Multivariate Statistical Analysis Mid Term 2008
Reference Solution

1. (6%)

Answer:

$$CI_{95}:\left(\bar{x}-t_{n-1}(0.025)\frac{s}{\sqrt{n}},\bar{x}+t_{n-1}(0.025)\frac{s}{\sqrt{n}}\right)$$

i.e., $\left(56-t_{15}(0.025)\frac{15}{\sqrt{16}},56+t_{15}(0.025)\frac{15}{\sqrt{16}}\right)$

$t_{15}(0.025)\approx 2.131$

Thus, $CI_{95}:(48.01,63.99)$, in cm.

2. (6%)

Answer:

In all samples of size 16 from all collegiate basketball players in Taiwan if they take *MagicPill*, 95% of the intervals determined by CI_{95} computed in the same way as in Problem 1 will include the actual average height jumped by all collegiate basketball players in Taiwan if they take *MagicPill.*

3. (6%)

Answer:

Power = 90% = $1-\beta$. Thus the type II error $\beta=0.1$

For critical value CV to achieve type I error $\alpha=0.05$, and type II error $\beta=0.1$, we have $\dfrac{CV-50}{15/\sqrt{n}}=z(0.025)=1.96=z_\alpha$, $\dfrac{56-CV}{15/\sqrt{n}}=z(0.1)=1.28=z_\beta$

Thus, $n=\left(\dfrac{\sigma(z_\alpha+z_\beta)}{\Delta}\right)^2=\left(\dfrac{15*(1.96+1.28)}{56-50}\right)^2\approx 65.61$. Take $n=66$.

4. (6%)

Answer:

The sample correlation coefficient $r_{ik}=\dfrac{s_{ik}}{\sqrt{s_{ii}}\sqrt{s_{kk}}}$ can be regarded as cosine of the angle formed by the deviation vectors $\mathbf{d}_i=\mathbf{y}_i-\bar{x}_i\mathbf{1}$ and $\mathbf{d}_k=\mathbf{y}_k-\bar{x}_k\mathbf{1}$ in the n-space..

5. (4%)

Answer:

The generalized sample variance $|\mathbf{S}| = (n-1)^{-p}(volume)^2$, where *volume* is the volume generated in *n*-space by the *p* deviation vectors $\mathbf{d}_1 = \mathbf{y}_1 - \bar{x}_1\mathbf{1}$, $\mathbf{d}_2 = \mathbf{y}_2 - \bar{x}_2\mathbf{1}, \ldots, \mathbf{d}_p = \mathbf{y}_p - \bar{x}_p\mathbf{1}$.

6. (6%)

Answer:

By spectral decomposition, $\mathbf{\Sigma} = \lambda_1\mathbf{e}_1\mathbf{e}_1' + \lambda_2\mathbf{e}_2\mathbf{e}_2'$, and by Result 4.1 in Textbook, $\mathbf{\Sigma}^{-1} = \frac{1}{\lambda_1}\mathbf{e}_1\mathbf{e}_1' + \frac{1}{\lambda_2}\mathbf{e}_2\mathbf{e}_2'$. Thus, the equation $(\mathbf{x}-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = c^2$ can be expressed as $(\mathbf{x}-\boldsymbol{\mu})'\left(\frac{1}{\lambda_1}\mathbf{e}_1\mathbf{e}_1' + \frac{1}{\lambda_2}\mathbf{e}_2\mathbf{e}_2'\right)(\mathbf{x}-\boldsymbol{\mu}) = c^2$, with $\mathbf{y} = [y_1 \quad y_2]' = \begin{bmatrix} \mathbf{e}_1' \\ \mathbf{e}_2' \end{bmatrix}(\mathbf{x}-\boldsymbol{\mu})$, the ellipse equation may be written as $\frac{1}{\lambda_1}\left(\mathbf{e}_1'(\mathbf{x}-\boldsymbol{\mu})\mathbf{e}_1\right)'\left(\mathbf{e}_1'(\mathbf{x}-\boldsymbol{\mu})\mathbf{e}_1\right) + \frac{1}{\lambda_1}\left(\mathbf{e}_2'(\mathbf{x}-\boldsymbol{\mu})\mathbf{e}_2\right)'\left(\mathbf{e}_2'(\mathbf{x}-\boldsymbol{\mu})\mathbf{e}_2\right) = c^2$, i.e., $\frac{y_1^2}{c^2\lambda_1} + \frac{y_2^2}{c^2\lambda_2} = 1$.

