วฐา มินเสน รหัส 48850226
http://beam.to/statistics

HW Chapter 5

5.1

(a) Evaluate T^2 ,for testing $H_o : \mu' = [7\ ,11]$,using the data

จากสูตร $T^2 = n(\bar{x}-\mu_o)'S^{-1}(\bar{x}-\mu_o)$

$$X = \begin{pmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{pmatrix}$$

$$\bar{x} = \begin{bmatrix} 6 & 10 \end{bmatrix}$$

$$S = \begin{bmatrix} 8 & -3.33333 \\ -3.33333 & 2 \end{bmatrix}$$

$$S^{-1} = \begin{bmatrix} 0.409091 & 0.681818 \\ 0.681818 & 1.636364 \end{bmatrix}$$

$$T^2 = 4 \begin{bmatrix} -1 & -1 \end{bmatrix} \begin{bmatrix} 0.409091 & 0.681818 \\ 0.681818 & 1.636364 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$$= 4 \begin{bmatrix} -1.09091 & -2.31818 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$$= 4 \quad x \quad 3.409091$$

$$= 13.63636$$

(b) Specify the distribution of T^2 for the situation in (a)

T^2 is distributed as $\dfrac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha)$

(c) Using (a) and (b), test H_o at the $\alpha = 0.05$ level. What conclusion do you reach?

1. $$H_0 : \mu' = [7\ ,11]$$
$$H_1 : \mu' \neq [7\ ,11]$$

2. $\alpha = 0.05$

3. T^2 = 13.63636

4. $\frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha) = \frac{(n-1)p}{(n-p)} F_{p,n-p}(0.05)$ = 3 x 19 = 57

5. T^2 = 13.63636 < $\frac{(n-1)p}{(n-p)} F_{p,n-p}(0.05)$ = 57

We not reject H_o at the 5% level of significance. So, $\mu' = [7\ ,11]$

5.9. Harry Roberts, a naturalist for the Alaska Fish and Game department, studies grizzly bears with the goal of maintaining a healthy population. Measurements on n=61 bears provided the following summary statistics (see also Exercise 8.23):

Variable	Weight (kg)	Body length (cm)	Neck (cm)	Girth (cm)	Head length (cm)	Head width (cm)
Sample mean $\bar{x}$	95.52	164.38	55.69	93.39	17.98	31.13

$$S = \begin{pmatrix} 3266.46 & 1343.97 & 731.54 & 1175.5 & 162.68 & 238.37 \\ 1343.97 & 721.91 & 324.25 & 537.35 & 80.17 & 117.73 \\ 731.54 & 324.25 & 179.28 & 281.17 & 39.15 & 56.8 \\ 1175.5 & 537.35 & 281.17 & 474.98 & 63.73 & 94.85 \\ 162.68 & 80.17 & 39.15 & 63.73 & 9.95 & 13.88 \\ 238.37 & 117.73 & 56.8 & 94.85 & 13.88 & 21.26 \end{pmatrix}$$

(a) Obtain the **large sample** 95% simultaneous confidence intervals for the six population mean body measurements.

$\chi_p^2(\alpha) = 12.59$

$$\mu_1 = \bar{x}_1 \pm \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{11}}{n}} \qquad = 95.52 \pm \sqrt{12.59}\sqrt{\frac{3266.46}{61}}$$

$$\boxed{69.553465 \quad \le \mu_1 \le \quad 121.48653}$$

$$\mu_2 = \bar{x}_2 \pm \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{22}}{n}} \qquad = 164.38 \pm \sqrt{12.59}\sqrt{\frac{721.91}{61}}$$

$$\boxed{152.17278 \quad \le \mu_2 \le \quad 176.58722}$$

$$\mu_3 = \bar{x}_3 \pm \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{33}}{n}} \qquad = 55.69 \pm \sqrt{12.59}\sqrt{\frac{179.28}{61}}$$

$$\boxed{49.606672 \quad \le \mu_3 \le \quad 61.773328}$$

$$\mu_4 = \bar{x}_4 \pm \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{44}}{n}} \qquad = 93.39 \pm \sqrt{12.59}\sqrt{\frac{474.98}{61}}$$

$$\boxed{83.488227 \quad \le \mu_4 \le \quad 103.29177}$$

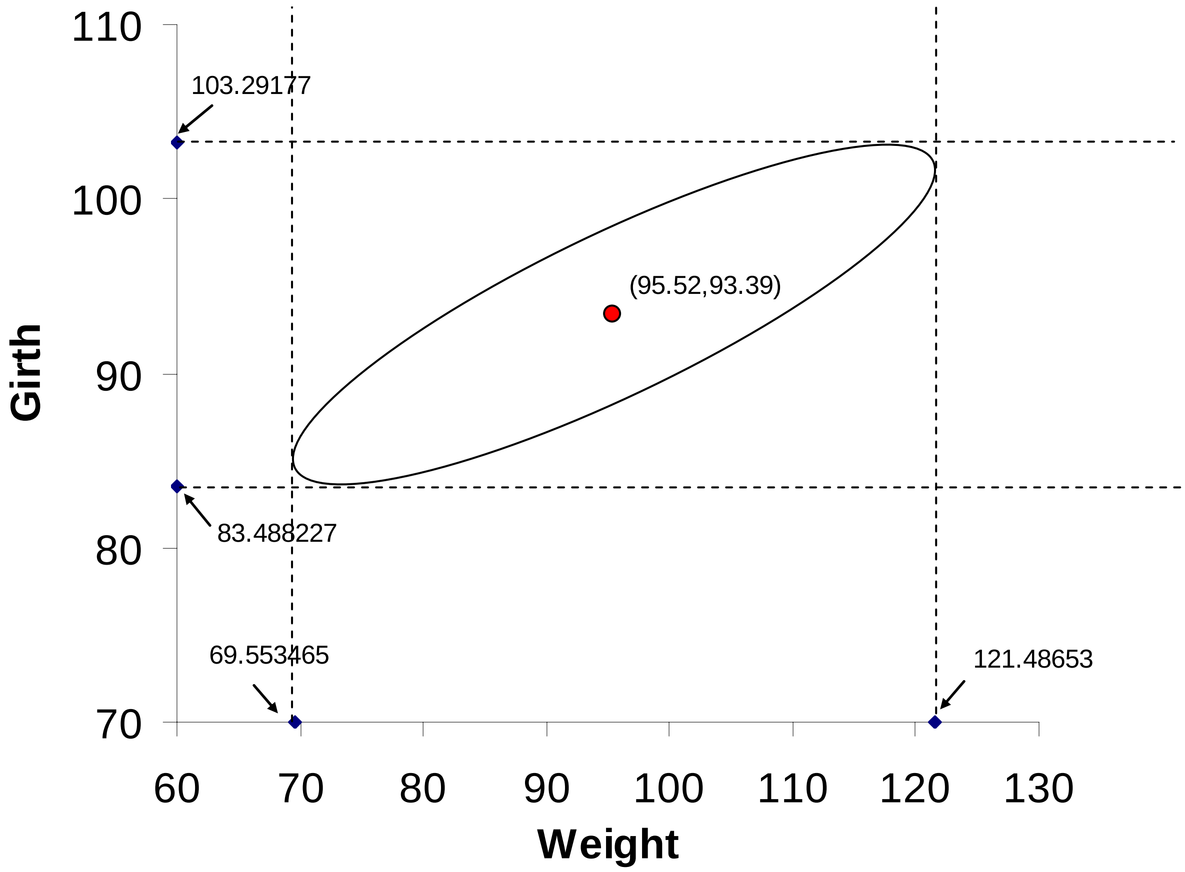
$$\mu_5 = \bar{x}_5 \pm \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{55}}{n}} \qquad = 17.98 \pm \sqrt{12.59}\sqrt{\frac{9.95}{61}}$$

$$\boxed{16.546866 \quad \le \mu_5 \le \quad 19.413134}$$

$$\mu_6 = \bar{x}_6 \pm \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{66}}{n}} \qquad = 31.13 \pm \sqrt{12.59}\sqrt{\frac{21.26}{61}}$$

$$\boxed{29.035131 \quad \le \mu_6 \le \quad 33.224869}$$

(b) Obtain the **large sample** 95% simultaneous confidence ellipse for mean weight and mean girth.



