Lecture 4 Inferences About Mean

Shiwei Lan[1]

[1]School of Mathematical and Statistical Sciences
Arizona State University

STP533 Multivariate Analysis
Spring 2025

Table of Contents

T-Test of Univariate Normal Population Mean

- 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

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

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

F-Test of Univariate Normal Population Mean

- 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

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

Multivariate Generalization of F Test Statistic

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

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

- The 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

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

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

Hotelling's T^2 Test

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 $H_0: \boldsymbol{\mu}' = [4, 50, 10]$ against $H_1: \boldsymbol{\mu}' \neq [4, 50, 10]$ at level of significance $\alpha = .10$.

Computer calculations provide

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

and

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

We evaluate

$$T^2 = 20[4.640 - 4, \ 45.400 - 50, \ 9.965 - 10] \begin{bmatrix} .586 & -.022 & .258 \\ -.022 & .006 & -.002 \\ .258 & -.002 & .402 \end{bmatrix} \begin{bmatrix} 4.640 - 4 \\ 45.400 - 50 \\ 9.965 - 10 \end{bmatrix}$$

$$= 20[.640, \ -4.600, \ -.035] \begin{bmatrix} .467 \\ -.042 \\ .160 \end{bmatrix} = 9.74$$

Hotelling's T^2 Statistic

- 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

$$T^2_{\mathbf{Y}} = T^2_{\mathbf{X}}$$

Likelihood Ratio Test

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

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

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

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

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

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

Hotelling's T^2 Test as LRT

- Now we consider the following statistic for LRT $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ against $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$:

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

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

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

- Hint: consider the determinant of $\mathbf{A} = \begin{bmatrix} (n-1)\mathbf{S} & \sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) \\ \sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T & -1 \end{bmatrix}$.

Table of Contents

Confidence Regions

- 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

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

Confidence Region of Population Mean Vector

- The CR of MVN mean vector $\boldsymbol{\mu}$ is determined by

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

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

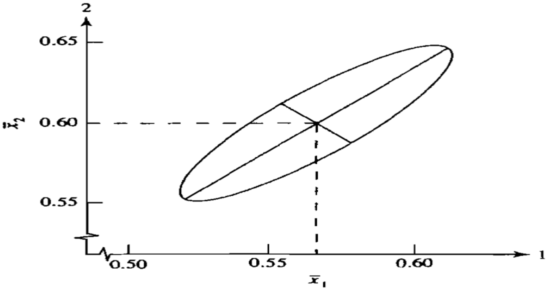

Figure 5.1 A 95% confidence ellipse for μ based on microwave-radiation data.

Example 5.3 (Constructing a confidence ellipse for μ) Data for radiation from microwave ovens were introduced in Examples 4.10 and 4.17. Let

$$x_1 = \sqrt[4]{\text{measured radiation with door closed}}$$

and

$$x_2 = \sqrt[4]{\text{measured radiation with door open}}$$

Simultaneous Confidence Intervals

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

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

Simultaneous Confidence Intervals

- Note, the CI for the component mean, e.g. μ_j can be obtained by setting $\mathbf{a} = \mathbf{e}_j = [\underbrace{0, \cdots, 0}_{j-1}, 1, \underbrace{0, \cdots, 0}_{p-j}]^T$.
- 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 $T^2 \sim \frac{(n-1)p}{n-p}F(p, n-p)$.

Simultaneous T^2 Confidence Intervals

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

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

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

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

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

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

Simultaneous CIs As Projections of CR

Example 5.4 (Simultaneous confidence intervals as shadows of the confidence ellipsoid) In Example 5.3, we obtained the 95% confidence ellipse for the means of the fourth roots of the door-closed and door-open microwave radiation measurements. The 95% simultaneous T^2 intervals for the two component means are, from (5-24),

$$\left(\bar{x}_1 - \sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(.05)} \sqrt{\frac{s_{11}}{n}},\quad \bar{x}_1 + \sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(.05)} \sqrt{\frac{s_{11}}{n}}\right)$$

$$= \left(.564 - \sqrt{\frac{2(41)}{40} 3.23} \sqrt{\frac{.0144}{42}},\quad .564 + \sqrt{\frac{2(41)}{40} 3.23} \sqrt{\frac{.0144}{42}}\right) \quad \text{or} \quad (.516,\quad .612)$$

$$\left(\bar{x}_2 - \sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(.05)} \sqrt{\frac{s_{22}}{n}},\quad \bar{x}_2 + \sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(.05)} \sqrt{\frac{s_{22}}{n}}\right)$$

$$= \left(.603 - \sqrt{\frac{2(41)}{40} 3.23} \sqrt{\frac{.0146}{42}},\quad .603 + \sqrt{\frac{2(41)}{40} 3.23} \sqrt{\frac{.0146}{42}}\right) \quad \text{or} \quad (.555,\quad .651)$$