7. (4%)

Answer:

$$\begin{bmatrix} 12 & 6 \\ 6 & 7 \end{bmatrix}\begin{bmatrix} 3/\sqrt{13} \\ 2/\sqrt{13} \end{bmatrix} = \begin{bmatrix} 48/\sqrt{13} \\ 32/\sqrt{13} \end{bmatrix} = 16\begin{bmatrix} 3/\sqrt{13} \\ 2/\sqrt{13} \end{bmatrix}, \text{ i.e., } \begin{bmatrix} 12 & 6 \\ 6 & 7 \end{bmatrix}\mathbf{e}_1 = \lambda_1\mathbf{e}_1$$

$$\begin{bmatrix} 12 & 6 \\ 6 & 7 \end{bmatrix}\begin{bmatrix} -2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix} = \begin{bmatrix} -6/\sqrt{13} \\ 9/\sqrt{13} \end{bmatrix} = 3\begin{bmatrix} -2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}, \text{ i.e., } \begin{bmatrix} 12 & 6 \\ 6 & 7 \end{bmatrix}\mathbf{e}_2 = \lambda_2\mathbf{e}_2$$

8. . (6%)

Answer:

Based on Eq. (5 - 6) or (5 - 7),

$$T^2 = n(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0) = 30[0.5 \quad -0.5]\frac{1}{48}\begin{bmatrix} 7 & -6 \\ -6 & 12 \end{bmatrix}\begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}$$

$$= \frac{5}{8}[0.5 \quad -0.5]\begin{bmatrix} 6.5 \\ -9 \end{bmatrix} = 7.75 \times \frac{5}{8} \approx 4.8438$$

$$\text{Critical value} = \frac{(n-1)p}{n-p}F_{p,n-p}(0.1) = \frac{29 \times 2}{28}F_{2,28}(0.1) = \frac{58}{28} \times 2.50 = 5.17857$$

$T^2 <$ critical value. $\therefore$ can not reject H_0

9. (6%)

Answer:

From Eq. (5-18),

$$n(\bar{\mathbf{x}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}) \le \frac{p(n-1)}{n-p}F_{p,n-p}(0.1) = 5.17857$$

$$(\bar{\mathbf{x}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}) \le \frac{5.17857}{30} = 0.172619$$

By Eq. (5-19), major axes $\sqrt{\lambda_i}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(0.1)}$ and directions $\mathbf{e}_i$

namely, major axes $4\sqrt{0.172619} \approx 1.66$ and $\sqrt{3}\sqrt{0.172619} \approx 0.72$.

directions $\frac{1}{\sqrt{13}}\begin{bmatrix} 3 \\ 2 \end{bmatrix} \approx \begin{bmatrix} 0.83 \\ 0.55 \end{bmatrix}$ and $\frac{1}{\sqrt{13}}\begin{bmatrix} -2 \\ 3 \end{bmatrix} \approx \begin{bmatrix} -0.55 \\ 0.83 \end{bmatrix}$

10. (6%)

Answer: In all samples of size 30 drawn from the population, 90% of the Hotelling's T^2 confidence region determined in the same way as in Problem 7 will include the population mean vector $\boldsymbol{\mu}$.

