

Lecture 4 Inferences About Mean

Shiwei Lan¹

¹School of Mathematical and Statistical Sciences
Arizona State University

STP533 Multivariate Analysis
Spring 2025

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

- 1 Hypothesis Test of Normal Population Mean
 - T-Test of Univariate Normal Population Mean
 - Hotelling's T^2 of Multivariate Normal Population Mean
 - Hotelling's T^2 as Likelihood Ratio Test
- 2 Confidence Regions of Mean Vector
 - Confidence Regions
 - Simultaneous Comparisons of Means
 - The Bonferroni Method of Multiple Comparisons
- 3 Large Sample Inference about a Population Mean Vector
 - Large Sample Inference of Mean Vector

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

- In this lecture, we concern about the inference of a mean vector.
- Let us start with the one-sample t -test for a univariate normal population and consider the following hypothesis:

$$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0$$

- Suppose we collect a random sample $\{X_i\}_{i=1}^n$ from the normal population with mean μ . Then the test statistic is

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}, \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

- This test statistic follow t -distribution with degree of freedom $(n-1)$ under the null hypothesis, i.e. $t \stackrel{H_0}{\sim} t(n-1)$.

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

- Reject H_0 when $|t| > t_{1-\alpha/2}(n-1)$ at the confidence level of $\alpha 100\%$.
- This is equivalent to considering the F test statistic

$$F = t^2 = n(\bar{X} - \mu_0)s^{-2}(\bar{X} - \mu_0) \sim F(1, n-1)$$

- Note the region of rejecting H_0 is

$$\mu_0 \in (-\infty, \bar{x} - t_{1-\alpha/2}(n-1)s/\sqrt{n}) \cup (\bar{x} + t_{1-\alpha/2}(n-1)s/\sqrt{n}, +\infty)$$

- Or equivalently, the $(1 - \alpha)100\%$ confidence interval for μ is

$$\mu \in [\bar{x} - t_{1-\alpha/2}(n-1)s/\sqrt{n}, \bar{x} + t_{1-\alpha/2}(n-1)s/\sqrt{n}]$$

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

- Now we generalize the above F test statistic to multivariate case.
- Recall sample mean $\bar{\mathbf{X}} = \frac{1}{n} \sum \mathbf{X}_i$ and sample covariance $\mathbf{S} = \frac{1}{n-1} \sum (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T$.
- When the sample $\mathbf{X}_{n \times p}$ is taken from multivariate normal, i.e. $\mathbf{X}_i \stackrel{iid}{\sim} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we have

$$\bar{\mathbf{X}} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n), \quad (n-1)\mathbf{S} \sim W_{n-1}(\boldsymbol{\Sigma}), \quad n(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim \chi_p^2.$$

- The quadratic form $T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)$ is called *Hotelling's T^2 statistic* which follows a F -distribution

$$T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0) \stackrel{H_0}{\sim} \frac{(n-1)p}{n-p} F(p, n-p)$$

- In general, $T^2(p, n-1) = N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})^T [W_{p, n-1}(\boldsymbol{\Sigma})/(n-1)]^{-1} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Example 5.2 (Testing a multivariate mean vector with T^2) Perspiration from 20 healthy females was analyzed. Three components, X_1 = sweat rate, X_2 = sodium content, and X_3 = potassium content, were measured, and the results, which we call the *sweat data*, are presented in Table 5.1.

Test the hypothesis $H_0: \mu' = [4, 50, 10]$ against $H_1: \mu' \neq [4, 50, 10]$ at level of significance $\alpha = .10$.

