UNIVERSITY OF CALGARY & DEPARTMENT OF MATHEMATICS AND STATISTICS

STATISTICS 625 Assignment 2

Qian Wang

October 9, 2018

1 PROBLEM 1(EXERCISE 3.17)

Proof:
According to the hint, by independence, we have

$$\begin{aligned} &P(X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_p \leqslant x_p \; and \; Z_1 \leqslant z_1, Z_2 \leqslant z_2, \cdots, Z_q \leqslant z_q) \\ &= P(X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_p \leqslant x_p) \cdot P(Z_1 \leqslant z_1, Z_2 \leqslant z_2, \cdots, Z_q \leqslant z_q) \end{aligned}$$

Let $x_1, x_2, \cdots, x_p$ and $z_1, z_2, \cdots, z_q$ tend to infinity, then we have

$$\begin{aligned} &P(X_1 \leqslant x_1, X_2 \leqslant \infty, \cdots, X_p \leqslant \infty \; and \; Z_1 \leqslant z_1, Z_2 \leqslant \infty, \cdots, Z_q \leqslant \infty) \\ &= P(X_1 \leqslant x_1, X_2 \leqslant \infty, \cdots, X_p \leqslant \infty) \cdot P(Z_1 \leqslant z_1, Z_2 \leqslant \infty, \cdots, Z_q \leqslant \infty) \\ &\Leftrightarrow P(X_1 \leqslant x_1 \; and \; Z_1 \leqslant z_1) = P(X_1 \leqslant x_1) \cdot P(Z_1 \leqslant z_1) \end{aligned}$$

for all x_1, z_1. So X_1, Z_1 are independent. Similarly, we can repeat for other pairs. Let $x_1, \cdots x_{k-1}, x_{k+1}, \cdot x_p$ and $z_1, \cdots z_{k-1}, z_{k+1}, \cdot x_q$ tend to infinity. $(k = 1, 2, \cdots, max(p, q))$. We can obtain

$$\begin{aligned} &P(X_1 \leqslant \infty, \cdots, X_{k-1} \leqslant \infty, X_k \leqslant x_k \cdots, X_p \leqslant \infty \; and \; Z_1 \leqslant \infty, \cdots, Z_{k-1} \leqslant \infty, Z_k \leqslant z_k \cdots, Z_p \leqslant \infty) \\ &= P(X_1 \leqslant \infty, \cdots, X_{k-1} \leqslant \infty, X_k \leqslant x_k \cdots, X_p \leqslant \infty) \cdot P(Z_1 \leqslant \infty, \cdots, Z_{k-1} \leqslant \infty, Z_k \leqslant z_k \cdots, Z_p \leqslant \infty) \\ &\Leftrightarrow P(X_k \leqslant x_k \; and \; Z_k \leqslant z_k) = P(X_k \leqslant x_k) \cdot P(Z_k \leqslant z_k) \end{aligned}$$

for all x_k, z_k. Therefore, each component of $\boldsymbol{X}$ is independent of each component of $\boldsymbol{Z}$.

2 PROBLEM 2

(a)

$$(\boldsymbol{I} - \boldsymbol{H})^2 = \boldsymbol{I}^2 - 2\boldsymbol{I}\boldsymbol{H} + \boldsymbol{H}^2 = \boldsymbol{I} - 2\boldsymbol{H} + \boldsymbol{H} = \boldsymbol{I} - \boldsymbol{H}$$

$$(\boldsymbol{I}-\boldsymbol{H})^\top = \boldsymbol{I}^\top - \boldsymbol{H}^\top = \boldsymbol{I}-\boldsymbol{H}$$

So we have $\boldsymbol{I}-\boldsymbol{H} = (\boldsymbol{I}-\boldsymbol{H})^2 = (\boldsymbol{I}-\boldsymbol{H})^\top$,i.e., $\boldsymbol{I}-\boldsymbol{H}$ is idempotent. Then,

$$(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{H} = \boldsymbol{I}\boldsymbol{H} - \boldsymbol{H}^2 = \boldsymbol{H}-\boldsymbol{H} = \boldsymbol{0}$$
$$\boldsymbol{H}(\boldsymbol{I}-\boldsymbol{H}) = \boldsymbol{H}\boldsymbol{I} - \boldsymbol{H}^2 = \boldsymbol{H}-\boldsymbol{H} = \boldsymbol{0}$$

Therefore, $\boldsymbol{I}-\boldsymbol{H}$ and $\boldsymbol{H}$ are mutually orthogonal.