11. (6%)

Answer:

Use Eq. (5.24),

$$\bar{x}_1 - \sqrt{\frac{p(n-1)}{n-p}F_{p,n-p}(0.1)}\sqrt{\frac{s_{11}}{n}} \le \mu_1 \le \bar{x}_1 + \sqrt{\frac{p(n-1)}{n-p}F_{p,n-p}(0.1)}\sqrt{\frac{s_{11}}{n}}$$

$$\bar{x}_2 - \sqrt{\frac{p(n-1)}{n-p}F_{p,n-p}(0.1)}\sqrt{\frac{s_{22}}{n}} \le \mu_2 \le \bar{x}_2 + \sqrt{\frac{p(n-1)}{n-p}F_{p,n-p}(0.1)}\sqrt{\frac{s_{22}}{n}}$$

$$\sqrt{\frac{p(n-1)}{n-p}F_{p,n-p}(0.1)} = \sqrt{5.17857} = 2.27565$$

$$-0.5 - \sqrt{5.17857}\sqrt{\frac{12}{30}} \le \mu_1 \le -0.5 + \sqrt{5.17857}\sqrt{\frac{12}{30}}, \quad \mu_1 \in (-1.94, 0.94)$$

$$0.5 - \sqrt{5.17857}\sqrt{\frac{7}{30}} \le \mu_2 \le 0.5 + \sqrt{5.17857}\sqrt{\frac{7}{30}}, \quad \mu_2 \in (-0.60, 1.60)$$

12. . (6%)

Answer:

In all samples of size 30 drawn from the population, 90% of the simultatneous T^2 confidence intervals determined in the same way as in Problem 9 will include the population mean vector $\boldsymbol{\mu}$.

13. (6%)

Answer:

From Eq. (5 - 29),

$$\bar{x}_1 - t_{n-1}(\frac{0.1}{2p})\sqrt{\frac{s_{11}}{n}} \le \mu_1 \le \bar{x}_1 + t_{n-1}(\frac{0.1}{2p})\sqrt{\frac{s_{11}}{n}}$$

$$\bar{x}_2 - t_{n-1}(\frac{0.1}{2p})\sqrt{\frac{s_{22}}{n}} \le \mu_2 \le \bar{x}_2 + t_{n-1}(\frac{0.1}{2p})\sqrt{\frac{s_{22}}{n}}$$

$$t_{29}(0.025) = 2.045$$

$$-0.5 - 2.045\sqrt{\frac{12}{30}} \le \mu_1 \le -0.5 + 2.045\sqrt{\frac{12}{30}}, \quad \mu_1 \in (-1.79, 0.79)$$

$$0.5 - 2.045\sqrt{\frac{7}{30}} \le \mu_2 \le 0.5 + 2.045\sqrt{\frac{7}{30}}, \quad \mu_2 \in (-0.49, 1.49)$$

14. (6%)

Answer:

$$\text{Likelihood} = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x_j-\mu)^2/(2\sigma^2)} = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\sum_{j=1}^{n}(x_j-\mu)^2/2\sigma^2} = L(\mu,\sigma^2)$$

Since the μ and σ^2 which maximize $L(\mu,\sigma^2)$ are the same as those maximize, $l(\mu,\sigma^2) = \ln L(\mu,\sigma^2) = -\sum_{j=1}^{n}(x_j-\mu)^2/2\sigma^2 - \frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2$, we set

$$\frac{\partial l(\mu,\sigma^2)}{\partial\mu} = \frac{1}{\sigma^2}\sum_{j=1}^{n}(x_j-\mu) = 0 \quad , \quad \text{and} \quad \frac{\partial l(\mu,\sigma^2)}{\partial\sigma^2} = \frac{1}{2\sigma^4}\sum_{j=1}^{n}(x_j-\mu)^2 - \frac{n}{2}\frac{1}{\sigma^2} = 0 \quad .$$

From the first equation, we have the maximum likelihood estimate of μ:

$\hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} x_j = \bar{x}$. Substitute it into the second equation, we obtain the maximum likelihood estimate of σ^2 : $\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j-\bar{x})^2$ .The corresponding maximum likelihood is thus $\max_{\mu,\sigma^2} L(\mu,\sigma^2) = \frac{1}{(2\pi)^{n/2}\hat{\sigma}^n} e^{-n/2}$

15. . (6%)

Answer:

$$\text{Likelihood} = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x_j-\mu_0)^2/(2\sigma^2)} = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\sum_{j=1}^{n}(x_j-\mu_0)^2/2\sigma^2} = L(\sigma^2)$$