Computer calculations provide

$$\bar{\mathbf{x}} = \begin{bmatrix} 4.640 \\ 45.400 \\ 9.965 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 2.879 & 10.010 & -1.810 \\ 10.010 & 199.788 & -5.640 \\ -1.810 & -5.640 & 3.628 \end{bmatrix}$$

and

$$\mathbf{S}^{-1} = \begin{bmatrix} .586 & -.022 & .258 \\ -.022 & .006 & -.002 \\ .258 & -.002 & .402 \end{bmatrix}$$

We evaluate

$T^2 =$

$$\begin{aligned} 20 [4.640 - 4, 45.400 - 50, 9.965 - 10] & \begin{bmatrix} .586 & -.022 & .258 \\ -.022 & .006 & -.002 \\ .258 & -.002 & .402 \end{bmatrix} \begin{bmatrix} 4.640 - 4 \\ 45.400 - 50 \\ 9.965 - 10 \end{bmatrix} \\ & = 20 [.640, -4.600, -.035] \begin{bmatrix} .467 \\ -.042 \\ .160 \end{bmatrix} = 9.74 \end{aligned}$$

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate
Normal Population
Mean

Hotelling's T^2 of
Multivariate Normal
Population Mean

Hotelling's T^2 as
Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous
Comparisons of Means

The Bonferroni Method
of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference
of Mean Vector

- Hotelling's T^2 statistics is invariant to affine transformation. That is, if $\mathbf{Y}_{p \times 1} = \mathbf{C}_{p \times p} \mathbf{X}_{p \times 1} + \mathbf{d}_{p \times 1}$ with \mathbf{C} nondegenerate, then

$$T_{\mathbf{Y}}^2 = T_{\mathbf{X}}^2$$

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

- The above Hotelling's T^2 test statistic can also be derived from likelihood ratio test (LRT),
- Recall the MLE $\hat{\mu} = \bar{\mathbf{X}}$, $\hat{\Sigma} = \frac{n-1}{n}\mathbf{S}$ of an MVN $N_p(\mu, \Sigma)$ is the maximum of

$$\max_{\mu, \Sigma} L(\mu, \Sigma) = (2\pi)^{np/2} |\hat{\Sigma}|^{-n/2} e^{-np/2}$$

where $L(\mu, \Sigma) = (2\pi)^{np/2} |\Sigma|^{-n/2} \exp\{-\frac{1}{2}\text{tr}[(\mathbf{X} - \mu)\Sigma^{-1}(\mathbf{X} - \mu)^T]\}$.

- By similar argument of MLE for $\hat{\Sigma}$, we have

$$\max_{\Sigma} L(\mu_0, \Sigma) = (2\pi)^{np/2} |\hat{\Sigma}_0|^{-n/2} e^{-np/2}$$

where $\hat{\Sigma}_0 = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \mu_0)(\mathbf{x}_i - \mu_0)^T = \frac{1}{n}(\mathbf{X} - \mu_0)^T(\mathbf{X} - \mu_0)$.

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

- Now we consider the following statistic for LRT $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$:

$$\Lambda = \frac{\max_{\Sigma} L(\mu_0, \Sigma)}{\max_{\mu, \Sigma} L(\mu, \Sigma)} = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} \right)^{n/2}$$

- The statistic $\Lambda^{2/n}$ is called *Wilk's lambda*.
- Based on the MVN assumption, i.e. $\mathbf{X}_i \stackrel{iid}{\sim} N_p(\mu, \Sigma)$, we have

$$\Lambda^{2/n} = \left[1 + \frac{T^2}{n-1} \right]^{-1}$$

- Hint: consider the determinant of $\mathbf{A} = \begin{bmatrix} (n-1)\mathbf{S} & \sqrt{n}(\bar{\mathbf{X}} - \mu_0) \\ \sqrt{n}(\bar{\mathbf{X}} - \mu_0)^T & -1 \end{bmatrix}$.

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

- 1 Hypothesis Test of Normal Population Mean
 - T-Test of Univariate Normal Population Mean
 - Hotelling's T^2 of Multivariate Normal Population Mean
 - Hotelling's T^2 as Likelihood Ratio Test
- 2 Confidence Regions of Mean Vector
 - Confidence Regions
 - Simultaneous Comparisons of Means
 - The Bonferroni Method of Multiple Comparisons
- 3 Large Sample Inference about a Population Mean Vector
 - Large Sample Inference of Mean Vector