(c) Obtain the 95% **<u>Bonferroni confidence intervals</u>** for the six means in **Part a**.

**** เนื่องจากโจทย์ ข้อนี้ไม่แน่ใจว่าต้องการให้ใช้ Large Sample หรือไม่ จึงทำทั้งสองแบบ คือ**

$$\bar{x}_p - t_{n-1,(\alpha/2p)}\sqrt{\frac{s_{pp}}{n}} \le \mu_p \le \bar{x}_p + t_{n-1,(\alpha/2p)}\sqrt{\frac{s_{pp}}{n}}$$

และ $$\bar{x}_p - Z_{(\alpha/2p)}\sqrt{\frac{s_{pp}}{n}} \le \mu_p \le \bar{x}_p + Z_{(\alpha/2p)}\sqrt{\frac{s_{pp}}{n}}$$

ตัวอย่างขนาดเล็ก $t_{n-1,(\alpha/2p)} = t_{60,(0.004167)} = 2.728549$

$$\bar{x}_p - t_{n-1,(\alpha/2p)}\sqrt{\frac{s_{pp}}{n}} \le \mu_p \le \bar{x}_p + t_{n-1,(\alpha/2p)}\sqrt{\frac{s_{pp}}{n}}$$

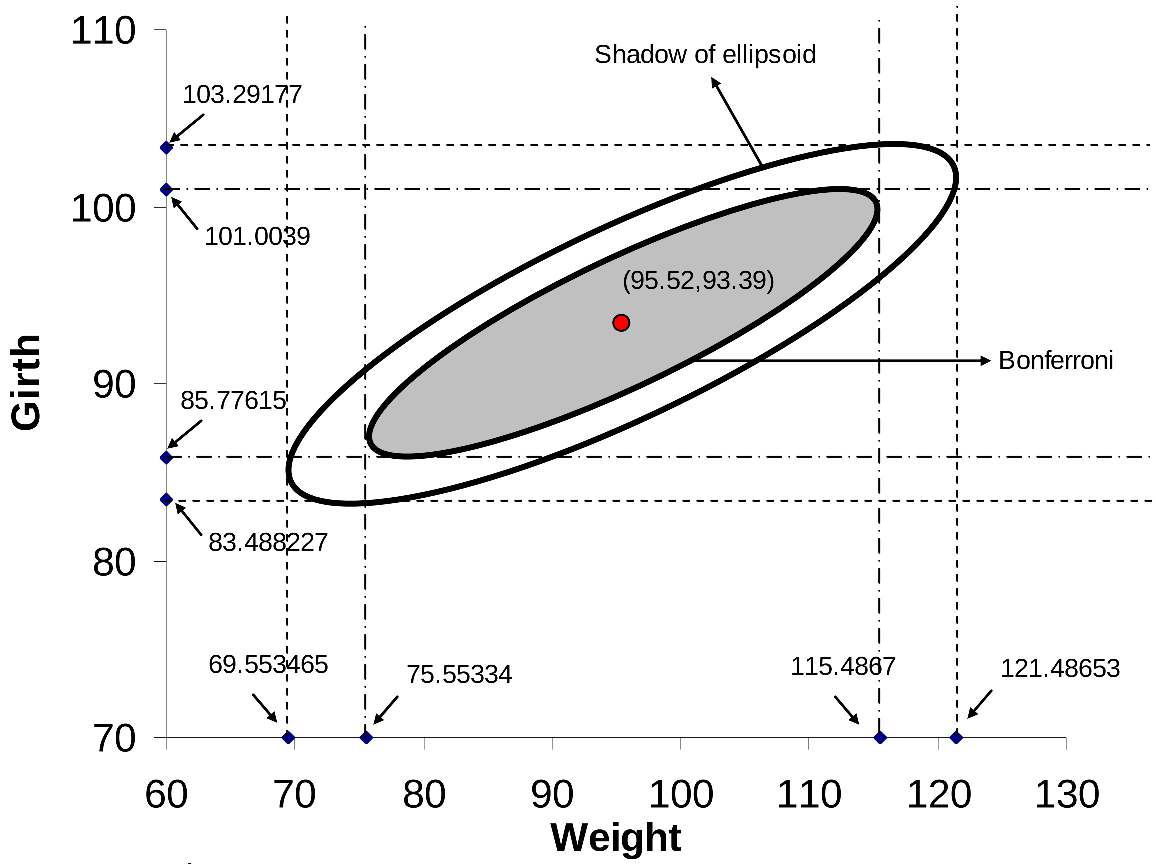
75.55334	$\le \mu_1 \le$	115.4867
154.9934	$\le \mu_2 \le$	173.7666
51.0123	$\le \mu_3 \le$	60.3677
85.77615	$\le \mu_4 \le$	101.0039
16.87801	$\le \mu_5 \le$	19.08199
29.51917	$\le \mu_6 \le$	32.74083

ตัวอย่างขนาดใหญ่ $Z_{(0.004167)} = 2.638255$

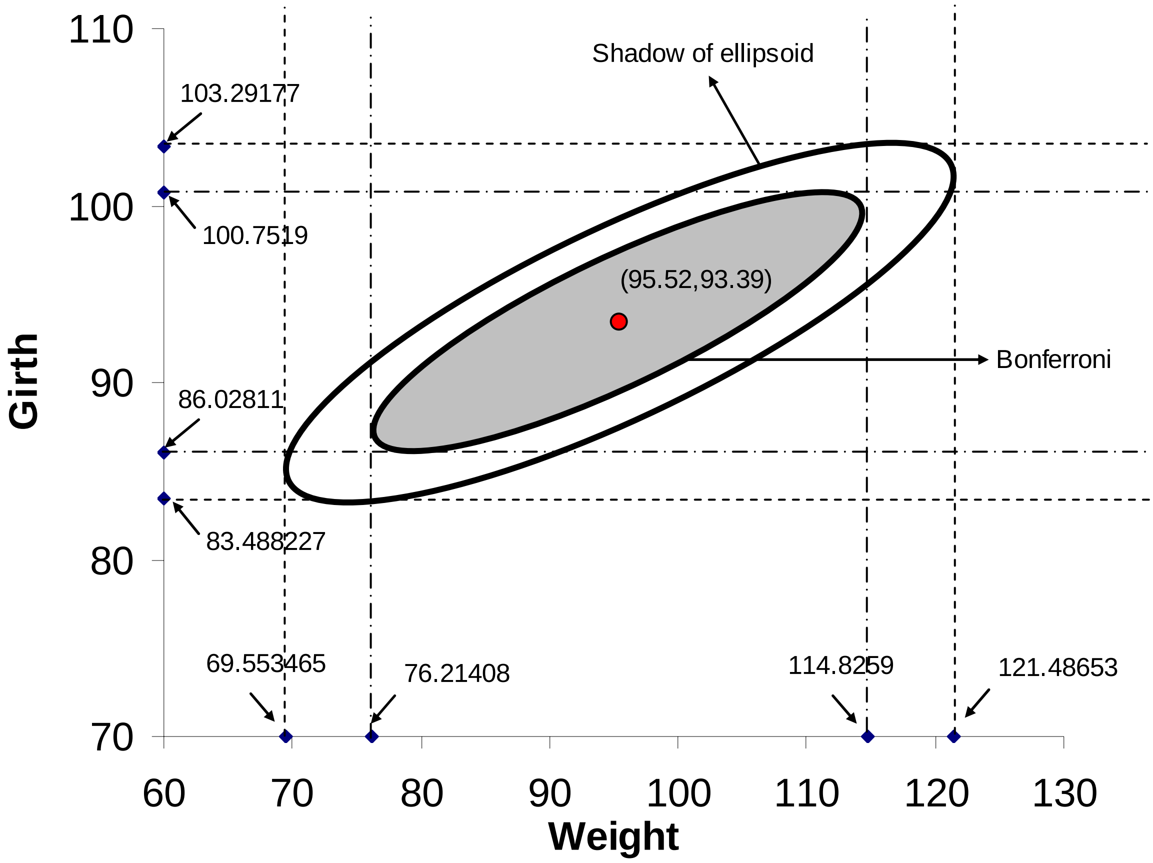
$$\bar{x}_p - Z_{(\alpha/2p)}\sqrt{\frac{s_{pp}}{n}} \le \mu_p \le \bar{x}_p + Z_{(\alpha/2p)}\sqrt{\frac{s_{pp}}{n}}$$