Since the σ^2 which maximize $L(\sigma^2)$ are the same as those maximize, $l(\sigma^2) = \ln L(\sigma^2) = -\sum_{j=1}^{n}(x_j-\mu_0)^2/2\sigma^2 - \frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2$, we set

$\frac{\partial l(\sigma^2)}{\partial\sigma^2} = \frac{1}{2\sigma^4}\sum_{j=1}^{n}(x_j-\mu_0)^2 - \frac{n}{2}\frac{1}{\sigma^2} = 0$. From this equation, we can obtain the maximum likelihood estimate of σ^2 : $\hat{\sigma}_0^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j-\mu_0)^2$ .The corresponding maximum likelihood is thus $\max_{\sigma^2} L(\sigma^2) = \frac{1}{(2\pi)^{n/2}\hat{\sigma}_0^n} e^{-n/2}$

16. (6%)

Answer:

$$\Lambda = \frac{\max_{\sigma^2} L(\sigma^2)}{\max_{\mu,\sigma^2} L(\mu,\sigma^2)} = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2}. \quad \text{Thus,} \quad \Lambda^{2/n} = \frac{\hat{\sigma}^2}{\hat{\sigma}_0^2} = \frac{\sum_{j=1}^{n}(x_j - \bar{x})^2}{\sum_{j=1}^{n}(x_j - \mu_0)^2}$$

$$\Lambda^{2/n} = \frac{\sum_{j=1}^{n}(x_j - \bar{x})^2}{\sum_{j=1}^{n}(x_j - \mu_0)^2} = \frac{\sum_{j=1}^{n}(x_j - \bar{x})^2}{\sum_{j=1}^{n}(x_j - \bar{x} + \bar{x} - \mu_0)^2}.$$ Since

$$\sum_{j=1}^{n}(x_j - \bar{x} + \bar{x} - \mu_0)^2 = \sum_{j=1}^{n}\left[(x_j - \bar{x})^2 + 2(x_j - \bar{x})(\bar{x} - \mu_0) + (\bar{x} - \mu_0)^2\right]$$

$$= \sum_{j=1}^{n}(x_j - \bar{x})^2 + 2(\bar{x} - \mu_0)\sum_{j=1}^{n}(x_j - \bar{x}) + \sum_{j=1}^{n}(\bar{x} - \mu_0)^2$$, because

$$= \sum_{j=1}^{n}(x_j - \bar{x})^2 + n(\bar{x} - \mu_0)^2$$

$\sum_{j=1}^{n}(x_j - \bar{x}) = \sum_{j=1}^{n} x_j - n\bar{x} = 0$. Thus,

$$\Lambda^{2/n} = \frac{1}{1 + \frac{n(\bar{x} - \mu_0)^2}{\sum_{j=1}^{n}(x_j - \bar{x})^2}} = \left(1 + \frac{n(\bar{x} - \mu_0)^2}{(n-1)s^2}\right)^{-1} = \left(1 + \frac{t^2}{n-1}\right)^{-1}, \quad t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

17. (6%)

Answer:

Since $\Lambda = \frac{\max_{\sigma^2} L(\sigma^2)}{\max_{\mu,\sigma^2} L(\mu,\sigma^2)} \le 1$, and if H_0 holds, Λ will be close to 1. Thus if $\Lambda \le c_\alpha$, we may reject H_0 at significance level α. From Problem 16, we have

$$\Lambda^{2/n} = \left(1 + \frac{t^2}{n-1}\right)^{-1}, \text{ i.e., } \left(1 + \frac{t^2}{n-1}\right)^{-1} \le c_\alpha^{2/n}, \text{ or } 1 + \frac{t^2}{n-1} \ge c_\alpha^{-2/n}.$$

Thus, $t^2 \ge (n-1)\left(c_\alpha^{-2/n} - 1\right)$ is required to reject H_0 at significance level α. This is achieved by choosing $\sqrt{(n-1)\left(c_\alpha^{-2/n}-1\right)} = t_{n-1}\left(\frac{\alpha}{2}\right)$, or $c_\alpha = \left(1+\frac{t_{n-1}^2(\alpha/2)}{n-1}\right)^{-n/2}$.