- We extend the concept of univariate *confidence interval* to a multivariate *confidence region*.
- Let $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ be a vector of unknown population parameter. A *confidence region (CR)* of $\boldsymbol{\theta}$ based on sample \mathbf{X} at $100(1 - \alpha)\%$ confidence level, denoted as $R(\mathbf{X})$, is defined as

$$\Pr[\boldsymbol{\theta} \in R(\mathbf{X})] = 1 - \alpha$$

- Recall Hotelling's $T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim \frac{(n-1)p}{n-p} F(p, n-p)$.
- Therefore CR for $\boldsymbol{\mu}$ is computed based on

$$\Pr \left[T^2 \leq \frac{(n-1)p}{n-p} F_{1-\alpha}(p, n-p) \right] = 1 - \alpha$$

- The CR of MVN mean vector μ is determined by

$$(\bar{\mathbf{X}} - \mu)^T \mathbf{S}^{-1} (\bar{\mathbf{X}} - \mu) = c^2 = \frac{(n-1)p}{n(n-p)} F_{1-\alpha}(p, n-p)$$

- This is an ellipsoid centered at $\bar{\mathbf{X}}$ and having axes $\pm \sqrt{\lambda_i} \mathbf{c} \mathbf{v}_i$ with eigen-pairs $\{\lambda_i, \mathbf{v}_i\}$ of \mathbf{S} .

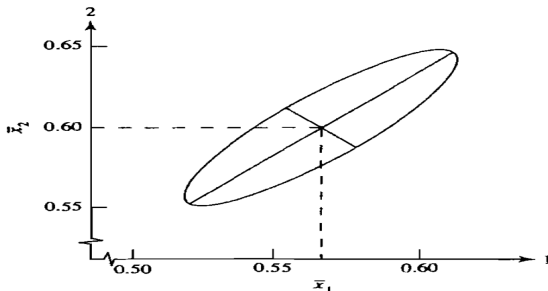


Figure 5.1 A 95% confidence ellipse for μ based on microwave-radiation data.

- CR is a joint statement of all the plausible values of population parameter, e.g. mean μ .
- Often we are concerned about separate confidence statements holding simultaneously, i.e. *simultaneous confidence intervals (SCI)*.
- We consider a linear combination of random vector $X \sim N_p(\mu, \Sigma)$: $Z = \mathbf{a}^T X \sim N(\mathbf{a}^T \mu, \mathbf{a}^T \Sigma \mathbf{a})$.
- Given a random sample $\mathbf{X}_{n \times p}$, we have corresponding sample $\mathbf{Z} = \mathbf{X}\mathbf{a}$, and hence $\bar{\mathbf{Z}} = \bar{\mathbf{X}}\mathbf{a}$ and $s_Z^2 = \mathbf{a}^T \mathbf{S} \mathbf{a}$.
- Therefore, the $100(1 - \alpha)\%$ CI for $\mu_Z = \mathbf{a}^T \mu$ can be obtained based on $|t| = \left| \frac{\bar{\mathbf{Z}} - \mu_Z}{s_Z / \sqrt{n}} \right| \leq t_{1-\alpha/2}(n-1)$:

$$\bar{\mathbf{X}}\mathbf{a} - t_{1-\alpha/2}(n-1)\sqrt{\mathbf{a}^T \mathbf{S} \mathbf{a}/n} \leq \mu_Z \leq \bar{\mathbf{X}}\mathbf{a} + t_{1-\alpha/2}(n-1)\sqrt{\mathbf{a}^T \mathbf{S} \mathbf{a}/n} \quad (*)$$

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

- Note, the CI for the component mean, e.g. μ_j can be obtained by setting $\mathbf{a} = \mathbf{e}_j = \underbrace{[0, \dots, 0]_{j-1}}_{j-1}, 1, \underbrace{[0, \dots, 0]_{n-j}}_{n-j}^T$.
- However, the confidence associated with all of the statements taken together is not $1 - \alpha$.
- It would be desirable to associate a 'collective' confidence coefficient of $1 - \alpha$ with the CIs generated by any \mathbf{a} . For this purpose, we consider

$$\max_{\mathbf{a}} t^2 = \max_{\mathbf{a}} \frac{n(\mathbf{a}^T(\bar{\mathbf{X}} - \boldsymbol{\mu}))^2}{\mathbf{a}^T \mathbf{S} \mathbf{a}} = n(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) = T^2$$

- Therefore, SCI can be obtained based on the previous Hotelling's test statistics $T^2 \sim \frac{(n-1)p}{n-p} F(p, n-p)$.

