STAT 524
HW4

Satoshi Ido
34788706
10/03/2023

5.2

5.2. Using the data in Example 5.1, verify that T^2 remains unchanged if each observation $\mathbf{x}_j, j = 1, 2, 3;$ is replaced by $\mathbf{C}\mathbf{x}_j$, where

$$\mathbf{C} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \qquad X = \begin{bmatrix} 6 & 9 \\ 10 & 6 \\ 8 & 3 \end{bmatrix}$$

Note that the observations

$$\mathbf{C}\mathbf{x}_j = \begin{bmatrix} x_{j1} - x_{j2} \\ x_{j1} + x_{j2} \end{bmatrix}$$

yield the data matrix

$$\begin{bmatrix} (6-9) & (10-6) & (8-3) \\ (6+9) & (10+6) & (8+3) \end{bmatrix}'$$

$$CX = \begin{bmatrix} -3 & 4 & 5 \\ 15 & 16 & 11 \end{bmatrix}^T = \begin{bmatrix} -3 & 15 \\ 4 & 16 \\ 5 & 11 \end{bmatrix}$$

$$\overline{CX} = \begin{bmatrix} \overline{CX_1} \\ \overline{CX_2} \end{bmatrix} = \begin{bmatrix} (-3+4+5)/3 \\ (15+16+11)/3 \end{bmatrix} = \begin{bmatrix} 2 \\ 14 \end{bmatrix}$$

$$C\mu_0 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 9 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 14 \end{bmatrix}$$

$$S_{11} = \frac{1}{2}\left\{(-3-2)^2 + (4-2)^2 + (5-2)^2\right\} = 19$$

$$S_{22} = \frac{1}{2}\left\{(15-14)^2 + (16-14)^2 + (11-14)^2\right\} = 7$$

$$S_{12} = S_{21} = \frac{1}{2}\left\{-5 \times 1 + 2 \times 2 + 3 \times (-3)\right\} = -5$$

$$\therefore S = \begin{bmatrix} 19 & -5 \\ -5 & 7 \end{bmatrix}$$

$$\Rightarrow S^{-1} = \frac{1}{19 \times 7 - (-5)^2} \begin{bmatrix} 7 & 5 \\ 5 & 19 \end{bmatrix} = \begin{bmatrix} 7/108 & 5/108 \\ 5/108 & 19/108 \end{bmatrix}$$

$$T^2 = (\overline{CX} - C\mu_0)^T \cdot n \cdot S^{-1} \cdot (\overline{CX} - C\mu_0)$$

$$= 3 \cdot \begin{bmatrix} \underset{-2}{2-4} & \underset{0}{14-14} \end{bmatrix} \cdot \begin{bmatrix} 7/108 & 5/108 \\ 5/108 & 19/108 \end{bmatrix} \cdot \begin{bmatrix} 2-4 \\ 14-14 \end{bmatrix}$$

$$= 3 \cdot \begin{bmatrix} -7/54 & -5/54 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

$$= 7/9 \qquad \therefore \; T^2 \text{ is unchanged. } //$$

5.9

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

Covariance matrix

$$\mathbf{S} = \begin{bmatrix} 3266.46 & 1343.97 & 731.54 & 1175.50 & 162.68 & 238.37 \\ 1343.97 & 721.91 & 324.25 & 537.35 & 80.17 & 117.73 \\ 731.54 & 324.25 & 179.28 & 281.17 & 39.15 & 56.80 \\ 1175.50 & 537.35 & 281.17 & 474.98 & 63.73 & 94.85 \\ 162.68 & 80.17 & 39.15 & 63.73 & 9.95 & 13.88 \\ 238.37 & 117.73 & 56.80 & 94.85 & 13.88 & 21.26 \end{bmatrix}$$

(a) Obtain the large sample 95% simultaneous confidence intervals for the six population mean body measurements.
(b) Obtain the large sample 95% simultaneous confidence ellipse for mean weight and mean girth.
(c) Obtain the 95% Bonferroni confidence intervals for the six means in Part a.
(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.
(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.

