Y, 3 Y)$$

$$= 25 \sigma_X^2 + 30 \sigma_{XY} + 9 \sigma_Y^2 = 25 \sigma_X^2 + 30 \rho \sigma_X \sigma_Y + 9 \sigma_Y^2$$

$$= 25 \cdot 5.6^2 + 30 \cdot 0.8 \cdot 5.6 \cdot 5 + 9 \cdot 5^2 = 1681$$

(standard deviation = $\sqrt{1681} = 41$).

$$P(\text{Portfolio} < 250) = P\left(Z < \frac{250 - 300}{41} \right) = P(Z < -1.22) = \Phi(-1.22)$$

$$= \mathbf{0.1112}.$$

d) What is the probability that 1 share of company A stock is worth more than 2 shares of company B stock?

Want $P(X > 2 Y) = P(X - 2 Y > 0) = ?$

$X - 2 Y$ has Normal distribution,

$E(X - 2 Y) = \mu_X - 2 \mu_Y = 45 - 2 \cdot 25 = -5$,

$$\mathrm{Var}(X - 2 Y) = \sigma_X^2 - 4 \sigma_{XY} + 4 \sigma_Y^2 = \sigma_X^2 - 4 \rho \sigma_X \sigma_Y + 4 \sigma_Y^2$$

$$= 5.6^2 - 4 \cdot 0.8 \cdot 5.6 \cdot 5 + 4 \cdot 5^2 = 41.76$$ (standard deviation ≈ 6.462).

$$P(X - 2 Y > 0) = P(Z > \frac{0 + 5}{6.462}) = P(Z > 0.77) = 1 - \Phi(0.77)$$

$$= 1 - 0.7794 = \mathbf{0.2206}.$$

3. In a college health fitness program, let X denote the weight in kilograms of a male freshman at the beginning of the program and let Y denote his weight change during a semester. Assume that X and Y have a bivariate normal distribution with $\mu_X = 75$, $\sigma_X = 9$, $\mu_Y = 2.5$, $\sigma_Y = 1.5$, $\rho = -0.6$. (The lighter students tend to gain weight, while the heavier students tend to lose weight.)

a) What proportion of the students that weigh 85 kg end up losing weight during the semester? That is, find $P(Y < 0 \mid X = 85)$.

Given $X = 85$, Y has Normal distribution

with mean $$\mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) = 2.5 + (-0.6) \cdot \frac{1.5}{9} \cdot (85 - 75) = 1.5$$

and variance $$\left(1 - \rho^2\right) \cdot \sigma_Y^2 = \left(1 - (-0.6)^2\right) \cdot 1.5^2 = 1.44$$

(standard deviation = 1.2).

$$P(Y < 0 \mid X = 85) = P\left(Z < \frac{0 - 1.5}{1.2}\right) = P(Z < -1.25) = \Phi(-1.25) = \mathbf{0.1056}.$$

b) What proportion of the students that weigh over 87 kg at the end of the semester? That is, find $P(X + Y > 87)$.

X + Y has Normal distribution,

$$E(X + Y) = \mu_X + \mu_Y = 75 + 2.5 = 77.5,$$

$$\mathrm{Var}(X + Y) = \sigma_X^2 + 2\sigma_{XY} + \sigma_Y^2 = \sigma_X^2 + 2\rho\sigma_X\sigma_Y + \sigma_Y^2$$

$$= 9^2 + 2 \cdot (-0.6) \cdot 9 \cdot 1.5 + 1.5^2 = 67.05.$$

$$SD(X + Y) \approx 8.1884.$$

$$P(X + Y > 87) = P\left(Z > \frac{87 - 77.5}{8.1884}\right) = P(Z > 1.16) = \mathbf{0.1230}.$$

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \cdots \\ X_n \end{pmatrix} \quad - \quad n\text{-dimensional random vector} \qquad \mathrm{E}(\mathbf{X}) = \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \cdots \\ \mu_n \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \cdots & \cdots & & \cdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}$$