76.21408	$\le \mu_1 \le$	114.8259
155.304	$\le \mu_2 \le$	173.456
51.16709	$\le \mu_3 \le$	60.21291
86.02811	$\le \mu_4 \le$	100.7519
16.91448	$\le \mu_5 \le$	19.04552
29.57248	$\le \mu_6 \le$	32.68752

(d) Refer to Part b. Construct the 95% Bonferroni confidence rectangle for the mean weight and mean girth using m = 6. Compare this rectangle with the confidence ellipse in Part b.



** **Bonfferoni เมื่อตัวอย่างขนาดเล็ก**



**** Bonfferoni เมื่อตัวอย่างขนาดใหญ่**

สรุปได้ว่า 95% Bonferroni Simultaneous confidence interval for component means จะใช้ตัวสถิติ t เมื่อตัวอย่างขนาดเล็ก หรือ ใช้ตัวสถิติ Z เมื่อตัวอย่างขนาดใหญ่ ก็จะมีรูป Ellipsoid เล็กกว่าการใช้ 95% T^2 Simultaneous confidence interval for component means

(e) Obtain the 95% Bonferroni confidence interval for

mean head width – mean head length

Using m = 6 + 1 = 7 to allow for this statement as well as statements about each individual mean.

**** ใช้ Bonfferoni ตัวอย่างขนาดเล็กเพราะโจทย์ไม่ได้อ้างอิงข้อย่อยอื่นๆ และไม่ได้บอกให้ใช้ Large Sample**

$$\mu_{width} - \mu_{length} = \overline{x}_{width} - \overline{x}_{length} \pm t_{n-1,(\alpha/2m)} \sqrt{\frac{s_{width} - 2s_{width,length} + s_{length}}{n}}$$

$$\mu_{width} - \mu_{length} = 31.13 - 17.98 \pm t_{61-1,(0.05/14)} \sqrt{\frac{21.26 - 2(13.88) + 9.95}{61}}$$

$$t_{61-1,(0.05/14)} = 2.78547$$

$$\mu_{width} - \mu_{length} = 31.13 - 17.98 \pm 2.78547 \sqrt{\frac{21.26 - 2(13.88) + 9.95}{61}}$$

$$\boxed{12.4875 \le \mu_{width} - \mu_{length} \le 13.8124}$$

Note: ถ้าใช้ **Bonfferoni เมื่อตัวอย่างขนาดใหญ่** ก็ได้คำตอบใกล้เคียงกัน

$$\boxed{12.51024 \le \mu_{width} - \mu_{length} \le 13.78976}$$

5.11. A physical anthropologist performed a mineral analysis of nine ancient Peruvian hairs. The results for the chromium(x_1) and strontium (x_2) levels, in parts per million (ppm), were as follows:

X_1	0.48	40.53	2.19	0.55	0.74	0.66	0.93	0.37	0.22
X_2	12.57	73.68	11.13	20.03	20.29	0.78	4.64	0.43	1.08

It is known that low levels (less than or equal to .100 ppm) of chromium suggest the presence of diabetes. While strontium is an indication of animal protein intake.

(a) Construct and plot a 90% joint confidence ellipse for the population mean random sample form individuals belonging to a particular ancient Peruvian culture.

$$\bar{x} = \begin{bmatrix} 5.1856 \\ 16.07 \end{bmatrix}$$

$$S = \begin{bmatrix} 176 & 287.24 \\ 287.24 & 527.85 \end{bmatrix}$$

$$\lambda_i = \begin{bmatrix} 688.76 \\ 15.094 \end{bmatrix}$$

$$\lambda_1 = 688.7594 \qquad e_1 = \begin{bmatrix} 0.4887 \\ 0.8724 \end{bmatrix}$$

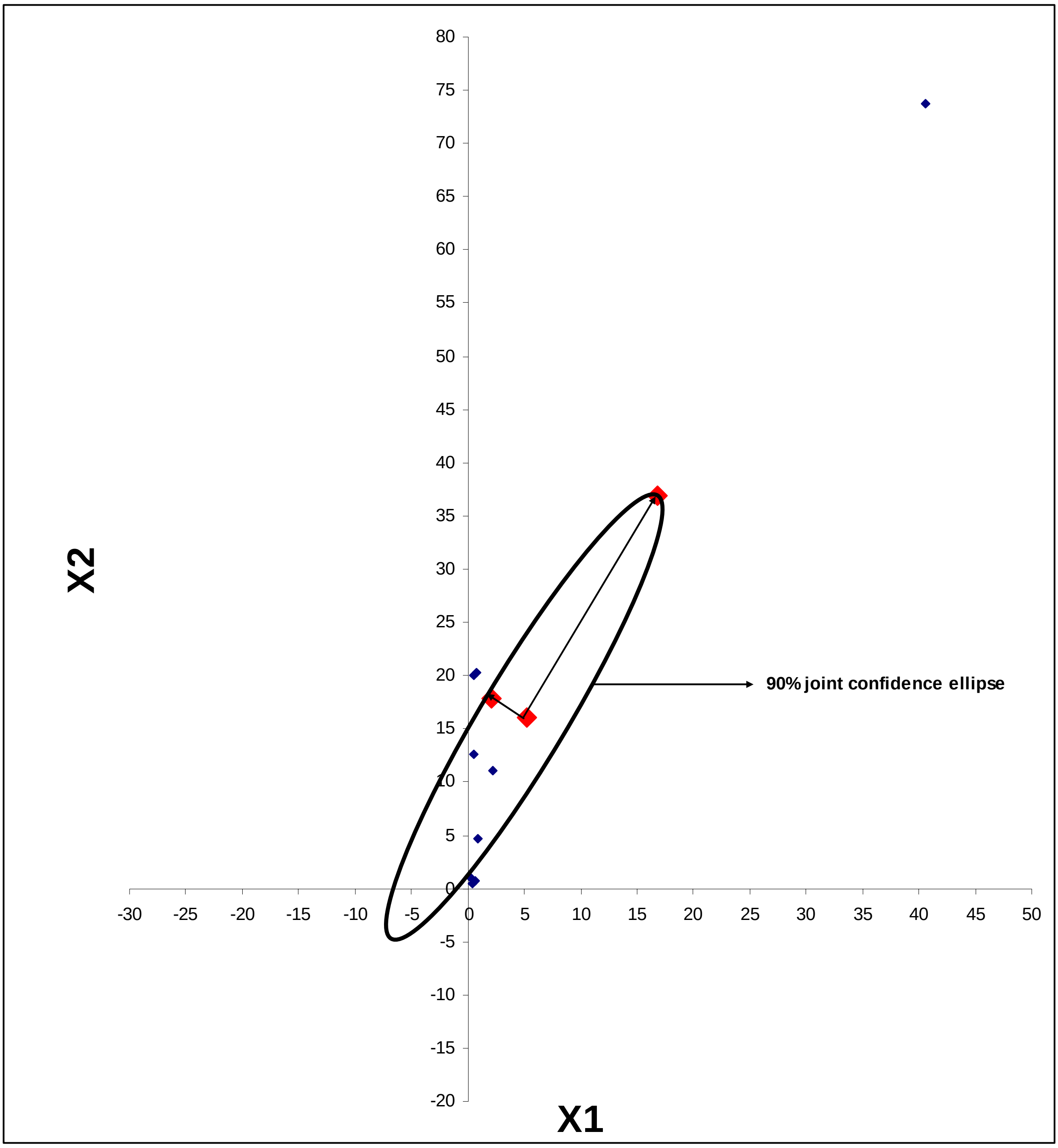
$$\lambda_2 = 15.0941 \qquad e_2 = \begin{bmatrix} -0.872 \\ 0.4887 \end{bmatrix}$$