(a) Since $n-p$ is large enough, we use a formula,

$$\text{Simultaneous C.I.} = a^T \bar{x} \pm \sqrt{\chi_p^2(\alpha)} \sqrt{\frac{a^T S a}{n}}$$

$$95\% \text{ SCI} \iff \alpha = 0.05 \; , \quad \chi_6^2(0.05)$$

For $\bar{x}_1$,

$$SCI = x_1 \pm \sqrt{\chi_6^2(0.05)}\sqrt{\frac{S_{11}}{n}}$$

$$= 95.52 \pm \sqrt{12.592}\sqrt{\frac{3266.46}{61}} = [69.55,\ 121.49]$$

For $\bar{x}_2$,

$$SCI = 164.38 \pm \sqrt{12.592}\sqrt{\frac{721.91}{61}} = [152.17,\ 176.59]$$

For $\bar{x}_3$,

$$SCI = 55.69 \pm \sqrt{12.592}\sqrt{\frac{179.28}{61}} = [49.61,\ 61.77]$$

For $\bar{x}_4$,

$$SCI = 93.39 \pm \sqrt{12.592}\sqrt{\frac{474.98}{61}} = [83.49,\ 103.29]$$

For $\bar{x}_5$,

$$SCI = 17.98 \pm \sqrt{12.592}\sqrt{\frac{9.95}{61}} = [16.55,\ 19.41]$$

For $\bar{x}_6$,

$$SCI = 31.13 \pm \sqrt{12.592}\sqrt{\frac{21.26}{61}} = [29.04,\ 33.22]$$

(d) For large sample,

95% Simultaneous Confidence Ellipse for mean weight and mean girth can be calculated as:

$$61\begin{bmatrix}95.52-\mu_1, & 93.39-\mu_4\end{bmatrix}\begin{bmatrix}3266.46 & 1175.50\\ 1175.50 & 474.98\end{bmatrix}^{-1}\begin{bmatrix}95.52-\mu_1\\ 93.39-\mu_4\end{bmatrix} \le \chi_6^2(0.05)$$

$$\cong 61 \begin{bmatrix} 95.52-\mu_1, & 93.39-\mu_4 \end{bmatrix} \begin{bmatrix} 0.0028 & -0.0069 \\ -0.0069 & 0.0193 \end{bmatrix} \begin{bmatrix} 95.52-\mu_1 \\ 93.39-\mu_4 \end{bmatrix} \le 12.952$$

$$\cong \begin{bmatrix} 0.0028(95.52-\mu_1) & -0.0069(95.52-\mu_1) \\ -0.0069(93.39-\mu_4) & 0.0193(93.39-\mu_4) \end{bmatrix} \begin{bmatrix} 95.52-\mu_1 \\ 93.39-\mu_4 \end{bmatrix} \le 0.206$$

$$\cong \begin{bmatrix} 0.0028(95.52-\mu_1)^2 - 0.0069(95.52-\mu_1)(93.39-\mu_4) \\ -0.0069(93.39-\mu_4)(95.52-\mu_1) + 0.0193(93.39-\mu_4)^2 \end{bmatrix} \le 0.206$$

The left hand side is 95% confidence ellipse and the right hand side is C^2.

$$S_{14} = \begin{bmatrix} \sigma_{11} & \sigma_{41} \\ \sigma_{14} & \sigma_{44} \end{bmatrix} = \begin{bmatrix} 3266.46 & 1175.50 \\ 1175.50 & 474.98 \end{bmatrix}$$

$$|S_{14} - \lambda I| = \begin{vmatrix} 3266.46-\lambda & 1175.50 \\ 1175.50 & 474.98-\lambda \end{vmatrix}$$

$$= (3266.46-\lambda)(474.98-\lambda) - 1175.50^2$$

$$= \lambda^2 - 3741.44\lambda + 169702.92$$

$$|S_{14} - \lambda I| = 0$$

$$\therefore \lambda = -(-1870.72) \pm \sqrt{(-1870.72)^2 - 169702.92}$$

$$= 1870.72 \pm 1824.8$$

$$\therefore \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 3695.52 \\ 45.92 \end{bmatrix}$$