$\sigma_{ij} = \mathrm{Cov}(X_i, X_j)$

$\boldsymbol{\Sigma} = \mathrm{E}[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})']$

$\boldsymbol{\Sigma}$ – symmetric, positive-definite

$\mathbf{Y} = \mathbf{A}\,\mathbf{X} + \mathbf{b} \qquad \mathbf{A} - m \times n \qquad \mathbf{b} \in \mathbf{R}^m$

$\Rightarrow \qquad \boldsymbol{\mu}_{\mathbf{Y}} = \mathrm{E}(\mathbf{Y}) = \mathbf{A}\,\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b}, \qquad \boldsymbol{\Sigma}_{\mathbf{Y}} = \mathbf{A}\,\boldsymbol{\Sigma}_{\mathbf{X}}\,\mathbf{A}'$

<u>Example</u>:

Consider a random vector $\vec{\mathbf{X}} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$ with mean $\mathrm{E}(\vec{\mathbf{X}}) = \vec{\boldsymbol{\mu}} = \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}$

and variance-covariance matrix $\mathrm{Cov}(\vec{\mathbf{X}}) = \boldsymbol{\Sigma} = \begin{pmatrix} 9 & 2 & -3 \\ 2 & 4 & -2 \\ -3 & -2 & 16 \end{pmatrix}$.

Then $\qquad \mathrm{Var}(X_1) = 9, \qquad \mathrm{Var}(X_2) = 4, \qquad \mathrm{Var}(X_3) = 16,$

$$\rho_{12} = \frac{2}{\sqrt{9}\cdot\sqrt{4}} = \frac{1}{3}, \qquad \rho_{13} = \frac{-3}{\sqrt{9}\cdot\sqrt{16}} = -\frac{1}{4}, \qquad \rho_{23} = \frac{-2}{\sqrt{4}\cdot\sqrt{16}} = -\frac{1}{4}.$$

Consider $2X_1 - 3X_2 - X_3 = \begin{pmatrix} 2 & -3 & -1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$.

$$\mathrm{E}(2X_1 - 3X_2 - X_3) = \begin{pmatrix} 2 & -3 & -1 \end{pmatrix} \vec{\boldsymbol{\mu}} = 2\mu_1 - 3\mu_2 - \mu_3 = 2\cdot 5 - 3\cdot 1 - 1\cdot 3 = 4.$$

$$\mathrm{Var}(2X_1 - 3X_2 - X_3) = \begin{pmatrix} 2 & -3 & -1 \end{pmatrix} \begin{pmatrix} 9 & 2 & -3 \\ 2 & 4 & -2 \\ -3 & -2 & 16 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ -1 \end{pmatrix}$$

$$= \begin{pmatrix} 15 & -6 & -16 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ -1 \end{pmatrix} = 64.$$

OR

$$\mathrm{Var}(2X_1 - 3X_2 - X_3) = 4\,\mathrm{Var}(X_1) + 9\,\mathrm{Var}(X_2) + \mathrm{Var}(X_3)$$
$$- 12\,\mathrm{Cov}(X_1, X_2) - 4\,\mathrm{Cov}(X_1, X_3) + 6\,\mathrm{Cov}(X_2, X_3)$$
$$= 36 + 36 + 16 - 24 + 12 - 12 = 64.$$

Multivariate Normal Distribution:

$\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ $\qquad f_{\mathbf{X}}(\mathbf{x}) = \dfrac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left\{ -\dfrac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$

$\mathbf{Z} \sim N_n(\mathbf{0}, \mathbf{I}_n)$ $\qquad f_{\mathbf{Z}}(\mathbf{z}) = \dfrac{1}{(2\pi)^{n/2}} \exp\left\{ -\dfrac{1}{2} \mathbf{z}' \mathbf{z} \right\} = \prod_{i=1}^{n} \dfrac{1}{\sqrt{2\pi}} \exp\left\{ -\dfrac{1}{2} z_i^2 \right\}$