$$\pm\sqrt{\lambda_1}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)} = 26.244 \quad \text{x} \quad 0.9099 \quad = \quad 23.88$$

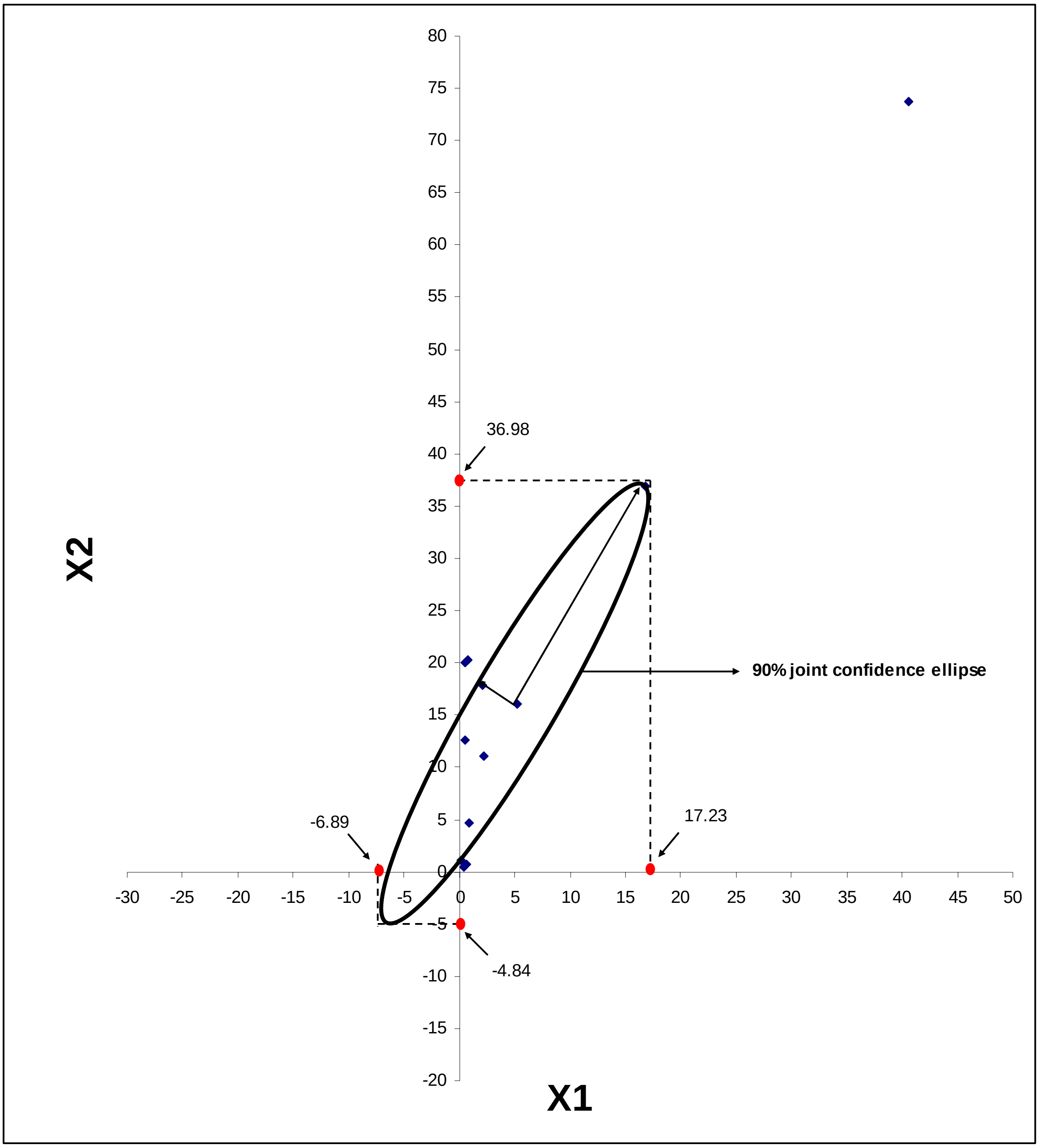
$$\pm\sqrt{\lambda_1}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)}e_1 = \begin{bmatrix} 11.67 \\ 20.833 \end{bmatrix}$$

$$\pm\sqrt{\lambda_2}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)} = 3.8851 \quad \text{x} \quad 0.9099 \quad = \quad 3.5351$$

$$\pm\sqrt{\lambda_2}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)}e_2 = \begin{bmatrix} -3.084 \\ 1.7276 \end{bmatrix}$$


X2
X1
90% joint confidence ellipse
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
-5
-10
-15
-20
-30
-25
-20
-15
-10
-5
0
5
10
15
20
25
30
35
40
45
50

(b) Obtain the individual simultaneous 90% confidence intervals for μ_1 and μ_2 by "projecting" the ellipse constructed in Part a on each coordinate axis.(Alternatively, we could use Result 5.3.) Does it appear as if this Peruvian culture has a mean strontium level of 10? That is , are any of the points (μ_1 arbitrary,10) in the confidence regions? Is $[.30, 10]'$ a plausible value for μ ? Discuss.



$-6.89 \le \mu_1 \le 17.23$ $-4.84 \le \mu_2 \le 36.98$

Does it appear as if this Peruvian culture has a mean strontium level of 10?

Yes, Peruvian culture has a mean strontium level of 10.