(b)

$$\begin{aligned}\boldsymbol{J}^2 = \boldsymbol{J}^\top\boldsymbol{J} &= (\frac{1}{n}\mathbb{1}\mathbb{1}^\top)^\top(\frac{1}{n}\mathbb{1}\mathbb{1}^\top) = (\frac{1}{n}\mathbb{1}\mathbb{1}^\top)(\frac{1}{n}\mathbb{1}\mathbb{1}^\top) = (\frac{1}{n^2}\mathbb{1}\mathbb{1}^\top\mathbb{1}\mathbb{1}^\top)\\ &= \frac{1}{n^2}\mathbb{1}(\mathbb{1}^\top\mathbb{1})\mathbb{1}^\top = \frac{1}{n^2}\mathbb{1}(n)\mathbb{1}^\top = \frac{1}{n}\mathbb{1}\mathbb{1}^\top = \boldsymbol{J}\end{aligned}$$

$$\boldsymbol{J}^\top = (\frac{1}{n}\mathbb{1}\mathbb{1}^\top)^\top = \frac{1}{n}\mathbb{1}\mathbb{1}^\top = \boldsymbol{J}$$

Therefore, $\boldsymbol{J}$ is idempotent.

(c) For an arbitary $\boldsymbol{A}$, which is idempotent, let λ be an eigenvalue of $\boldsymbol{A}$ and $\boldsymbol{x}$ be the eigenvector of λ. Then we have $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, $\boldsymbol{x} \neq \boldsymbol{0}$. From two perspective to deduce $\boldsymbol{A}^2\boldsymbol{x}$,

$$\boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \tag{2.1}$$

$$\boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}(\lambda\boldsymbol{x}) = \lambda(\boldsymbol{A}\boldsymbol{x}) = \lambda(\lambda\boldsymbol{x}) = \lambda^2\boldsymbol{x} \tag{2.2}$$

Since equation 2.1=2.2, then we have $\lambda\boldsymbol{x} = \lambda^2\boldsymbol{x}$, $\boldsymbol{x} \neq \boldsymbol{0} \Rightarrow \lambda^2\boldsymbol{x} - \lambda\boldsymbol{x} = 0 \Rightarrow \lambda^2 - \lambda = 0$. So we can get $\lambda_1 = 1$, $\lambda_2 = 0$.

(d) According to **Spectral Decomposition Theorem (SDT)**,

$$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{D}_{\boldsymbol{\lambda}}\boldsymbol{X}^\top \Rightarrow \boldsymbol{X}^{-\mathbf{1}}\boldsymbol{A} = \boldsymbol{D}_{\boldsymbol{\lambda}}\boldsymbol{X}^\top \Rightarrow \boldsymbol{D}_{\boldsymbol{\lambda}} = \boldsymbol{X}^{-\mathbf{1}}\boldsymbol{A}(\boldsymbol{X}^\top)^{-\mathbf{1}}$$

where $\boldsymbol{X} = (\boldsymbol{x_1}, ..., \boldsymbol{x_n})$ is the orthogonal matrix whose columns are the orthonormal eigenvectors of $\boldsymbol{A}$ and $\boldsymbol{D}_{\boldsymbol{\lambda}} = diag(\lambda_1, ..., \lambda_n)$ is the diagonal matrix of eigenvalues of $\boldsymbol{A}$. Then we can get,

$$\begin{aligned}\sum_{i=1}^{n}\lambda_i &= tr(\boldsymbol{D}_{\boldsymbol{\lambda}})\\ &= tr(\boldsymbol{X}^{-\mathbf{1}}\boldsymbol{A}(\boldsymbol{X}^\top)^{-\mathbf{1}})\\ &= tr(\boldsymbol{A}(\boldsymbol{X}^\top)^{-\mathbf{1}}\boldsymbol{X}^{-\mathbf{1}})\\ &= tr(\boldsymbol{A}(\boldsymbol{X}^{-\mathbf{1}})^\top\boldsymbol{X}^{-\mathbf{1}})\\ &= tr(\boldsymbol{A})\end{aligned}$$

Since for $\boldsymbol{A}$, $\lambda_i = 1$ or $\lambda_i = 0$. So $tr(\boldsymbol{A}) = the\ number\ of\ \{\lambda_i = 1\}$.

(e) According to **Spectral Decomposition Theorem (SDT)**,

$$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{D_\lambda}\boldsymbol{X}^\top$$

where $\boldsymbol{X} = (\boldsymbol{x_1}, ..., \boldsymbol{x_n})$ is the orthogonal matrix whose columns are the orthonormal eigenvectors of $\boldsymbol{A}$. Therefore, $|\boldsymbol{X}| = \pm 1 \neq 0$, and $|\boldsymbol{X}^\top| = \pm 1 \neq 0$, i.e., $\boldsymbol{X}$ and $\boldsymbol{X}^\top$ are non-singular. So, we have

$$rank(\boldsymbol{A}) = rank(\boldsymbol{X}\boldsymbol{D_\lambda}\boldsymbol{X}^\top) = rank(\boldsymbol{X}\boldsymbol{D_\lambda}) = rank(\boldsymbol{D_\lambda})$$