For $\lambda_1 = 3695.52$, $S_{14} e = \lambda_1 e$

$$\Rightarrow \begin{bmatrix} 3266.46 & 1175.50 \\ 1175.50 & 474.98 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = 3695.52 \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$\begin{cases} 3266.46 e_1 + 1175.50 e_2 = 3695.52 e_1 \\ 1175.50 e_1 + 474.98 e_2 = 3695.52 e_2 \end{cases}$$

$$\begin{cases} -429.06 e_1 + 1175.50 e_2 = 0 \\ 1175.50 e_1 - 3220.54 e_2 = 0 \end{cases}$$

$$\Rightarrow e_1 \cong 2.74 e_2 \quad \therefore e = \begin{bmatrix} 2.74 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.365 \end{bmatrix}$$

For $\lambda_2 = 45.92$, $S_{14} e = \lambda_2 e$

$$\Rightarrow \begin{bmatrix} 3266.46 & 1175.50 \\ 1175.50 & 474.98 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = 45.92 \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$\begin{cases} 3266.46 e_1 + 1175.50 e_2 = 45.92 e_1 \\ 1175.50 e_1 + 474.98 e_2 = 45.92 e_2 \end{cases}$$

$$\begin{cases} 3220.54 e_1 + 1175.50 e_2 = 0 \\ 1175.50 e_1 + 429.06 e_2 = 0 \end{cases}$$

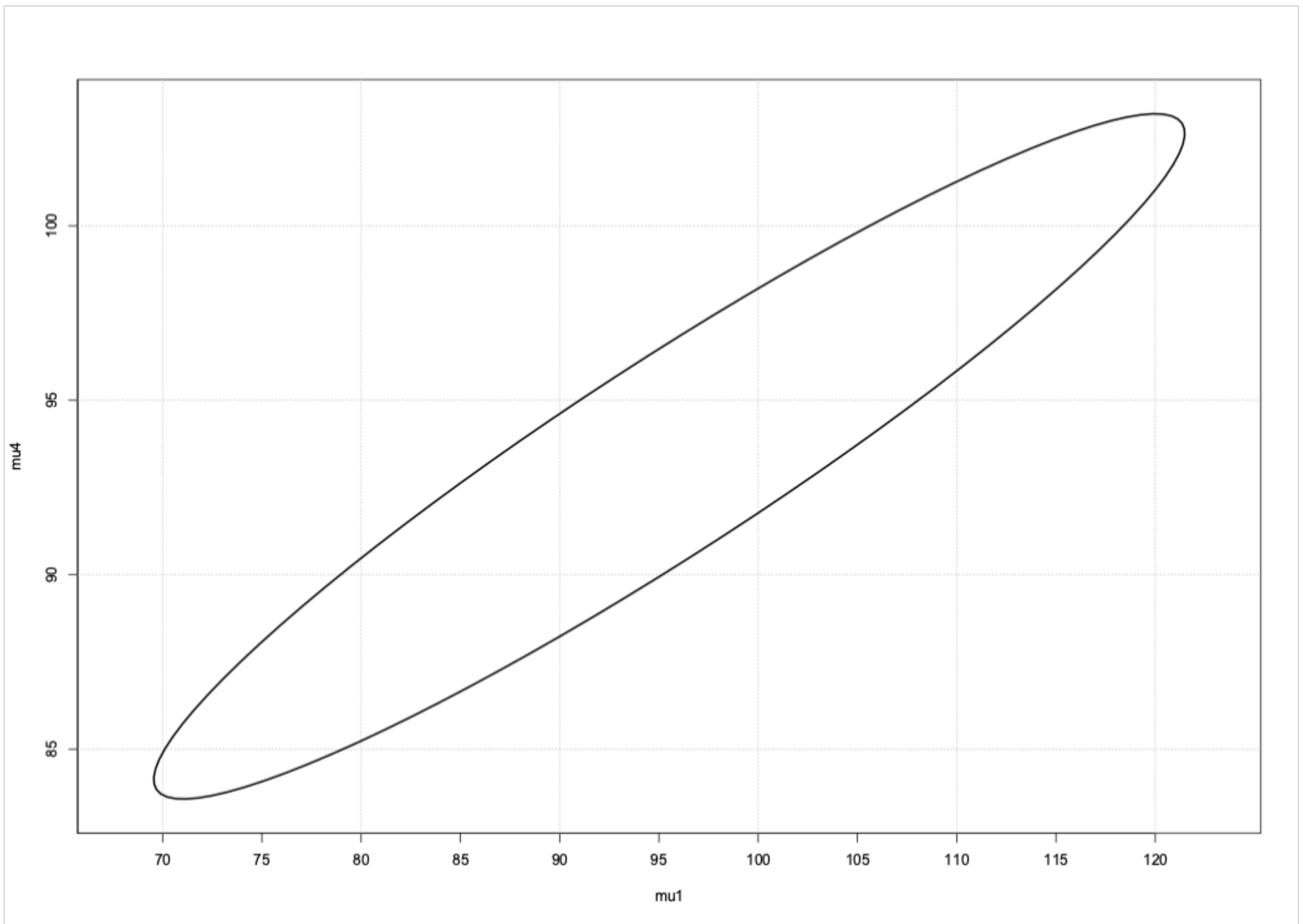
$$\Rightarrow e_1 \cong -0.365 e_2 \quad \therefore e = \begin{bmatrix} -0.365 \\ 1 \end{bmatrix}$$

