MVA HW#2

1. Problem 4.18 and 4.19 of the textbook, p. 205

4.18. Find the maximum likelihood estimates of the 2×1 mean vector $\boldsymbol{\mu}$ and the 2×2 covariance matrix $\boldsymbol{\Sigma}$ based on the random sample

$$\mathbf{X} = \begin{bmatrix} 3 & 6 \\ 4 & 4 \\ 5 & 7 \\ 4 & 7 \end{bmatrix}$$

from a bivariate normal population.

4.19. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{20}$ be a random sample of size $n = 20$ from an $N_6(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ population. Specify each of the following completely.

(a) The distribution of $(\mathbf{X}_1 - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}_1 - \boldsymbol{\mu})$

(b) The distributions of $\overline{\mathbf{X}}$ and $\sqrt{n}(\overline{\mathbf{X}} - \boldsymbol{\mu})$

(c) The distribution of $(n - 1)\,\mathbf{S}$

2. Problem 4.26 of the textbook, p. 206

4.26. Exercise 1.2 gives the age x_1, measured in years, as well as the selling price x_2, measured in thousands of dollars, for $n = 10$ used cars. These data are reproduced as follows:

x_1	1	2	3	3	4	5	6	8	9	11
x_2	18.95	19.00	17.95	15.54	14.00	12.95	8.94	7.49	6.00	3.99

(a) Use the results of Exercise 1.2 to calculate the squared statistical distances $(\mathbf{x}_j - \bar{\mathbf{x}})'\mathbf{S}^{-1}(\mathbf{x}_j - \bar{\mathbf{x}})$, $j = 1, 2, \ldots, 10$, where $\mathbf{x}_j' = [x_{j1}, x_{j2}]$.

(b) Using the distances in Part a, determine the proportion of the observations falling within the estimated 50% probability contour of a bivariate normal distribution.

(c) Order the distances in Part a and construct a chi-square plot.

(d) Given the results in Parts b and c, are these data approximately bivariate normal? Explain.

3. Problem 4.40 of the textbook, p. 208

4.40. Consider the data on national parks in Exercise 1.27.

(a) Comment on any possible outliers in a scatter plot of the original variables.

(b) Determine the power transformation $\hat{\lambda}_1$ the makes the x_1 values approximately normal. Construct a $Q-Q$ plot of the transformed observations.

(c) Determine the power transformation $\hat{\lambda}_2$ the makes the x_2 values approximately normal. Construct a $Q-Q$ plot of the transformed observations.

(d) Determine the power transformation for approximate bivariate normality using (4-40).

Table 1.11 Attendance and Size of National Parks

National Park	Size (acres)	Visitors (millions)
Arcadia	47.4	2.05
Bruce Canyon	35.8	1.02
Cuyahoga Valley	32.9	2.53
Everglades	1508.5	1.23
Grand Canyon	1217.4	4.40
Grand Teton	310.0	2.46
Great Smoky	521.8	9.19
Hot Springs	5.6	1.34
Olympic	922.7	3.14
Mount Rainier	235.6	1.17
Rocky Mountain	265.8	2.80
Shenandoah	199.0	1.09
Yellowstone	2219.8	2.84
Yosemite	761.3	3.30
Zion	146.6	2.59

4. Let $\mathbf{X}$ be a normally distributed random vector with

$$\boldsymbol{\mu} = \begin{bmatrix} -3 \\ 1 \\ -2 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

(a) Which of the following are pairs of independent random variables

i. X_1 and X_2

ii. X_1 and X_3

iii. X_1 and $X_1 + 3X_2 - 2X_3$

iv. (X_1, X_3) and X_2

v. $X_1 + X_3$ and $X_1 - 2X_2$

vi. $X_1 + X_2 + X_3$ and $4X_1 - 2X_2 + 3X_3$

(b) What is the distribution of $\mathbf{Y} = (X_1, X_2)'$

(c) What is the conditional distribution of $\mathbf{Y} = (X_1, X_2)'$ given $X_3 = x_3$?

(d) What is the conditional distribution of X_2 given $X_1 = x_1$ and $X_3 = x_3$

5. Suppose $\mathbf{X} \sim N_4(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\mu} = \begin{bmatrix} -4 \\ 3 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 4 & 0 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 2 & 1 & 9 & -2 \\ 2 & 0 & -2 & 4 \end{bmatrix}$$

(a) What is the distribution of $Z_1 = 3X_1 - 5X_2 + X_4$

(b) What is the joint distribution of Z_1 in part (a) and $Z_2 = 2X_1 - X_3 + 3X_4$

(c) Find the conditional distribution of $(X_1, X_4)'$ given $X_3 = 1$

6. Suppose $\mathbf{X} \sim N_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $\mathbf{Y} \sim N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ where $\mathbf{X}$ and $\mathbf{Y}$ are independent and

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 2 \\ 6 \end{bmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{bmatrix} 8 & -2 \\ -2 & 4 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$

(a) Evaluate $\mathrm{Cov}(\mathbf{X} - \mathbf{Y}, \mathbf{X} + \mathbf{Y})$

(b) Are $\mathbf{X} - \mathbf{Y}$ and $\mathbf{X} + \mathbf{Y}$ independent random vectors? Explain.

(c) Show that the joint distribution for the four dimensional random vector

$$\begin{bmatrix} \mathbf{X} - \mathbf{Y} \\ \mathbf{X} + \mathbf{Y} \end{bmatrix}$$

is a multivariate normal distribution