Since the hint told us for arbitrary matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, the rank of $\boldsymbol{AB}$ equals the rank of $\boldsymbol{A}$ if $\boldsymbol{B}$ is nonsingular. Then,

$$rank(\boldsymbol{D_\lambda}) = rank(diag(\lambda_1, ..., \lambda_n)) = the\ number\ of\ \{\lambda_i = 1\} = tr(\boldsymbol{A})$$

(f) We can easily find that $\exists \boldsymbol{b}^\top = (b_1, b_2, ..., b_n) = (1, -1, 0, ..., 0) \neq \boldsymbol{0}$ s.t., $\sum_{i=1}^n b_i \boldsymbol{J}_i = 0$, where $\boldsymbol{J_i}^\top = (\frac{1}{n}, ... \frac{1}{n})$. Therefore, $\boldsymbol{J_1}, ..., \boldsymbol{J_n}$ are not linear independent. Similarly, we can find *n-1*, *n-2*, ..., *2* of $\boldsymbol{J_i}$ are all not linear independent. Therefore $rank(\boldsymbol{J}) = 1$.

3 Problem 3

(a) First, we calculate the marginal PDFs of X_1 and X_2.

$$\begin{aligned}
f_{X1}(x_1) &= 2\int_0^\infty \phi_2(x_1, x_2; \boldsymbol{0}, \boldsymbol{I}_2) dx_2 I\{x_1 : x_1 \geqslant 0\} \\
&= 2\int_0^\infty \frac{1}{2\pi^{2/2}|\boldsymbol{I_2}|^{1/2}} exp\{-\frac{1}{2}(\boldsymbol{x}^\top \boldsymbol{I_2}^{-1}\boldsymbol{x})\} dx_2 I\{x_1 : x_1 \geqslant 0\} \\
&= 2\int_0^\infty \frac{1}{2\pi 1} exp\{-\frac{1}{2}(x_1^2 + x_2^2)\} dx_2 I\{x_1 : x_1 \geqslant 0\} \\
&= \frac{1}{\pi} exp\{-\frac{1}{2}x_1^2\} \int_0^\infty exp\{-\frac{1}{2}x_2^2\} dx_2 I\{x_1 : x_1 \geqslant 0\} \\
&= \frac{1}{\pi}\sqrt{2\pi} exp\{-\frac{1}{2}x_1^2\} \int_0^\infty \frac{1}{\sqrt{2\pi}} exp\{-\frac{1}{2}x_2^2\} dx_2 I\{x_1 : x_1 \geqslant 0\} \\
&= \frac{1}{\pi}\sqrt{2\pi} exp\{-\frac{1}{2}x_1^2\} \frac{1}{2} I\{x_1 : x_1 \geqslant 0\} \\
&= \frac{1}{\sqrt{2\pi}} exp\{-\frac{1}{2}x_1^2\} I\{x_1 : x_1 \geqslant 0\}
\end{aligned}$$

Similarly, we can get $f_{X_2}(x_2)$.

$$f_{X_2}(x_2) = \frac{1}{\sqrt{2\pi}} exp\{-\frac{1}{2}x_2^2\} I\{x_2 : x_2 \geqslant 0\}$$

(b) According to the marginal PDFs X_1 and X_2, we have that X_1 and X_2 are (marginally) normally distributed. $X_1 \sim N(0, 1)$ and $X_2 \sim N(0, 1)$.

4 Problem 4

Proof:
We can rewrite the np component vector.

$$[X_{11},\cdots X_{1p},X_{21},\cdots X_{2p},\cdots,X_{n1},\cdots,X_{np}] = [\boldsymbol{X_1}^\top,\cdots,\boldsymbol{X_n}^\top] = \underset{(1\times np)}{\boldsymbol{X}^\top}$$

And $\boldsymbol{X}^\top$ is multivariate normal with $\boldsymbol{X} \sim N_{np}(\boldsymbol{\mu},\boldsymbol{\Sigma_X})$, where

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu_1} \\ \boldsymbol{\mu_2} \\ \vdots \\ \boldsymbol{\mu_n} \end{pmatrix}, \quad \boldsymbol{\Sigma_X} = \begin{pmatrix} \boldsymbol{\Sigma} & 0 & \cdots & 0 \\ 0 & \boldsymbol{\Sigma} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{\Sigma} \end{pmatrix}$$

Let

$$\underset{(2p\times np)}{\boldsymbol{A}} = \begin{pmatrix} a_1\boldsymbol{I} & a_2\boldsymbol{I} & \cdots & a_n\boldsymbol{I} \\ b_1\boldsymbol{I} & b_2\boldsymbol{I} & \cdots & b_n\boldsymbol{I} \end{pmatrix}$$

where $\boldsymbol{I}$ is the $p\times p$ identity matrix, gives

$$\boldsymbol{AX} = \begin{pmatrix} a_1\boldsymbol{I} & a_2\boldsymbol{I} & \cdots & a_n\boldsymbol{I} \\ b_1\boldsymbol{I} & b_2\boldsymbol{I} & \cdots & b_n\boldsymbol{I} \end{pmatrix}\begin{pmatrix} \boldsymbol{X_1} \\ \vdots \\ \boldsymbol{X_n} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_j\boldsymbol{X_j} \\ \sum_{j=1}^n b_j\boldsymbol{X_j} \end{pmatrix} = \begin{pmatrix} \boldsymbol{V_1} \\ \boldsymbol{V_2} \end{pmatrix}$$