In Figure 5.2, we have redrawn the 95% confidence ellipse from Example 5.3. The 95% simultaneous intervals are shown as shadows, or projections, of this ellipse on the axes of the component means. ∎

Simultaneous CIs As Projections of CR

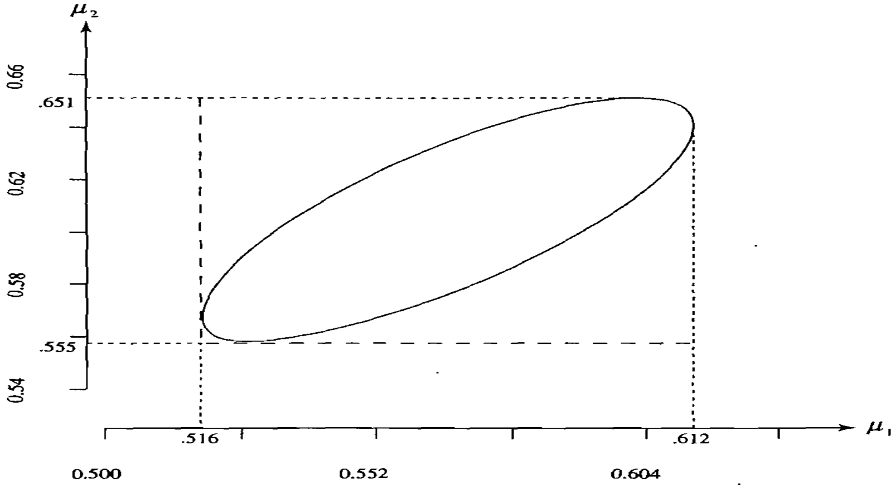

Figure 5.2 Simultaneous T^2-intervals for the component means as shadows of the confidence ellipse on the axes—microwave radiation data.

Example 5.5 (Constructing simultaneous confidence intervals and ellipses) The scores obtained by $n = 87$ college students on the College Level Examination Program (CLEP) subtest X_1 and the College Qualification Test (CQT) subtests X_2 and X_3 are given in Table 5.2 on page 228 for X_1 = social science and history, X_2 = verbal, and X_3 = science. These data give

One-at-a-Time Confidence Intervals

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

$$
\begin{aligned}
\bar{\mathbf{X}}_1 - t_{1-\alpha/2}(n-1)s_1/\sqrt{n} \leq &\ \mu_1 \leq \bar{\mathbf{X}}_1 + t_{1-\alpha/2}(n-1)s_1/\sqrt{n} \\
\vdots \qquad\qquad & \ \ \vdots \qquad\qquad \vdots \\
\bar{\mathbf{X}}_j - t_{1-\alpha/2}(n-1)s_j/\sqrt{n} \leq &\ \mu_j \leq \bar{\mathbf{X}}_j + t_{1-\alpha/2}(n-1)s_j/\sqrt{n} \\
\vdots \qquad\qquad & \ \ \vdots \qquad\qquad \vdots \\
\bar{\mathbf{X}}_p - t_{1-\alpha/2}(n-1)s_p/\sqrt{n} \leq &\ \mu_p \leq \bar{\mathbf{X}}_p + t_{1-\alpha/2}(n-1)s_p/\sqrt{n}
\end{aligned}
$$

- What is the issue?

One-at-a-Time Confidence Intervals

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

$$\begin{aligned}
\bar{\mathbf{X}}_1 - t_{1-\alpha/2}(n-1)s_1/\sqrt{n} &\le \mu_1 \le \bar{\mathbf{X}}_1 + t_{1-\alpha/2}(n-1)s_1/\sqrt{n} \\
&\vdots \\
\bar{\mathbf{X}}_j - t_{1-\alpha/2}(n-1)s_j/\sqrt{n} &\le \mu_j \le \bar{\mathbf{X}}_j + t_{1-\alpha/2}(n-1)s_j/\sqrt{n} \\
&\vdots \\
\bar{\mathbf{X}}_p - t_{1-\alpha/2}(n-1)s_p/\sqrt{n} &\le \mu_p \le \bar{\mathbf{X}}_p + t_{1-\alpha/2}(n-1)s_p/\sqrt{n}
\end{aligned}$$

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

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

The Bonferroni Method

- Often, we are concerned about a limited number, m, of linear combinations of means, i.e. $\mathbf{a}_1^T\boldsymbol{\mu}, \cdots, \mathbf{a}_m^T\boldsymbol{\mu}$.
- 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