$\therefore$ The half length of the major axis is given by

$\sqrt{\lambda_1}\, c \cong 60.790 \cdot 0.4539 \cong 27.59$ with a direction of $\begin{bmatrix} 1 \\ 0.365 \end{bmatrix}$

The half length of the minor axis is given by

$\sqrt{\lambda_2}\, c \cong 6.776 \cdot 0.4539 \cong 3.076$ with a direction of $\begin{bmatrix} -0.365 \\ 1 \end{bmatrix}$



(c) 95% Simultaneous Bonferroni Confidence Interval is given by:

$$\bar{x}_i \pm t_{n-1}\left(\frac{\alpha}{2\cdot p}\right)\sqrt{\frac{S_{ii}}{n}} \qquad i = 1, \dots, 6,\ n = 61,\ p = 6$$

For x_1,

$$95.52 \pm 2.73 \cdot \sqrt{\frac{3266.46}{61}} = [75.56,\ 115.49]$$ //

For x_2,

$$164.38 \pm 2.73 \cdot \sqrt{\frac{721.91}{61}} = [154.99,\ 173.77]$$ //

For x_3,

$$55.69 \pm 2.73 \cdot \sqrt{\frac{179.28}{61}} = [51.01,\ 60.37]$$ //

For x_4,

$$93.39 \pm 2.73 \cdot \sqrt{\frac{474.98}{61}} = [85.73,\ 101.00]$$ //

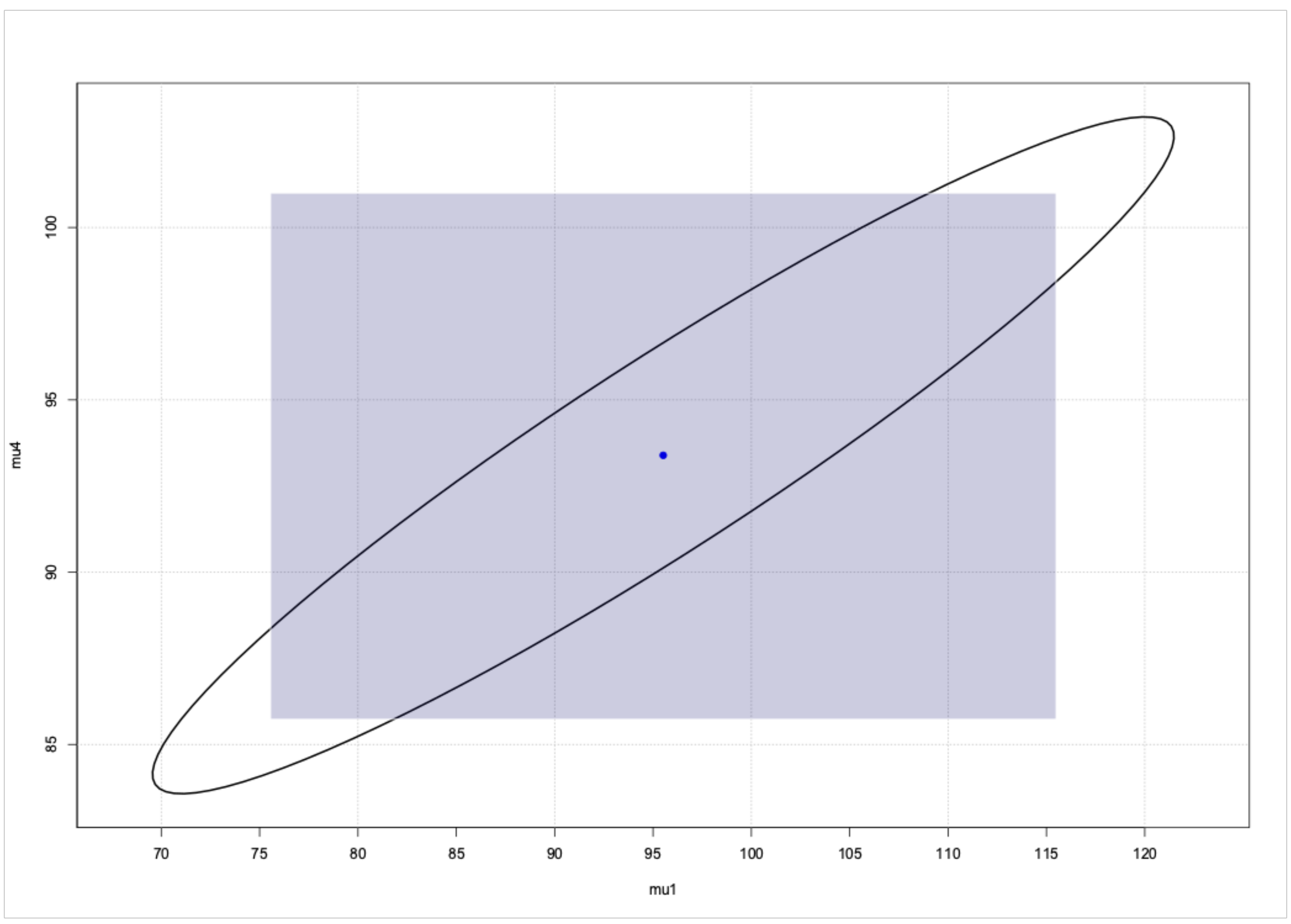
For x_5,

$$17.98 \pm 2.73 \cdot \sqrt{\frac{9.95}{61}} = [16.88, 19.08]$$ //

For x_6,

$$31.13 \pm 2.73 \sqrt{\frac{21.26}{61}} = [29.52, 32.74]$$ //

(d) As we can see, the Bonferroni Simultaneous Confidence Rectangle is shorter both in x-axis and y-axis when $m = p = 6$ since the method is more conservative and provide shorter range.