By the Result of 4.3,

$$\boldsymbol{AX} \sim N_{2p}(\boldsymbol{A\mu},\boldsymbol{A\Sigma_X A}^\top)$$

Straightforward block multiplication shows that $\boldsymbol{A\Sigma_X A}^\top$ has the first block diagonal term:

$$[a_1\boldsymbol{\Sigma},a_2\boldsymbol{\Sigma},\cdots,a_n\boldsymbol{\Sigma}][a_1\boldsymbol{I},a_2\boldsymbol{I},\cdots,a_n\boldsymbol{I}]^\top = (\boldsymbol{a}^\top\boldsymbol{a})\boldsymbol{\Sigma}$$

Another diagonal term is:

$$[b_1\boldsymbol{\Sigma},b_2\boldsymbol{\Sigma},\cdots,b_n\boldsymbol{\Sigma}][b_1\boldsymbol{I},b_2\boldsymbol{I},\cdots,b_n\boldsymbol{I}]^\top = (\boldsymbol{b}^\top\boldsymbol{b})\boldsymbol{\Sigma}$$

And the off-diagonal term is:

$$[a_1\boldsymbol{\Sigma},a_2\boldsymbol{\Sigma},\cdots,a_n\boldsymbol{\Sigma}][b_1\boldsymbol{I},b_2\boldsymbol{I},\cdots,b_n\boldsymbol{I}]^\top = (\boldsymbol{a}^\top\boldsymbol{b})\boldsymbol{\Sigma}$$

These term are the covariance matrix of $\boldsymbol{V_1}$, $\boldsymbol{V_2}$. Consequently, $\boldsymbol{V_1}$ and $\boldsymbol{V_2}$ are independent if and only if $\boldsymbol{a}^\top\boldsymbol{b} = 0$, so that $(\boldsymbol{a}^\top\boldsymbol{b})\boldsymbol{\Sigma} = 0$, i.e., $\boldsymbol{a}^\top$ is orthogonal to $\boldsymbol{b}^\top$.

5 Problem 5(Exercise 4.3)

(a) X_1 and X_2 are not independent. Because $\sigma_{12} = \sigma_{21} = -2 \neq 0$.

(b) X_2 and X_3 are independent. Because $\sigma_{23} = \sigma_{32} = 0$.

(c) (X_1, X_2) and X_3 are independent. Because $\Sigma_{12} = 0$.

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X_1^\top} \\ \boldsymbol{X_2^\top} \\ \boldsymbol{X_2^\top} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X_1'} \\ \boldsymbol{X_2'} \end{pmatrix}$$

Then we can get $\Sigma_{12} = cov(\boldsymbol{X_1'}, \boldsymbol{X_2'}) = (\sigma_{13}, \sigma_{23})^\top = (0,0)^\top$.

(d) $(X_1, X_2)/2$ and X_3 are independent. Because $\sigma = 0$.

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{AX_1'} \\ \boldsymbol{X_2'} \end{pmatrix}, \quad \boldsymbol{A} = (1/2, 1/2)$$

Therefore, $\sigma = \frac{1}{2}\sigma_{13} + \frac{1}{2}\sigma_{23} = 0$

(e) X_2 and $-\frac{2}{5}\boldsymbol{X_1} + \boldsymbol{X_2} - \boldsymbol{X_3}$ are not independent. Because $\sigma_{12}' \neq 0$, $\sigma_{21}' \neq 0$.

$$\boldsymbol{AX} = \begin{pmatrix} 0 & 1 & 0 \\ -\frac{2}{5} & 1 & -1 \end{pmatrix} \begin{pmatrix} \boldsymbol{X_1} \\ \boldsymbol{X_2} \\ \boldsymbol{X_3} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X_2} \\ -\frac{2}{5}\boldsymbol{X_1} + \boldsymbol{X_2} - \boldsymbol{X_3} \end{pmatrix}$$

$$\Sigma' = \boldsymbol{A}\Sigma\boldsymbol{A}^\top = \begin{pmatrix} 0 & 1 & 0 \\ -\frac{2}{5} & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ -2 & -5 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & -\frac{2}{5} \\ 1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -5 & -\frac{5}{21} \\ -\frac{5}{21} & -\frac{41}{5} \end{pmatrix}$$

Therefore, $\sigma_{12}' \neq 0$, $\sigma_{21}' \neq 0$.