- For $j = 1, \dots, p$, successfully choose $\mathbf{a} = \mathbf{e}_j$ to obtain CI for μ_j simultaneously:

$$\bar{\mathbf{X}}_1 - \sqrt{\frac{(n-1)p}{n(n-p)} F_{1-\alpha}(p, n-p) s_1} \leq \mu_1 \leq \bar{\mathbf{X}}_1 + \sqrt{\frac{(n-1)p}{n(n-p)} F_{1-\alpha}(p, n-p) s_1}$$

$\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots$

$$\bar{\mathbf{X}}_j - \sqrt{\frac{(n-1)p}{n(n-p)} F_{1-\alpha}(p, n-p) s_j} \leq \mu_j \leq \bar{\mathbf{X}}_j + \sqrt{\frac{(n-1)p}{n(n-p)} F_{1-\alpha}(p, n-p) s_j}$$

$\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots$

$$\bar{\mathbf{X}}_p - \sqrt{\frac{(n-1)p}{n(n-p)} F_{1-\alpha}(p, n-p) s_p} \leq \mu_p \leq \bar{\mathbf{X}}_p + \sqrt{\frac{(n-1)p}{n(n-p)} F_{1-\alpha}(p, n-p) s_p}$$

Hypothesis Test
of Normal
Population
Mean

T-Test of Univariate
Normal Population
Mean

Hotelling's T^2 of
Multivariate Normal
Population Mean

Hotelling's T^2 as
Likelihood Ratio Test

Confidence
Regions of
Mean Vector

Confidence Regions

Simultaneous
Comparisons of Means

The Bonferroni Method
of Multiple Comparisons

Large Sample
Inference about
a Population
Mean Vector

Large Sample Inference
of Mean Vector

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

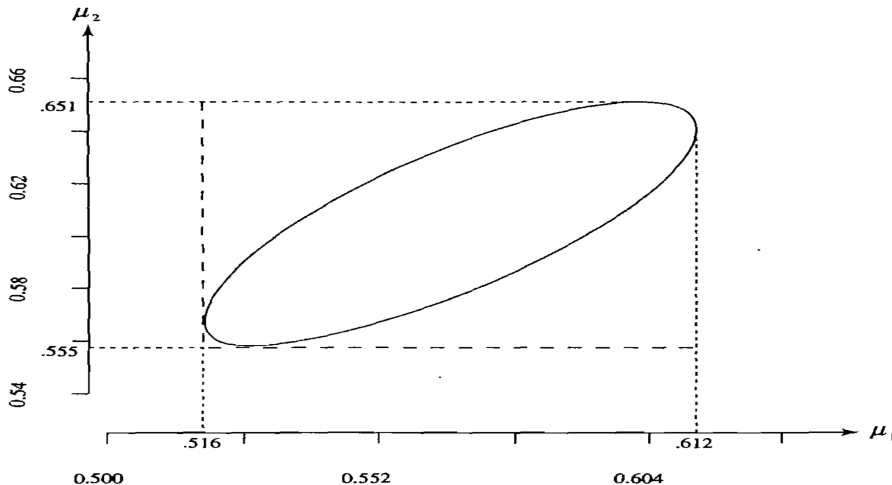


Figure 5.2 Simultaneous T^2 -intervals for the component means as shadows of the confidence ellipse on the axes—microwave radiation data.

- For $j = 1, \dots, p$, successfully choose $\mathbf{a} = \mathbf{e}_j$ in (*) to obtain CI for μ_j one at a time:

$$\bar{\mathbf{X}}_1 - t_{1-\alpha/2}(n-1)s_1/\sqrt{n} \leq \mu_1 \leq \bar{\mathbf{X}}_1 + t_{1-\alpha/2}(n-1)s_1/\sqrt{n}$$

$$\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots$$

$$\bar{\mathbf{X}}_j - t_{1-\alpha/2}(n-1)s_j/\sqrt{n} \leq \mu_j \leq \bar{\mathbf{X}}_j + t_{1-\alpha/2}(n-1)s_j/\sqrt{n}$$

$$\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots$$

$$\bar{\mathbf{X}}_p - t_{1-\alpha/2}(n-1)s_p/\sqrt{n} \leq \mu_p \leq \bar{\mathbf{X}}_p + t_{1-\alpha/2}(n-1)s_p/\sqrt{n}$$

- What is the issue?