(e) $\bar{x}_i \pm t_{n-1}\left(\frac{\alpha}{2 \cdot p}\right) \sqrt{\frac{\Sigma_{6-5}}{n}}$

$\bar{x}_{1-4}$ = (mean head width − mean head length) = 13.15

$\Sigma_{6-5} = Var(\bar{x}_6 - \bar{x}_5) = \dfrac{\Sigma_{55} + \Sigma_{66} - 2 \cdot \Sigma_{56}}{n = 61}$

$n = 61, \quad p = 7 \quad \Rightarrow t_{60}\left(\frac{0.05}{2\cdot 7}\right) = 2.79$

$\Rightarrow$ Bonferroni SCI $= 13.15 \pm 2.79 \cdot \sqrt{\frac{3.45}{61}}$

$= 13.15 \pm 0.66$

$= [12.49, \ 13.81]$ //

5.12 **5.12.** Given the data

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

with missing components, use the prediction–estimation algorithm of Section 5.7 to estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Determine the initial estimates, and iterate to find the *first* revised estimates.

$\tilde{\mu}_1 = 4, \quad \tilde{\mu}_2 = 6, \quad \tilde{\mu}_3 = 2$

The first revised estimates are

$$\tilde{\mu} = \frac{1}{n} \tilde{T}_1 = \begin{bmatrix} 4.08 \\ 6.00 \\ 2.25 \end{bmatrix} \text{//}$$

$$\tilde{\Sigma} = \begin{bmatrix} 0.60 & & \\ 0.17 & 2.50 & \\ 0.81 & 0.00 & 1.94 \end{bmatrix} \text{//}$$

5.15

5.15. Let X_{ji} and X_{jk} be the ith and kth components, respectively, of $\mathbf{X}_j$.

(a) Show that $\mu_i = E(X_{ji}) = p_i$ and $\sigma_{ii} = \operatorname{Var}(X_{ji}) = p_i(1 - p_i)$, $i = 1, 2, \ldots, p$.

(b) Show that $\sigma_{ik} = \operatorname{Cov}(X_{ji}, X_{jk}) = -p_i p_k$, $i \neq k$. Why must this covariance necessarily be negative?

(a)
$$x = \begin{cases} 1 & \text{if attribute present} \\ 0 & \text{if attribute absent} \end{cases}$$

$$E(x_{ji}) = 1 \times p_i + 0(1 - p_i) = p_i \quad //$$

$$\operatorname{Var}(x_{ji}) = (1 - p_i)^2 \cdot p_i + (0 - p_i)^2 \cdot (1 - p_i)$$

$$= p_i(1 - p_i) \quad //$$

(b) $\operatorname{Cov}(x_{ji}, x_{jk})$

$$= E(x_{ji}, x_{jk}) - E(x_{ji}) E(x_{jk})$$

$\because$ all the elements of x are iid.

$\therefore E(x_{ji}, x_{jk})$

$$= 0 - E(x_{ji}) E(x_{jk})$$

$$= -p_i p_k \quad \text{for } i \neq k$$

Since there is no correlation between the variables and the probabilities are positive //

6.1

6.1. Construct and sketch a joint 95% confidence region for the mean difference vector $\boldsymbol{\delta}$ using the effluent data and results in Example 6.1. Note that the point $\boldsymbol{\delta} = \mathbf{0}$ falls outside the 95% contour. Is this result consistent with the test of H_0: $\boldsymbol{\delta} = \mathbf{0}$ considered in Example 6.1? Explain.

$$\bar{d} = \begin{bmatrix} \bar{d}_1 \\ \bar{d}_2 \end{bmatrix} = \begin{bmatrix} -9.36 \\ 13.27 \end{bmatrix}$$

The confidence ellipse is centered at $[-9.36, 13.27]$

$$\lambda \text{ for } S_d = \begin{bmatrix} 449.79 \\ 165.09 \end{bmatrix}$$

$$e \text{ for } S_d = \begin{bmatrix} e_1^T \\ e_2^T \end{bmatrix} = \begin{bmatrix} 0.33, & 0.94 \\ -0.94, & 0.33 \end{bmatrix}$$

$\therefore$ The half length of the major axis is given by

$$\sqrt{\lambda_1}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)} = 19.65 \quad \text{with direction } \pm[0.33, 0.94]$$