That is , are any of the points (μ_1 arbitrary,10) in the confidence regions?
ถ้าให้ μ_1 arbitrary และให้ μ_2 = 10 จะพบว่าค่า μ_1 จะเป็นไปได้คือ
$-1.98 \le \mu_1 \le 5.74$ จึงจะทำให้การทดสอบไม่ปฏิเสธ H_o ที่ระดับนัยสำคัญ .10
ดังนั้นจะมีจุดอยู่ 1 จุด คือ ชุดข้อมูลที่ 2 ที่มีค่าอยู่นอกช่วงที่กำหนดส่วนจุดอื่นๆ อยู่ใน Confidence region

Is $[.30\ ,10]'$ a plausible value for μ ? Discuss.

$$n(\bar{x}-\mu_o)'S^{-1}(\bar{x}-\mu_o) \le \frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha)$$

$$n(\bar{x}-\mu_o)'S^{-1}(\bar{x}-\mu_o) = \quad 9\begin{bmatrix}4.8856 & 6.07\end{bmatrix}\begin{bmatrix}0.0508 & -0.028\\ -0.028 & 0.0169\end{bmatrix}\begin{bmatrix}4.8856\\ 6.07\end{bmatrix}$$

$$9\begin{bmatrix}0.0803 & -0.032\end{bmatrix}\begin{bmatrix}4.8856\\ 6.07\end{bmatrix}$$

$$9 \quad \text{x} \quad 0.1969$$

$$n(\bar{x}-\mu_o)'S^{-1}(\bar{x}-\mu_o) = \quad 1.7725$$

$$\frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha) = \ 7.4514$$

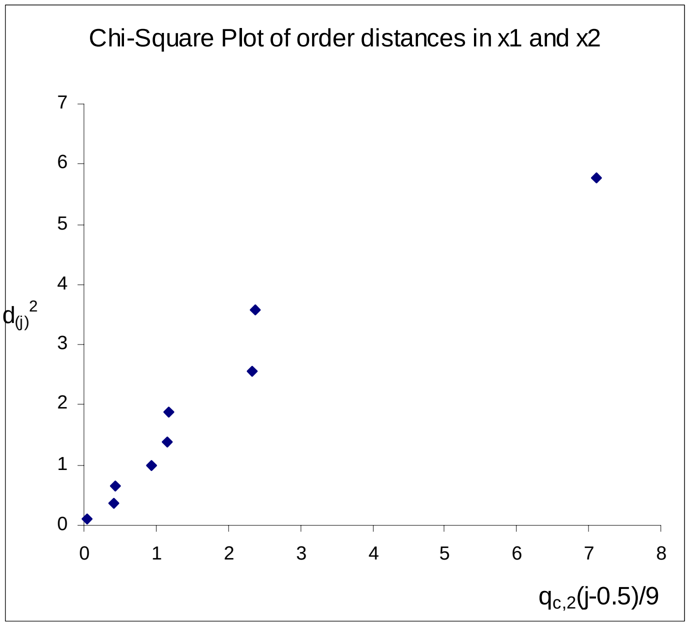
$$n(\bar{x}-\mu_o)'S^{-1}(\bar{x}-\mu_o) = 1.7725 \le \frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha) = 7.4514$$

We not reject H_o at the 10% level of significance. So, $\mu' = [.30\ ,10]'$ plausible value of mean chromium = .30 and plausible value of mean strontium = 10

(c) Do these data appear to be bivariate normal? Discuss their status with **reference to Q-Q plots and a scatter diagram**. If the data are not bivariate normal, what implications does this have for the results in Parts a and b?

การตรวจสอบ bivariate normal อยู่ในบทที่ 4 หน้า 183 มีวิธีการทำดังนี้

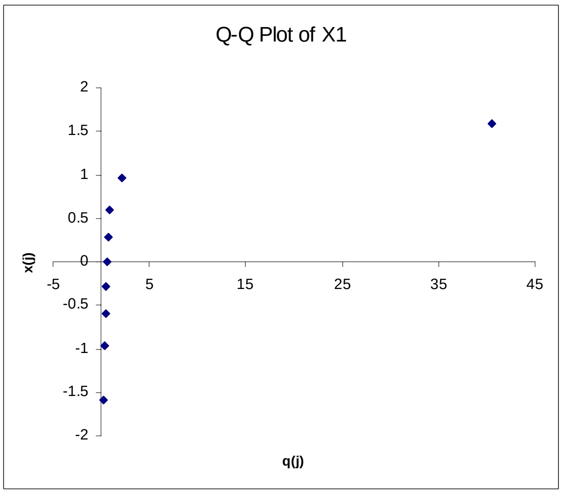
$(x-\bar{x})'S^{-1}(x-\bar{x}) \le \chi_2^2(.5)$ หรือ จัดทำ Chi-Square Plot

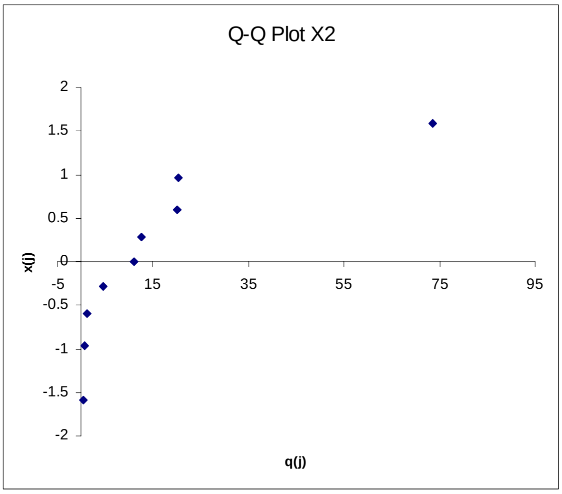


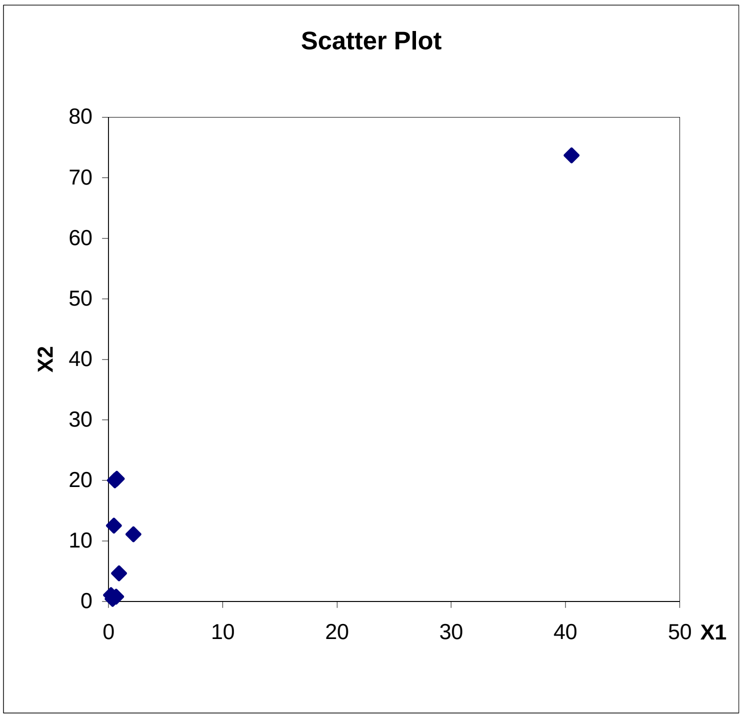
ข้อมูลชุดที่	คำนวณ		Chi-Square(0.5)
1	0.42	<	1.39
2	**7.10**	**>**	**1.39**
3	0.05	<	1.39
4	**2.37**	**>**	**1.39**
5	**2.34**	**>**	**1.39**
6	1.17	<	1.39
7	0.44	<	1.39
8	1.16	<	1.39
9	0.94	<	1.39

จากการทดสอบมี 3 จุดที่มากกว่า 1.39 หรือประมาณ 33.33% ซึ่งไม่เกิน 50% แสดงเป็น Bivariate normal

แต่โจทย์ให้พิจารณาจาก Q-Q Plot และ Scatter diagram ซึ่งพิจารณาได้ดังนี้







การตัดสินใจเกี่ยวกับ Bivariate normal จากรูป Q-Q Plot และ Scatter Diagram เป็นไปได้ยาก และมีโอกาสตัดสินใจพลาดสูง แต่จากรูป Q-Q Plot และ Scatter Diagram สามารถพิจารณาลักษณะของข้อมูลเพื่อตัดสินใจบางอย่างได้ดังนี้ มีข้อมูล 1 ชุดที่มีค่าเกินกว่าข้อมูลชุดอื่นๆ มาก คือข้อมูลชุดที่ 2 ดังนั้นควรตัดข้อมูลชุดนี้ออกจากการวิเคราะห์ เพราะอาจทำให้เกิดความผิดพลาดในการสรุปผลได้ ดังนั้น ใน (a) และ (b) ที่ได้จัดทำไปแล้วจะมีช่วงความเชื่อมั่นที่มากกว่าที่ควรจะเป็น

(d) **Repeat the analysis** with the obvious "outlying" observation removed. Do the inferences change? Comment.