$\mathbf{X} = \boldsymbol{\Sigma}^{1/2} \mathbf{Z} + \boldsymbol{\mu}$

$\mathrm{M}_{\mathbf{X}}(\mathbf{t}) = \exp\left\{ \mathbf{t}' \boldsymbol{\mu} + \dfrac{1}{2} \mathbf{t}' \boldsymbol{\Sigma} \mathbf{t} \right\}$ $\qquad \mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \\ \cdots \\ t_n \end{pmatrix} \in \mathbf{R}^n$

$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ $\qquad \mathbf{A} - m \times n$ $\qquad \mathbf{b} \in \mathbf{R}^m$

$\Rightarrow \qquad \mathbf{Y} \sim N_m(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \qquad \begin{matrix} \mathbf{X}_1 \text{ is of dimension } m < n \\ \mathbf{X}_2 \text{ is of dimension } n - m \end{matrix} \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$$

$$\mathbf{X}_1 \mid \mathbf{X}_2 \sim N_m\left(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\,\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\,\boldsymbol{\Sigma}_{22}^{-1}\,\boldsymbol{\Sigma}_{21}\right)$$

4*.

$$\mathbf{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \qquad \boldsymbol{\mu} = \begin{pmatrix} 5 \\ 3 \\ 7 \end{pmatrix} \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 9 \end{pmatrix}$$

a) Find $P(X_1 > 8)$.

$X_1 \sim N(5, 4)$

$$P(X_1 > 8) = P\left(Z > \frac{8-5}{2}\right) = P(Z > 1.5) = \mathbf{0.0668}.$$

b)* Find $P(X_1 > 8 \mid X_2 = 1, X_3 = 10)$.

$$\boldsymbol{\Sigma}_{22} = \begin{pmatrix} 4 & 2 \\ 2 & 9 \end{pmatrix} \qquad \boldsymbol{\Sigma}_{22}^{-1} = \frac{1}{32}\begin{pmatrix} 9 & -2 \\ -2 & 4 \end{pmatrix}$$

$$\boldsymbol{\Sigma}_{12}\,\boldsymbol{\Sigma}_{22}^{-1} = \frac{1}{32}\begin{pmatrix} -1 & 0 \end{pmatrix}\begin{pmatrix} 9 & -2 \\ -2 & 4 \end{pmatrix} = \frac{1}{32}\begin{pmatrix} -9 & 2 \end{pmatrix}$$

$$\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\,\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2) = 5 + \frac{1}{32}\begin{pmatrix} -9 & 2 \end{pmatrix}\begin{pmatrix} 1-3 \\ 10-7 \end{pmatrix} = 5.75.$$

$$\boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\,\boldsymbol{\Sigma}_{22}^{-1}\,\boldsymbol{\Sigma}_{21} = 4 - \frac{1}{32}\begin{pmatrix} -9 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \end{pmatrix} = 3.71875.$$

$X_1 \mid X_2 = 1, X_3 = 10 \sim N(5.75, 3.71875)$

$$P(X_1 > 8 \mid X_2 = 1, X_3 = 10) = P\left(Z > \frac{8-5.75}{\sqrt{3.71875}}\right) = P(Z > 1.17) = \mathbf{0.1210}.$$

c) Find $P(4X_1 - 3X_2 + 5X_3 < 63)$.

$$4\mu_1 - 3\mu_2 + 5\mu_3 = 4 \cdot 5 - 3 \cdot 3 + 5 \cdot 7 = 46.$$

$$\begin{pmatrix} 4 & -3 & 5 \end{pmatrix} \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 9 \end{pmatrix} \begin{pmatrix} 4 \\ -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 19 & -6 & 39 \end{pmatrix} \begin{pmatrix} 4 \\ -3 \\ 5 \end{pmatrix} = 289.$$

$$4X_1 - 3X_2 + 5X_3 \sim N(46, 289)$$

$$P(4X_1 - 3X_2 + 5X_3 < 63) = P\left(Z < \frac{63 - 46}{\sqrt{289}}\right)$$

$$= P(Z < 1.00) = \mathbf{0.8413}.$$