The half length of the minor axis is given by

$$\sqrt{\lambda_1}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)} = 12.03 \quad \text{with direction } \pm[-0.94, 0.33]$$

The 95% Confidence Region for $\bar{d}$ is given by

$$n(\bar{d}-\delta)^2 S_d^{-1}(\bar{d}-\delta) \le \frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)$$

$$= 11\left[-9.36-\delta_1,\ 13.27-\delta_2\right]\begin{bmatrix} 0.0055 & -0.0012 \\ -0.0012 & 0.0026 \end{bmatrix}\begin{bmatrix} -9.36-\delta_1 \\ 13.27-\delta_2 \end{bmatrix}$$

$$\le \frac{20}{9}F_{2,9}(0.05) = 0.86$$

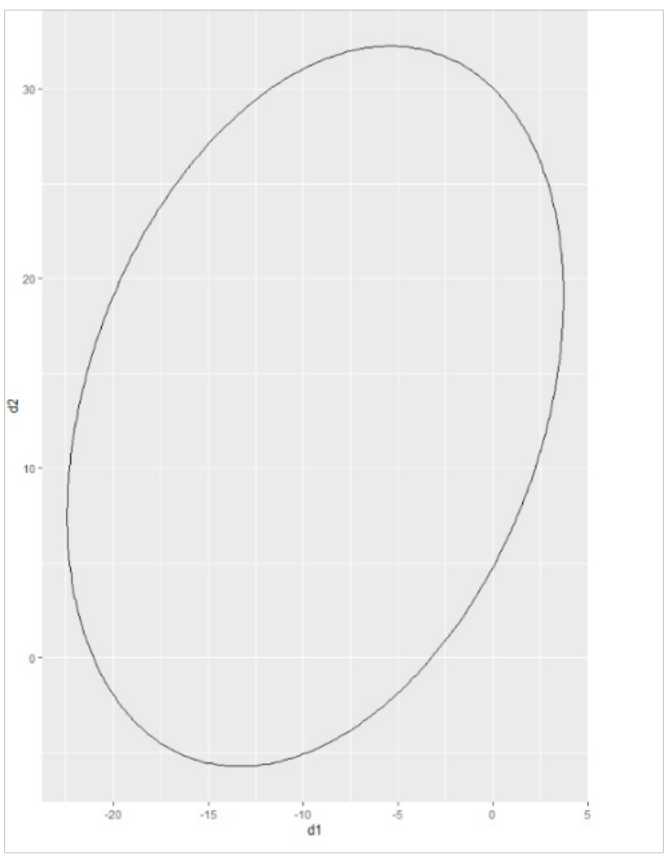
The left hand side

$$= \left[(-9.36)^2 (0.0055) + 2(-9.36)(13.27)(-0.0012) + (13.27)^2 (0.0026) \right]$$

$= 1.26$ which is larger than 0.46.

$\therefore$ $\delta = 0$ doesn't fall inside the confidence region.

$\Rightarrow$ H_0 is rejected, and the result is consistent with exp 6.1 //



6.6

6.6. Use the data for treatments 2 and 3 in Exercise 6.8.

(a) Calculate $\mathbf{S}_{\text{pooled}}$.

(b) Test H_0: $\boldsymbol{\mu}_2 - \boldsymbol{\mu}_3 = \mathbf{0}$ employing a two-sample approach with $\alpha = .01$.

(c) Construct 99% simultaneous confidence intervals for the differences $\mu_{2i} - \mu_{3i}$, $i = 1, 2$.