6 Problem 6(Exercise 4.5(b))

$$\begin{aligned} f_{X_2|X_1,X_3}(x_2|x_1,x_3) &= \frac{f_{X_1,X_2,X_3}(x_1,x_2,x_3)}{f_{X_1,X_3}(x_1,x_3)} \\ &= \frac{\frac{1}{(2\pi)^{3/2}|\Sigma|^{1/2}} exp\{-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^\top\Sigma^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\}}{\frac{1}{(2\pi)^{3/2}|\Sigma|^{1/2}}\int_{-\infty}^{\infty} exp\{-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^\top\Sigma^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\}dx_2} \\ &= \frac{exp\{-\frac{1}{2}(x_2^2+10x_2+4x_1x_2)\}}{\int_{-\infty}^{\infty} exp\{-\frac{1}{2}(x_2^2+10x_2+4x_1x_2)\}dx_2} \\ &= \frac{exp\{-\frac{1}{2}(x_2-(-2x_1-5))^2\}exp\{-\frac{1}{2}(-4x_1^2-20x_1-25)\}}{exp\{-\frac{1}{2}(-4x_1^2-20x_1-25)\}\int_{-\infty}^{\infty} exp\{-\frac{1}{2}(x_2-(-2x_1-5))^2\}dx_2} \\ &= \frac{1}{\sqrt{2\pi}} exp\{-\frac{1}{2}(x_2-(-2x_1-5))^2\} \end{aligned}$$

Therefore, $X_2|X_1,X_3 \sim N(-2x_1-5, 1)$.

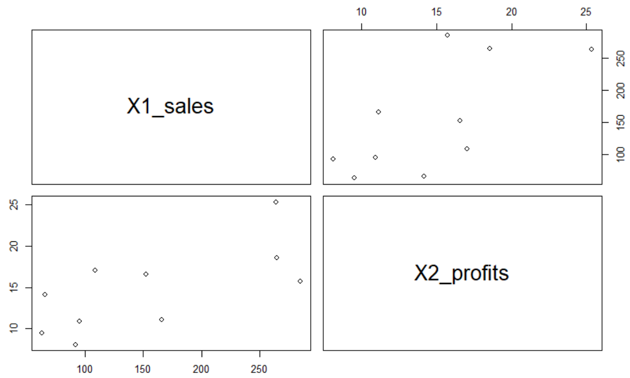

Figure 7.1: Pairs of sales and profits

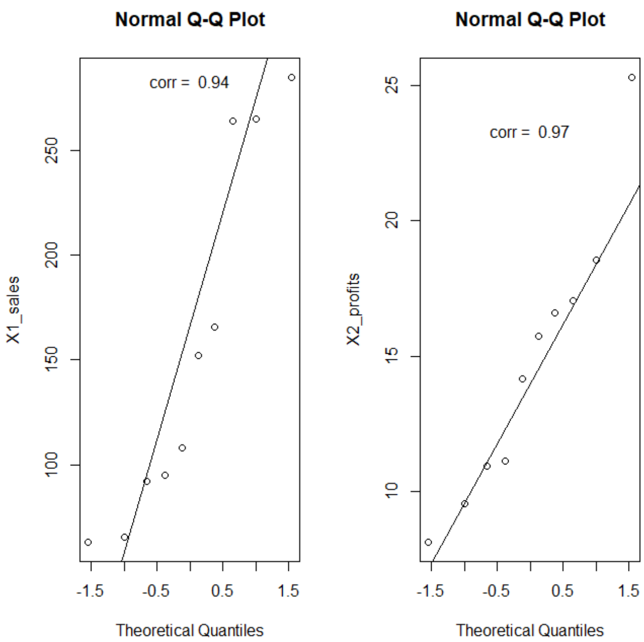

Figure 7.2: QQ plot of sales and profits

7 Problem 7(Exercise 4.24)

(a) From the figure 7.1 and 7.2, we can see all the point are displayed on the same line. That is to say x_i and q_i are approximately linearly related under normality. Therefore, the data appeared to be normally distributed.

(b) We have know that,

$$r_Q = \frac{\sum_{j=1}^{10}(x_j - \bar{x})(q_j - \bar{q})}{\sqrt{\sum_{j=1}^{n}(x_j - \bar{x})^2}\sqrt{\sum_{j=1}^{n}(q_j - \bar{q})^2}}$$

Since $\bar{q} = 0$. By calculating, $\sum_{j=1}^{10}(x_j - \bar{x})q_j = 262.2567$, $\sum_{j=1}^{n}(x_j - \bar{x})^2 = 67288.08$, $\sum_{j=1}^{n}(q_j - \bar{q})^2 = 8.797873$. Therefore, $r_Q = \frac{262.2567}{\sqrt{67288.08}\times\sqrt{8.797873}} = 0.3408543$. Since $r_Q > \alpha = 0.1$, we do not reject the hypothesis of normality, which means it corroborate the result of Part(a).