พิจารณาจากข้อ (c) จะตัดข้อมูลที่อยู่ห่างไกลออก คือ **<u>ข้อมูลชุดที่ 2</u>**
ดังนั้นสามารถวิเคราะห์ใหม่ทั้งหมดได้ดังนี้

$\overline{x} = \begin{bmatrix} 0.7675 \\ 8.8688 \end{bmatrix}$

$S = \begin{bmatrix} 0.3786 & 1.0303 \\ 1.0303 & 69.86 \end{bmatrix}$

$\lambda_i = \begin{bmatrix} 69.875 \\ 0.3633 \end{bmatrix}$

$\lambda_1 = 69.875 \quad e_1 = \begin{bmatrix} 0.0148 \\ 0.9999 \end{bmatrix}$

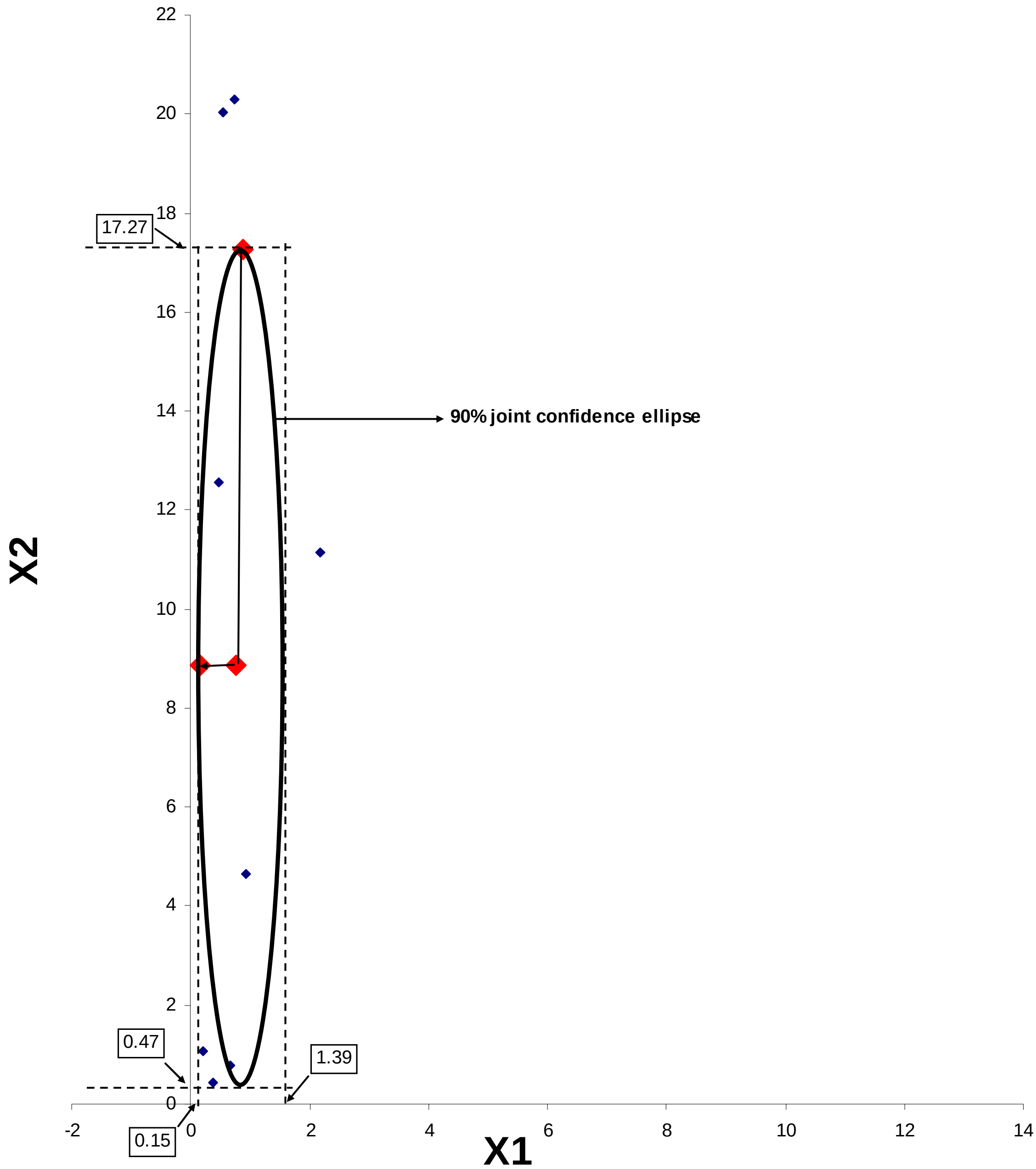
$\lambda_2 = 0.3633 \quad e_2 = \begin{bmatrix} -0.9999 \\ 0.0148 \end{bmatrix}$

$\pm\sqrt{\lambda_1}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)} =$ 8.3591 x 1.0046 = 8.3974

$\pm\sqrt{\lambda_1}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)}e_1 = \begin{bmatrix} 0.1243 \\ 8.3965 \end{bmatrix}$

$\pm\sqrt{\lambda_2}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)} =$ 0.6027 x 1.0046 = 0.6055

$\pm\sqrt{\lambda_2}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)}e_2 = \begin{bmatrix} -0.605 \\ 0.009 \end{bmatrix}$



$\boxed{0.15 \le \mu_1 \le 1.39}$ $\boxed{0.47 \le \mu_2 \le 17.27}$

Does it appear as if this Peruvian culture has a mean strontium level of 10?

Yes, Peruvian culture has a mean strontium level of 10.

Is $[.30\ ,10]'$ a plausible value for μ ? Discuss.

$$n(\overline{x}-\mu_o)'S^{-1}(\overline{x}-\mu_o) = 8\begin{bmatrix} 0.4675 & -1.131 \end{bmatrix}\begin{bmatrix} 2.752 & -0.041 \\ -0.041 & 0.0149 \end{bmatrix}\begin{bmatrix} 0.4675 \\ -1.131 \end{bmatrix}$$

$$8\begin{bmatrix} 1.3325 & -0.036 \end{bmatrix}\begin{bmatrix} 0.4675 \\ -1.131 \end{bmatrix}$$

$$8 \quad \text{x} \quad 0.6635$$

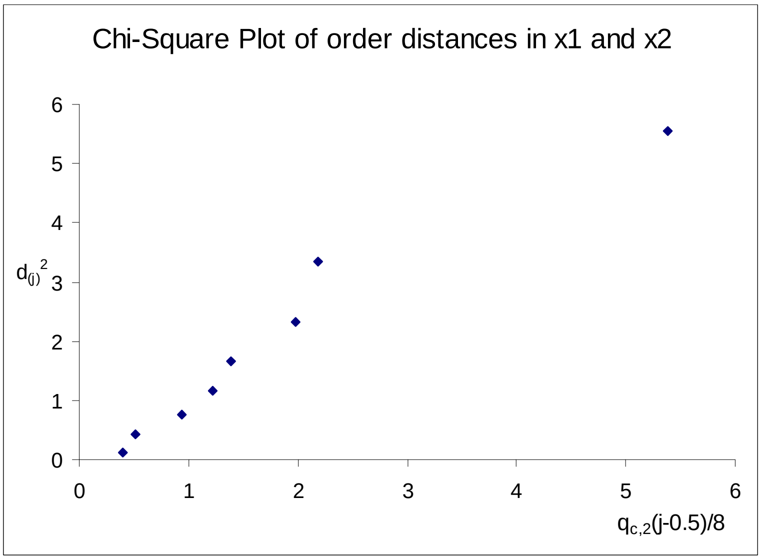
$$n(\overline{x}-\mu_o)'S^{-1}(\overline{x}-\mu_o) = 5.3079$$

$$\frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha) = 8.0733$$

We not reject H_o at the 10% level of significance. So, $\mu' = [.30\ ,10]'$ plausible value of mean chromium = .30 and plausible value of mean strontium = 10