(a) For treatment 2,

$$\bar{x}_2 = \begin{bmatrix} (3+1+2)/3 \\ (3+6+3)/3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$S_{2\,11} = \frac{(3-2)^2 + (1-2)^2 + (2-2)^2}{2} = 1$$

$$S_{2\,22} = \frac{(3-4)^2+(6-4)^2+(3-4)^2}{2} = 3$$

$$S_{2\,12} = \frac{(3-2)(3-4)+(1-2)(6-4)+0}{2} = \frac{3}{2}$$

$$\therefore S_2 = \begin{bmatrix} 1 & 3/2 \\ 3/2 & 3 \end{bmatrix}$$

For treatment 3,

$$\bar{x}_3 = \begin{bmatrix} (2+5+3+2)/4 \\ (3+1+1+3)/4 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$S_{3\,11} = \frac{(2-3)^2+(5-3)^2+(3-3)^0+(2-3)^2}{3} = 2$$

$$S_{3\,22} = \frac{(3-2)^2+(1-2)^2+(1-2)^2+(3-2)^2}{3} = \frac{4}{3}$$

$$S_{3\,12} = \frac{(2-3)(3-2)+(5-3)(1-2)+0+(2-3)(3-2)}{3} = \frac{-4}{3}$$

$$\therefore S_3 = \begin{bmatrix} 2 & -4/3 \\ -4/3 & 4/3 \end{bmatrix}$$

$$\therefore S_{pooled} = \frac{n_2-1}{n_2+n_3-2} S_2 + \frac{n_3-1}{n_2+n_3-2} S_3$$

$$= \frac{3-1}{3+4-2} \cdot \begin{bmatrix} 1 & 3/2 \\ 3/2 & 3 \end{bmatrix} + \frac{4-1}{3+4-2} \begin{bmatrix} 2 & -4/3 \\ -4/3 & 4/3 \end{bmatrix}$$

$$= \begin{bmatrix} 1.6 & -1.4 \\ -1.4 & 2 \end{bmatrix}$$

//

(b) $H_0 : \mu_2 - \mu_3 \neq 0$

$$T^2 = (\bar{x}_2 - \bar{x}_3 - \delta_0)^T \left[\left(\frac{1}{n_2} + \frac{1}{n_3} \right) S_{pooled} \right]^{-1} (\bar{x}_2 - \bar{x}_3 - \delta_0)$$

$$= \begin{bmatrix} 2-3-0 \\ 4-2-0 \end{bmatrix} \left[\left(\frac{1}{3} + \frac{1}{4} \right) \begin{bmatrix} 1.6 & -1.4 \\ -1.4 & 2 \end{bmatrix} \right]^{-1} \begin{bmatrix} 2-3-0 \\ 4-2-0 \end{bmatrix}$$

$$= 3.87$$

$$c^2 = \frac{(n_2+n_3-2)p}{n_2+n_3-p-1} F_{p,\, n_2+n_3-p-1}(\alpha)$$

$$= \frac{(3+4-2) \times 2}{3+4-2-1} F_{2,4}(0.01) = 45$$

$\therefore T^2 < c^2 \Rightarrow$ we are not able to reject H_0 //

(c) 99% CI of $\mu_{21} - \mu_{31}$ is given by

$$(\bar{x}_{21} - \bar{x}_{31}) \pm \sqrt{\frac{(n_2+n_3-2)p}{n_2+n_3-p-1} F_{p,\, n_2+n_3-p-1}(\alpha)} \sqrt{\left(\frac{1}{n_2} + \frac{1}{n_3} \right) S_{11,\, pooled}}$$

$$= (2-3) \pm \sqrt{45} \cdot \sqrt{\frac{7}{12} \times 1.6}$$

$$= [-7.48,\ 5.48] \ //$$

99% CI of $\mu_{22} - \mu_{32}$ is given by

$$(\bar{x}_{22} - \bar{x}_{32}) \pm \sqrt{\frac{(n_2+n_3-2)p}{n_2+n_3-p-1} F_{p,\, n_2+n_3-p-1}(\alpha)} \sqrt{\left(\frac{1}{n_2} + \frac{1}{n_3} \right) S_{22,\, pooled}}$$

$$= (4-2) \pm \sqrt{45} \cdot \sqrt{\frac{7}{12} \cdot 2}$$

$$= [-5.25,\ 9.25] \quad //$$