8 Problem 8(Exercise 4.29)

(a) We can use R to help us calculate this.

$$\boldsymbol{\mu} = \begin{pmatrix} 10.047619 \\ 9.404762 \end{pmatrix}, \quad \boldsymbol{S}^{-1} = \begin{pmatrix} 0.090514367 & -0.009135427 \\ -0.009135427 & 0.033202458 \end{pmatrix}$$

And the statistical distance,(**D** is a 42×1 vector)

$$\boldsymbol{D} = \begin{matrix} 0.55727327 & 0.62644612 & 2.47684081 & 1.36168462 & 0.37779733 & 0.71020612 & 0.98268825 \\ 11.89720210 & 0.09564260 & 1.20199256 & 1.64320425 & 0.91567081 & 5.73115561 & 0.49243545 \\ 0.37779733 & 0.03099507 & 0.97938027 & 5.55828764 & 3.37341190 & 0.62644612 & 3.06500907 \\ 0.75873480 & 0.76771508 & 2.98073447 & 5.58165729 & 0.75873480 & 0.09564260 & 0.56135063 \\ 0.19935617 & 2.01859764 & 0.30593753 & 0.74176804 & 1.36758817 & 6.94615941 & 1.44188162 \\ 0.03099507 & 1.75386295 & 3.05829543 & 7.80873831 & 0.47612607 & 0.53199062 & 1.41945974 \end{matrix}$$

(b) According to the hint, the desired χ^2 values are then the diagonal elements of the matrix. Using the *qchisq(0.5,2)* in R, we can calculate the that 0.5 quantile of χ^2 is 1.386294. i.e. $\chi_2^2(0.5) = 1.386294$. Then we calculate the number of values below 1.386294 is 26. Therefore, the proportion is:

$$p = 26/42 = 0.6190476$$

(c) The result can be seen in Figure 8.1.

9 Problem 9

(a) We consider the vatiate $Z = |Z_0|sgn(XY)$ in different stages. When $Z_0 \geqslant 0$,

$$Z = \begin{cases} Z_0, & X > 0,\ Y > 0 \ or\ X < 0,\ Y < 0 \\ -Z_0, & X > 0,\ Y < 0 \ or\ X < 0,\ Y > 0 \\ 0, & X = 0 \ or\ Y = 0 \end{cases}$$

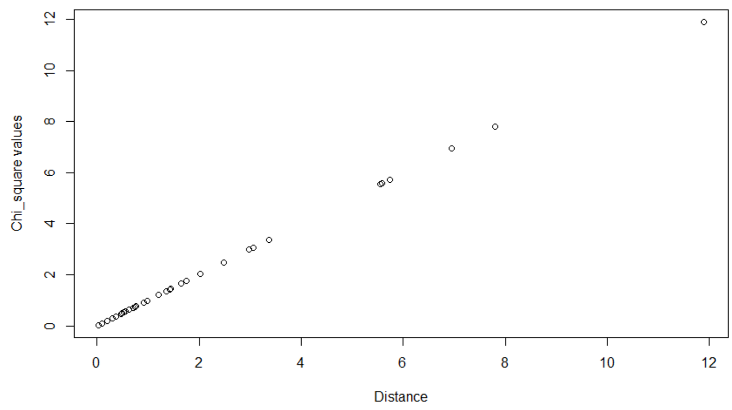

Figure 8.1: Chi-square plot

When $Z_0 < 0$,

$$Z = \begin{cases} -Z_0, & X > 0,\ Y > 0\ or\ X < 0,\ Y < 0 \\ Z_0, & X > 0,\ Y < 0\ or\ X < 0,\ Y > 0 \\ 0, & X = 0\ or\ Y = 0 \end{cases}$$

Therefore, when $X \neq 0$ and $Y \neq 0$, $Z \sim N(0,1)$, i.e., Z is also normally distributed.

(b) We consider the probability of X, Y, and Z (every time only two of them).
When $Z_0 \geqslant 0$, $P(|Z_0|) = P(Z_0)$.

$$P(Z) = P(|Z_0|sgn(XY)) = P(Z_0 sgn(XY)) = sgn(XY)P(Z_0)$$

$$P(XZ) = P(X|Z_0|sgn(XY)) = P(XZ_0 sgn(XY)) = sgn(XY)P(XZ_0) = sgn(XY)P(X)P(Z_0)$$

Since X and Z_0 are independent. Therefore,

$$P(XZ) = P(X)P(Z)$$

When $Z_0 < 0$, $P(|Z_0|) = P(-Z_0) = -P(Z_0)$

$$P(Z) = P(|Z_0|sgn(XY)) = P(-Z_0 sgn(XY)) = -sgn(XY)P(Z_0)$$

$$P(XZ) = P(X|Z_0|sgn(XY)) = P(X(-Z_0)sgn(XY)) = -sgn(XY)P(XZ_0) = -sgn(XY)P(X)P(Z_0)$$

Therefore,

$$P(XZ) = P(X)P(Z)$$

Similarly, we can get $P(YZ) = P(Y)P(Z)$. And we already known X and Y are independent, so X, Y , and Z are pairwise independent.