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

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

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

Example 5.6 (Constructing Bonferroni simultaneous confidence intervals and comparing them with T^2-intervals) Let us return to the microwave oven radiation data in Examples 5.3 and 5.4. We shall obtain the simultaneous 95% Bonferroni confidence intervals for the means, μ_1 and μ_2, of the fourth roots of the door-closed and door-open measurements with $\alpha_i = .05/2$, $i = 1, 2$. We make use of the results in Example 5.3, noting that $n = 42$ and $t_{41}(.05/2(2)) = t_{41}(.0125) = 2.327$, to get

$$\bar{x}_1 \pm t_{41}(.0125)\sqrt{\frac{s_{11}}{n}} = .564 \pm 2.327\sqrt{\frac{.0144}{42}} \quad \text{or} \quad .521 \le \mu_1 \le .607$$

$$\bar{x}_2 \pm t_{41}(.0125)\sqrt{\frac{s_{22}}{n}} = .603 \pm 2.327\sqrt{\frac{.0146}{42}} \quad \text{or} \quad .560 \le \mu_2 \le .646$$

Figure 5.4 shows the 95% T^2 simultaneous confidence intervals for μ_1, μ_2 from Figure 5.2, along with the corresponding 95% Bonferroni intervals. For each component mean, the Bonferroni interval falls within the T^2-interval. Consequently, the rectangular (joint) region formed by the two Bonferroni intervals is contained in the rectangular region formed by the two T^2-intervals. If we are interested only in the component means, the Bonferroni intervals provide more precise estimates than

The Bonferroni Method of Multiple Comparisons

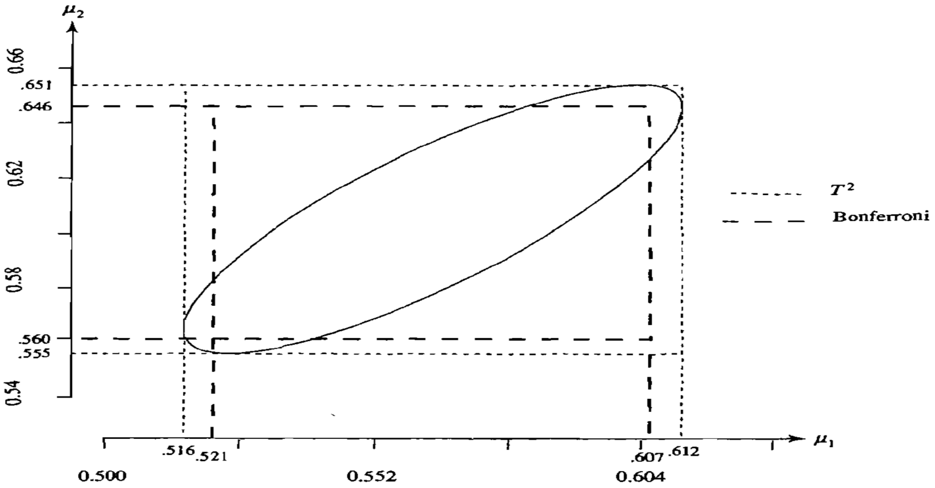

Figure 5.4 The 95% T^2 and 95% Bonferroni simultaneous confidence intervals for the component means—microwave radiation data.

Table of Contents

Large Sample Inference of Mean Vector

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

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

- 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

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

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

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

Large Sample Simultaneous Confidence Interval

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

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

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

$$\bar{\mathbf{X}}_j - z_{1-\alpha/2}\sqrt{s_{ii}/n} \le \mu_j \le \bar{\mathbf{X}}_j + z_{1-\alpha/2}\sqrt{s_{ii}/n}, \quad j = 1, \cdots, p$$

- The Bonferroni SCI is

$$\bar{\mathbf{X}}_j - z_{1-\alpha/(2p)}\sqrt{s_{ii}/n} \le \mu_j \le \bar{\mathbf{X}}_j + z_{1-\alpha/(2p)}\sqrt{s_{ii}/n}, \quad j = 1, \cdots, p$$

Example 5.7 (Constructing large sample simultaneous confidence intervals) A music educator tested thousands of Finnish students on their native musical ability in order to set national norms in Finland. Summary statistics for part of the data set are given in Table 5.5. These statistics are based on a sample of $n = 96$ Finnish 12th graders.

Table 5.5 Musical Aptitude Profile Means and Standard Deviations for 96 12th-Grade Finnish Students Participating in a Standardization Program

	Raw score	
Variable	Mean ($\bar{x}_i$)	Standard deviation ($\sqrt{s_{ii}}$)
X_1 = melody	28.1	5.76
X_2 = harmony	26.6	5.85
X_3 = tempo	35.4	3.82
X_4 = meter	34.2	5.12
X_5 = phrasing	23.6	3.76
X_6 = balance	22.0	3.93
X_7 = style	22.7	4.03

Source: Data courtesy of V. Sell.