- For $j = 1, \dots, p$, successfully choose $\mathbf{a} = \mathbf{e}_j$ in (*) to obtain CI for μ_j one at a time:

$$\bar{\mathbf{X}}_1 - t_{1-\alpha/2}(n-1)s_1/\sqrt{n} \leq \mu_1 \leq \bar{\mathbf{X}}_1 + t_{1-\alpha/2}(n-1)s_1/\sqrt{n}$$

$$\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots$$

$$\bar{\mathbf{X}}_j - t_{1-\alpha/2}(n-1)s_j/\sqrt{n} \leq \mu_j \leq \bar{\mathbf{X}}_j + t_{1-\alpha/2}(n-1)s_j/\sqrt{n}$$

$$\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots$$

$$\bar{\mathbf{X}}_p - t_{1-\alpha/2}(n-1)s_p/\sqrt{n} \leq \mu_p \leq \bar{\mathbf{X}}_p + t_{1-\alpha/2}(n-1)s_p/\sqrt{n}$$

- What is the issue?
- The probability of them holding simultaneously

$$\Pr[\bar{\mathbf{X}}_j - t_{1-\alpha/2}(n-1)s_j/\sqrt{n} \leq \mu_j \leq \bar{\mathbf{X}}_j + t_{1-\alpha/2}(n-1)s_j/\sqrt{n}, 1 \leq j \leq p] = (1-\alpha)^p$$

- Often, we are concerned about a limited number, m , of linear combinations of means, i.e. $\mathbf{a}_1^T \boldsymbol{\mu}, \dots, \mathbf{a}_m^T \boldsymbol{\mu}$.
- Let C_i be the $100(1 - \alpha_i)\%$ CI for $\mathbf{a}_i^T \boldsymbol{\mu}$, i.e. $\Pr[\mathbf{a}_i^T \boldsymbol{\mu} \in C_i] = 1 - \alpha_i$. Then we have

$$\begin{aligned} \Pr[\mathbf{a}_i^T \boldsymbol{\mu} \in C_i, j = 1, \dots, p] &= 1 - \Pr[\exists i_0, \text{ s.t. } \mathbf{a}_{i_0}^T \boldsymbol{\mu} \notin C_{i_0}] \\ &\geq 1 - \sum_{i=1}^m \Pr[\mathbf{a}_i^T \boldsymbol{\mu} \notin C_i] = 1 - \sum_{i=1}^m \alpha_i \end{aligned}$$

- Specifically, setting $\alpha_i = \frac{\alpha}{m}$ for $i = 1, \dots, m$ with $m = p$ we get the SCI for means with confidence level (at least) $1 - \alpha$:

$$\bar{\mathbf{X}}_j - t_{1-\alpha/(2p)}(n-1)s_j/\sqrt{n} \leq \mu_j \leq \bar{\mathbf{X}}_j + t_{1-\alpha/(2p)}(n-1)s_j/\sqrt{n}, \quad j = 1, \dots, p$$