(c) We can use counter-evidence to proof that $\boldsymbol{V} = (X, Y, Z)^\top$ is not normally distributed. If $\boldsymbol{V}$ is normally distributed, we have the pdf

$$f(\boldsymbol{V}) = \frac{1}{(\sqrt{2\pi})^3|\Sigma|^{1/2}} exp(-\boldsymbol{V}\Sigma^{-1}\boldsymbol{V}/2)$$

and

$$f(\mathbf{0}) = \frac{1}{(\sqrt{2\pi})^3}$$

However,

$$\begin{aligned} f(\mathbf{0}) &= f(X=0)\cdot f(Y=0)\cdot f(Z=0|X=0,Y=0) \\ &= \frac{1}{(\sqrt{2\pi})}\cdot\frac{1}{(\sqrt{2\pi})}\cdot 1 \\ &= \frac{1}{2\pi} \end{aligned}$$

Since under the condition of $X = 0$ and $Y = 0$, $P(Z = 0) = 1$. Now we have $\frac{1}{(\sqrt{2\pi})^3} \neq \frac{1}{2\pi}$. It contradicts. Therefore, $\boldsymbol{V} = (X, Y, Z)^\top$ is not normally distributed.

10 Problem 10

Proof[?]:

(a) According to what we have known,

$$\boldsymbol{Y}_{i-1} = (S_{i1}, \cdots, S_{i,i-1})^\top = (\sum_k^m Z_{ik}Z_{1k}, \cdots, \sum_k^m Z_{ik}Z_{i-1,k})^\top$$

Since $\boldsymbol{\Sigma} = \boldsymbol{I}_p$, $Z_{11}, \cdots, Z_{1m}, \cdots, Z_{p1}, \cdots, Z_{pm}$ are all iid $N(0,1)$. According to the hint, given the Z_{jk}, the S_{ij} are linear functions of jointly normally distributed RVs. It follows that conditional on Z_{jk}, $j = 1, \cdots, i-1$, $k = 1, \cdots, m$, $\boldsymbol{Y}_{i-1} \sim N(\mathbf{0}, \boldsymbol{S}_{[i-1]})$.

(b) According to the hint, $\boldsymbol{U}_{i-1,k} = (Z_{1k}, \cdots, Z_{i-1,k})^\top$, $i = 2, \cdots, p$, $k = 1, \cdots, m$. Then $\boldsymbol{S}_{[i-1]} = \sum_{k=1}^m = \boldsymbol{U}_{i-1,k}\boldsymbol{U}_{i-1,k}^\top$ and

$$S_{ii.1,\cdots,i-1} = \sum_k^m (Z_{ik} - \boldsymbol{U}_{i-1,k}^\top \boldsymbol{S}_{[i-1]}^{-1}\boldsymbol{U}_{i-1,k})$$

Now due to the mutual independence of $Z_{11}, \cdots, Z_{pm}$, for every $k = 1, \cdots, m$,

$$\begin{aligned} cov(\boldsymbol{Y}_{i-1}, Z_{ik}) &= [cov(\sum_{u=1}^m Z_{iu}Z_{1u}, Z_{ik}), \cdots, cov(\sum_{u=1}^m Z_{iu}Z_{i-1,u}, Z_{ik})|Z_{ik}, j = 1, \cdots, i-1, k = 2, \cdots, m]^\top \\ &= (Z_{1k}, \cdots, Z_{i-1,k})^\top \\ &= \boldsymbol{U}_{i-1,k} \end{aligned}$$

And we have $E(\boldsymbol{Y}_{i-1}\boldsymbol{Y}_{i-1}^\top | Z_{jk}, j=1,\cdots,i-1, k=1,\cdots,m]^\top) = \boldsymbol{S}_{[i-1]}$. Hence one gets,

$$cov(\boldsymbol{Y}_{i-1}, Z_{ik} - \boldsymbol{Y}_{i-1}^\top \boldsymbol{S}_{[i-1]}^{-1}\boldsymbol{U}_{i-1,k} | j=1,\cdots,i-1, k=1,\cdots,m) = \boldsymbol{U}_{i-1,k} - \boldsymbol{S}_{[i-1]}\boldsymbol{S}_{[i-1]}^{-1}\boldsymbol{U}_{i-1,k} = \boldsymbol{0}$$

Once again, from a property of linear function of normal variables, we can get the independence of $\boldsymbol{Y}_{i-1}$ and $(Z_{i1} - \boldsymbol{U}_{i-1,1}^\top \boldsymbol{S}_{[i-1]}^{-1}\boldsymbol{U}_{i-1,k}, \cdots, Z_{im} - \boldsymbol{U}_{i-1,m}^\top \boldsymbol{S}_{[i-1]}^{-1}\boldsymbol{U}_{i-1,k})$. This implies the independence of $\boldsymbol{Y}_{i-1}$ and $S_{ii.1,\cdots,i-1}$.