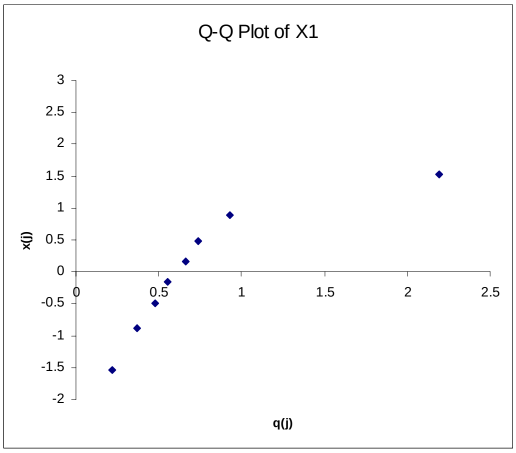
การตรวจสอบ bivariate normal อยู่ในบทที่ 4 หน้า 183 มีวิธีการทำดังนี้

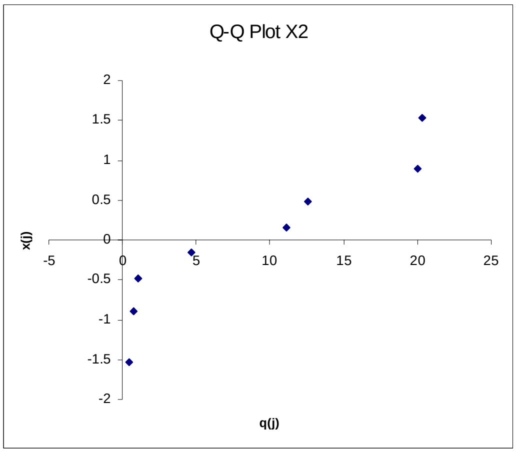
$(x-\overline{x})'S^{-1}(x-\overline{x}) \le \chi_2^2(.5)$ หรือ จัดทำ Chi-Square Plot

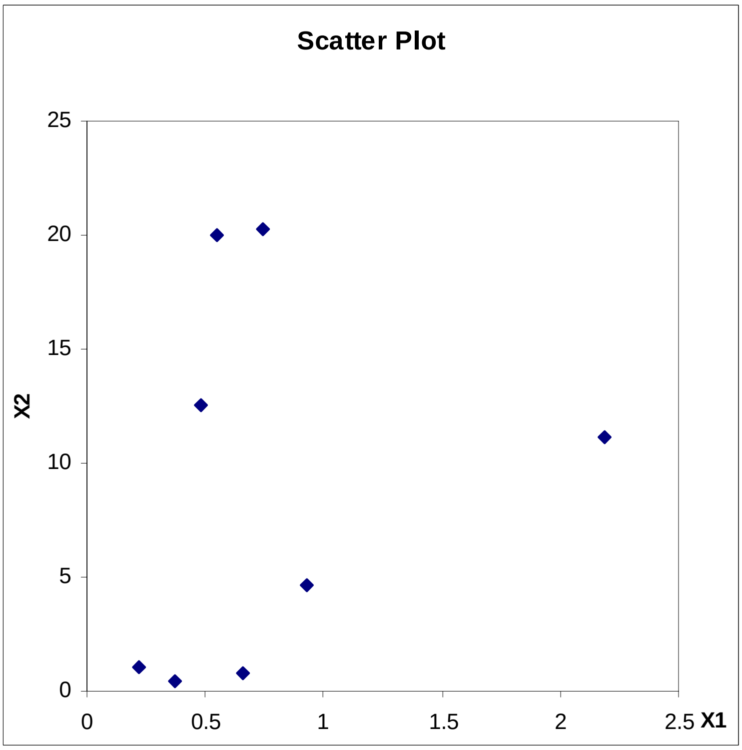


ข้อมูลชุดที่	คำนวณ		Chi-Square(0.5)
1	0.52	**<**	**1.39**
2	5.38	>	1.39
3	2.18	**>**	**1.39**
4	1.97	**>**	**1.39**
5	0.94	<	1.39
6	0.40	<	1.39
7	1.22	<	1.39
8	1.38	<	1.39

จากการทดสอบมี 3 จุดที่มากกว่า 1.39 หรือประมาณ 37.5% ซึ่งไม่เกิน 50% แสดงเป็น Bivariate normal







สรุปได้ว่าการตัดข้อมูลชุดที่ 2 ออกทำให้ตัวแปร x_1 และ ตัวแปร x_2 มีความสัมพันธ์กันน้อยลง (Correlation Coefficient = 0.2) และช่วงความเชื่อมั่นของทั้งสองตัวแปร แคบกว่าการไม่ตัดข้อมูลชุดที่ 2 ออก ส่วนการทดสอบสมมติฐานที่ $\mu' = [.30\ ,10]'$ ยังไม่สามารถปฏิเสธสมมติฐาน H_o ที่ระดับนัยสำคัญ 10% ได้ ทั้งนี้เนื่องจากตัวแปรมีความสัมพันธ์กันน้อย ดังนั้นการสร้างช่วงความเชื่อมั่น และการทดสอบสมมติฐานต่างๆ อาจจะสามารถแยกกันคำนวณได้โดยไม่ต้องพิจารณาแบบ Multivariate (เพื่อให้แน่ใจควรทดสอบความสัมพันธ์)

5.12. Given the data

$$X = \begin{bmatrix} 3 & 6 & 0 \\ 4 & 4 & 3 \\ - & 8 & 3 \\ 5 & - & - \end{bmatrix}$$

With missing components, use the prediction-estimation algorithm of Section 5.7 to estimate μ and Σ. Determine the initial estimates, and iterate to find the first revised estimates.