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

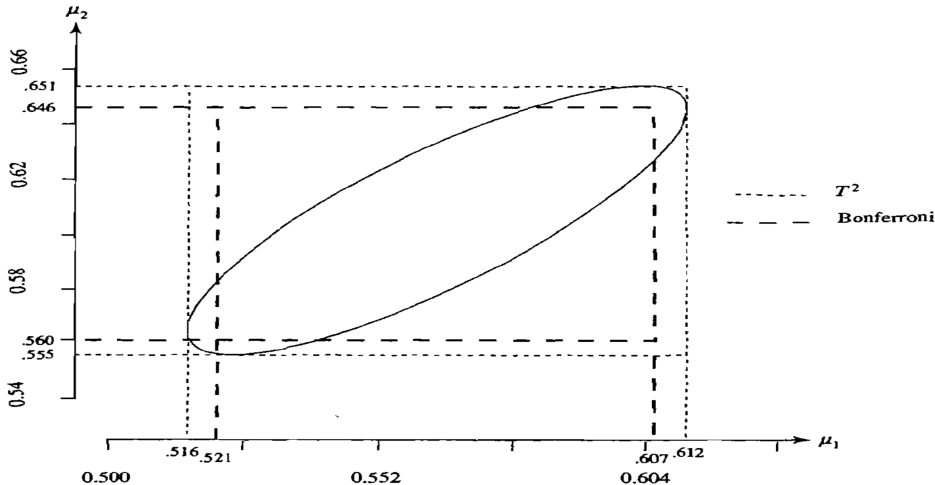


Figure 5.4 The 95% T^2 and 95% Bonferroni simultaneous confidence intervals for the component means—microwave radiation data.

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate Normal Population Mean

Hotelling's T^2 of Multivariate Normal Population Mean

Hotelling's T^2 as Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous Comparisons of Means

The Bonferroni Method of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference of Mean Vector

- 1 Hypothesis Test of Normal Population Mean
 - T-Test of Univariate Normal Population Mean
 - Hotelling's T^2 of Multivariate Normal Population Mean
 - Hotelling's T^2 as Likelihood Ratio Test
- 2 Confidence Regions of Mean Vector
 - Confidence Regions
 - Simultaneous Comparisons of Means
 - The Bonferroni Method of Multiple Comparisons
- 3 Large Sample Inference about a Population Mean Vector
 - Large Sample Inference of Mean Vector

- Recall that for a sample $\mathbf{X}_i \stackrel{iid}{\sim} (\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we have for large n

$$n(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim \chi_p^2.$$

- We can consider large sample inference of mean vector $\boldsymbol{\mu}$ regardless of the original distribution.
- Consider the hypothesis test $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ against $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$ and reject H_0 at the level of significance of α if

$$n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0) > \chi_{1-\alpha}^2(p)$$

- Alternatively, we can consider the (approximate) $100(1 - \alpha)\%$ CR of $\boldsymbol{\mu}$ based on

$$\Pr[n(\bar{\mathbf{X}} - \boldsymbol{\mu})^T \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \leq \chi_{1-\alpha}^2(p)] = 1 - \alpha$$

Lecture 4

S.Lan

Hypothesis Test of Normal Population Mean

T-Test of Univariate
Normal Population
Mean

Hotelling's T^2 of
Multivariate Normal
Population Mean

Hotelling's T^2 as
Likelihood Ratio Test

Confidence Regions of Mean Vector

Confidence Regions

Simultaneous
Comparisons of Means

The Bonferroni Method
of Multiple Comparisons

Large Sample Inference about a Population Mean Vector

Large Sample Inference
of Mean Vector

- For $\mathbf{X}_i \stackrel{iid}{\sim} (\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we can also consider the (approximate) SCI for $\mu_Z = \mathbf{a}^T \boldsymbol{\mu}$ when $n - p$ is large

$$\bar{\mathbf{X}}\mathbf{a} - \sqrt{\chi_{1-\alpha}^2(p)}\sqrt{\mathbf{a}^T \mathbf{S}\mathbf{a}/n} \leq \mu_Z \leq \bar{\mathbf{X}}\mathbf{a} + \sqrt{\chi_{1-\alpha}^2(p)}\sqrt{\mathbf{a}^T \mathbf{S}\mathbf{a}/n} \quad (*)$$

- We can consider the one-at-a-time confidence interval when n is large

$$\bar{\mathbf{X}}_j - z_{1-\alpha/2}^2 \sqrt{s_{jj}/n} \leq \mu_Z \leq \bar{\mathbf{X}}_j + z_{1-\alpha/2}^2 \sqrt{s_{jj}/n}, \quad j = 1, \dots, p$$

- The Bonferroni SCI is

$$\bar{\mathbf{X}}_j - z_{1-\alpha/(2p)}^2 \sqrt{s_{jj}/n} \leq \mu_Z \leq \bar{\mathbf{X}}_j + z_{1-\alpha/(2p)}^2 \sqrt{s_{jj}/n}, \quad j = 1, \dots, p$$