(c) We begin with the identity

$$S_{ii} = \boldsymbol{Y}_{i-1}^\top \boldsymbol{S}_{[i-1]}^{-1}\boldsymbol{Y}_{i-1} + S_{ii.1,\cdots,i-1}$$

By (b), conditional on Z_{jk}, $j=1,\cdots,i-1$, $k=1,\cdots,m$, $\boldsymbol{Y}_{i-1}^\top \boldsymbol{S}_{[i-1]}^{-1}\boldsymbol{Y}_{i-1}$ is distributed independently with $S_{ii.1,\cdots,i-1}$. Set $\phi_W(t)$ as a generic symbol for the MGF of a random variable of W. Now we have,

$$\phi_{S_{ii}}(t) = \phi_{\boldsymbol{Y}_{i-1}^\top \boldsymbol{S}_{[i-1]}^{-1}\boldsymbol{Y}_{i-1}}(t)\phi_{S_{ii.1,\cdots,i-1}}(t)$$

By (a), for the same conditioning variables $Z_{jk}, j=1,\cdots,i-1$, $k=1,\cdots,m$, $\boldsymbol{Y}_{i-1}^\top \boldsymbol{S}_{[i-1]}^{-1}\boldsymbol{Y}_{i-1} \sim \chi_{i-1}^2$. Due to the mutual independence of all the Z_{jk}, the conditional distribution of $S_{ii} = \sum_{k=1}^m Z_{ik}^2$, given Z_{jk}, $j=1,\cdots,i-1$, $k=1,\cdots,m$, is the same as its unconditional distribution which is χ_m^2. Hence, according to the MGF we stated before, conditional on Z_{jk}, $j=1,\cdots,i-1$, $k=1,\cdots,m$,

$$\phi_{S_{ii.1,\cdots,i-1}} = (1-2t) - \frac{m-i+1}{2}$$

which implies (c).

11 Appedix

This part include some **R** code of the assignment.

11.1 Problem 7(Exercise 4.24)

```
## Input the data
X1_sales<-c(108.28, 152.36, 95.04, 65.45, 62.97, 263.99, 265.19, 285.06, 92.01, 165.68)
X2_profits<-c(17.05, 16.59, 10.91, 14.14, 9.52, 25.33, 18.54, 15.73, 8.10, 11.13)
company<-as.matrix(cbind(X1_sales,X2_profits))
colMeans(company)
S<-cov(company)
apply(company, 2, var)
pairs(company)


##Construct QQ plot
dev.new()
par(mfrow = c(1,2))
for(i in 1:2) {
  y = company[,i]
  v=qqnorm(y, ylab = colnames(company)[i])
  text(0, max(v$y-2), paste("corr = ", round(cor(v$x,v$y),2)))
  qqline(y)
}
```

```
##Calculate the r_Q

##calculate sum(x_j-bar(x))^2
mean_x1_sales=mean(X1_sales)
mean_x2_profits=mean(X2_profits)
for(i in 1:10){
  x_1[i]=X1_sales[i]-mean_x1_sales
  x_22[i]=(X1_sales[i]-mean_x1_sales)^2
  x_33=sum(x_22)
}

##calculate sum(q_j-bar(q))^2
for(j in 1:10){
  pl[j]<-c((j-0.5)/10)
  q_j[j]<-qnorm(pl[j],mean=0, sd=1)
  q_bar<-0
  y_1[j]<-(q_j[j]-q_bar)
  y_22[j]=(q_j[j]-q_bar)^2
  y_33=sum(y_22)

}

##calculate sum(x_j-bar(x))(q_j-bar(q))
z<-sum(x_1*y_1)

##calculate r_Q
z_1<-sqrt(x_33)

z_2<-sqrt(y_33)

r_Q<-z/(z_1*z_2)
```

11.2 Problem 8(Exercise 4.29

```
##(a)
data<-read.csv('Table_1_5.csv',sep=',')
x_j=as.matrix(data[,c(5,6)])
mu=colMeans(x_j)
y_j=x_j-mu
S=cov(x_j)
S_1=solve(S)
for (i in 1:42){
  D[i]=t(y_j[i,])%*%S_1%*%y_j[i,]
}

#(b)
chi_y=y_j%*%S_1%*%t(y_j)
chi_value=diag(chi_y)
q=qchisq(0.5,2)
i<-length(which(chi_value<q))
pop=i/42
## 0.6190476

##(c)
plot(D, chi_value, xlab="Distance", ylab="Chi_square␣values")
```

11.3 Problem 10

References

[1] Malay Ghosh & Bimal K Sinha (2002) *A Simple Derivation of the Wishart Distribution*, The American Statistician, 56:2, 100-101, DOI: 10.1198/000313002317572754