$\tilde{\mu}_1 = 4$

$\tilde{\mu}_2 = 6$

$\tilde{\mu}_3 = 2$

$$\tilde{\Sigma} = \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 2 & 0 \\ 0.5 & 0 & 1.5 \end{bmatrix}$$

The first component of x_3 is missing, so we partition $\tilde{\mu}$ and $\tilde{\Sigma}$ as

$$\tilde{\mu} = \begin{bmatrix} \tilde{\mu}_1 \\ \hline \tilde{\mu}_2 \\ \tilde{\mu}_3 \end{bmatrix} = \begin{bmatrix} \tilde{\mu}^{(1)} \\ \hline \tilde{\mu}^{(2)} \end{bmatrix}, \quad \tilde{\Sigma} = \left[\begin{array}{c|cc} \tilde{\sigma}_{11} & \tilde{\sigma}_{12} & \tilde{\sigma}_{13} \\ \hline \tilde{\sigma}_{12} & \tilde{\sigma}_{22} & \tilde{\sigma}_{23} \\ \tilde{\sigma}_{13} & \tilde{\sigma}_{23} & \tilde{\sigma}_{33} \end{array}\right] = \left[\begin{array}{c|c} \tilde{\Sigma}_{11} & \tilde{\Sigma}_{12} \\ \hline \tilde{\Sigma}_{21} & \tilde{\Sigma}_{22} \end{array}\right]$$

And predict

$$\tilde{x}_{31} = \tilde{\mu}_1 + \tilde{\Sigma}_{12}\tilde{\Sigma}_{22}^{-1} \begin{bmatrix} x_{32} - \tilde{\mu}_2 \\ x_{33} - \tilde{\mu}_3 \end{bmatrix}$$

$$4 + \begin{pmatrix} 0 & 0.5 \end{pmatrix} \begin{pmatrix} 0.5 & 0 \\ 0 & 0.666667 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$4 + \begin{pmatrix} 0.00 & 0.33 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$4 + 0.333333$$

$$= 4.33$$

$$\widetilde{x_{31}^2} = \tilde{\Sigma}_{11} - \tilde{\Sigma}_{12}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} + \tilde{x}_{31}^2$$

$$0.5 \quad - \quad \begin{pmatrix} 0 & 0.5 \end{pmatrix} \begin{pmatrix} 0.5 & 0 \\ 0 & 0.666667 \end{pmatrix} \begin{pmatrix} 0 \\ 0.5 \end{pmatrix} + \quad 18.78$$

$$0.5 \quad - \quad \begin{pmatrix} 0.00 & 0.33 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 0.5 \end{pmatrix} \quad + \quad 18.78$$

$$0.5 \quad - \quad 0.1667 \quad + \quad 18.78$$

$$= \quad 19.1111$$

$$\widetilde{x_{31}\begin{bmatrix} x_{32} & x_{33} \end{bmatrix}} = \tilde{x}_{31}\begin{bmatrix} x_{32} & x_{33} \end{bmatrix} = 4.33\begin{bmatrix} 8 & 3 \end{bmatrix} = \begin{bmatrix} 34.64 & 12.99 \end{bmatrix}$$

For the two missing component of x_4 , we partition $\tilde{\mu}$ and $\tilde{\Sigma}$ as

$$\tilde{\mu} = \begin{bmatrix} \tilde{\mu}_2 \\ \tilde{\mu}_3 \\ \hline \tilde{\mu}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mu}^{(1)} \\ \hline \tilde{\mu}^{(2)} \end{bmatrix}, \ \tilde{\Sigma} = \left[\begin{array}{cc|c} \tilde{\sigma}_{22} & \tilde{\sigma}_{23} & \tilde{\sigma}_{12} \\ \tilde{\sigma}_{23} & \tilde{\sigma}_{33} & \tilde{\sigma}_{13} \\ \hline \tilde{\sigma}_{12} & \tilde{\sigma}_{13} & \tilde{\sigma}_{11} \end{array}\right] = \left[\begin{array}{c|c} \tilde{\Sigma}_{11} & \tilde{\Sigma}_{12} \\ \hline \tilde{\Sigma}_{21} & \tilde{\Sigma}_{22} \end{array}\right]$$

And predict

$$\widetilde{\begin{bmatrix} x_{42} \\ x_{43} \end{bmatrix}} = \begin{bmatrix} \tilde{\mu}_2 \\ \tilde{\mu}_3 \end{bmatrix} + \tilde{\Sigma}_{12}\tilde{\Sigma}_{22}^{-1}\begin{bmatrix} x_{41} - \tilde{\mu}_1 \end{bmatrix}$$

$$\begin{pmatrix} 6 \\ 2 \end{pmatrix} \quad + \quad \begin{pmatrix} 0 \\ 0.5 \end{pmatrix} \quad 2 \quad\quad 1$$

$$\begin{pmatrix} 6 \\ 2 \end{pmatrix} \quad + \quad \begin{pmatrix} 0.00 \\ 1 \end{pmatrix}$$

$$= \quad \begin{pmatrix} 6.0 \\ 3.0 \end{pmatrix}$$

For the contribution to T_1 . Also, from (5-39)

$$\widetilde{\begin{bmatrix} x_{42}^2 & x_{42}x_{43} \\ x_{42}x_{43} & x_{43}^2 \end{bmatrix}} = \tilde{\Sigma}_{11} - \tilde{\Sigma}_{12}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21} + \widetilde{\begin{bmatrix} x_{42} \\ x_{43} \end{bmatrix}}\widetilde{\begin{bmatrix} x_{42} & x_{43} \end{bmatrix}}$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 1.5 \end{pmatrix} - \begin{pmatrix} 0 \\ 0.5 \end{pmatrix} 2 \begin{pmatrix} 0 & 0.5 \end{pmatrix} + \begin{pmatrix} 6.0 \\ 3.0 \end{pmatrix} \begin{pmatrix} 6.0 & 3.0 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 1.5 \end{pmatrix} - \begin{pmatrix} 0 \\ 0.5 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} + \begin{pmatrix} 36 & 18 \\ 18 & 9 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 1.5 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 0.5 \end{pmatrix} + \begin{pmatrix} 36 & 18 \\ 18 & 9 \end{pmatrix}$$

$$= \begin{pmatrix} 38 & 18 \\ 18 & 10 \end{pmatrix}$$

$$\widetilde{\begin{bmatrix} x_{42} \\ x_{43} \end{bmatrix} x_{41}} = \widetilde{\begin{bmatrix} x_{42} \\ x_{43} \end{bmatrix}} x_{41} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} (5) = \begin{bmatrix} 30 \\ 15 \end{bmatrix}$$

Are the contributions to T_2. Thus, the predicted complete-data sufficient statistics are

$$\tilde{T}_1 = \begin{bmatrix} x_{11} + x_{21} + \tilde{x}_{31} + x_{41} \\ x_{12} + x_{22} + x_{32} + \tilde{x}_{42} \\ x_{13} + x_{23} + x_{33} + \tilde{x}_{43} \end{bmatrix} = \begin{bmatrix} 3+4+4.33+5 \\ 6+4+8+6 \\ 0+3+3+3 \end{bmatrix} = \begin{bmatrix} 16.33 \\ 24 \\ 9 \end{bmatrix}$$

$$\tilde{T}_2 = \begin{bmatrix} 9+16+19.11+25=69.11 & & \\ 3(6)+4(4)+34.64+30=98.64 & 36+16+64+38=154 & \\ 3(0)+4(3)+12.99+15=39.99 & 6(0)+4(3)+8(3)+18=54 & 9+9+10=28 \end{bmatrix}$$

$$\tilde{T}_2 = \begin{bmatrix} 69.11 & 98.64 & 39.99 \\ 98.64 & 154 & 54 \\ 39.99 & 54 & 28 \end{bmatrix}$$

This completes one prediction step.

$$\tilde{\mu} = \frac{1}{n}\tilde{T}_1 = \begin{bmatrix} 4.0825 \\ 6 \\ 2.25 \end{bmatrix}$$

$$\tilde{\Sigma} = \frac{1}{n}\tilde{T}_2 - \tilde{\mu}\tilde{\mu}'$$
$$= \frac{1}{4}\begin{bmatrix} 69.11 & 98.64 & 39.99 \\ 98.64 & 154 & 54 \\ 39.99 & 54 & 28 \end{bmatrix} - \begin{bmatrix} 4.0825 \\ 6 \\ 2.25 \end{bmatrix}\begin{bmatrix} 4.0825 & 6 & 2.25 \end{bmatrix}$$

$$\tilde{\Sigma} = \begin{bmatrix} 0.610694 & 0.165 & 0.811875 \\ 0.165 & 2.5 & 0 \\ 0.811875 & 0 & 1.9375 \end{bmatrix}$$