STT 843: Multivariate Analysis

4. Inferences about a Mean Vector (Chapter 5.1-5.4)

Guanqun Cao

Department of Statistics and Probability
Michigan State University

Spring 2026

Outline

1. The Plausibility of μ_0 as a Value for a Normal Population Mean
2. Confidence Regions and Simultaneous Comparisons of Component Means
3. Simultaneous Confidence Statements
4. Control regions for future individual observations

The Plausibility of μ_0 as a Value for a Normal Population Mean

- The univariate theory for determining whether a specific value of μ_0 is a plausible value for the population mean μ.
- Test of the competing hypotheses

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu \neq \mu_0$$

Here H_0 is the null hypothesis and H_1 is the (two-side) alternative hypothesis.

- Rejecting H_0 when $|t|$ is large is equivalent to rejecting H_0 in favor of H_1, at significance level α if

$$t^2 = \frac{\left(\bar{X} - \mu_0\right)^2}{s^2/n} = n\left(\bar{X} - \mu_0\right)\left(s^2\right)^{-1}\left(\bar{X} - \mu_0\right) > t_{n-1}^2(\alpha/2)$$

- If H_0 is not rejected, we conclude that μ_0 is a plausible value for the normal population mean.

From the well-known correspondence between acceptance region for test of $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ and confidence interval for μ we have

$$\left\{ \text{Do not reject } H_0 : \mu = \mu_0 \text{ at level } \alpha \text{ or } \left| \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right| \leq t_{n-1}(\alpha/2) \right\}$$

is equivalent to

$$\left\{ \mu_0 \text{ lies in the } 100(1-\alpha)\% \text{ confidence interval } \bar{x} \pm t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}} \right\}$$

Hotelling's T^2

- A natural generalization of the squared distance above is its multivariate analog

$$T^2 = \left(\hat{\boldsymbol{X}} - \boldsymbol{\mu}_0\right)^{\mathsf{T}} \left(\frac{1}{n}\mathbf{S}\right)^{-1} \left(\hat{\boldsymbol{X}} - \boldsymbol{\mu}_0\right) = n\left(\hat{\boldsymbol{X}} - \boldsymbol{\mu}_0\right)^{\mathsf{T}} \mathbf{S}^{-1}\left(\overline{\boldsymbol{X}} - \boldsymbol{\mu}_0\right)$$

The statistics T^2 is called Hotelling's T^2.

Hotelling's T^2

First consider univariate test of $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$ when σ is known. (Consider only two-sided tests, since one-sided don't readily generalize for $p > 1$)
Test statistic using r.s. $(x_1, \ldots, x_n)$:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \text{ under } H_0$$

or

$$\begin{aligned} z^2 &= \text{ square of standardized distance} \\ &= n\left(\frac{\bar{x} - \mu_0}{\sigma}\right)^2 \sim \chi^2_1 \text{ under } H_0 \end{aligned}$$

Multivariate generalization ($\boldsymbol{\Sigma}$ is known):

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 = \begin{bmatrix} \mu_{01} \\ \mu_{02} \\ \vdots \\ \mu_{0p} \end{bmatrix} \quad \text{vs.} \quad \underbrace{H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0}_{\substack{\text{At least one } \mu_i \\ \text{is not equal to } \mu_{0i}}}$$

Test statistic using r.s. $(\underset{p\times 1}{\mathbf{x}_1, \ldots, \mathbf{x}_n})$:

$$z^2 = n\left(\bar{\mathbf{x}} - \boldsymbol{\mu}_0\right)^{\mathsf{T}} \boldsymbol{\Sigma}^{-1} \left(\bar{\mathbf{x}} - \boldsymbol{\mu}_0\right) \sim \chi^2_p \quad \text{under } H_0$$

More frequently in practice, $\boldsymbol{\Sigma}$ is unknown.

- Univariate test statistic using r.s. $(x_1, \ldots, x_n)$:

$$\begin{aligned} t^2 &= n\left(\frac{\bar{x}-\mu_0}{s}\right)^2 \sim t_{n-1}^2 \quad \text{under } H_0 \\ &= \sqrt{n}\,(\bar{x}-\mu_0)\left(s^2\right)^{-1}(\bar{x}-\mu_0)\sqrt{n} \\ &= \left[N_1\left(0,\sigma^2\right)\right]\left[\frac{\text{scaled } \chi^2}{df}\right]^{-1}\left[N_1\left(0,\sigma^2\right)\right] \end{aligned}$$

- If the observed statistical distance T^2 is too large- that is, if $\hat{\mathbf{x}}$ is "too far" from μ_0-the hypothesis $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ is rejected.

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

where $F_{p,n-p}$ denotes a random variable with an F-distribution with p and $n-p$ degree of freedom.

To summarize, we have the following:

- Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ be a random sample from an $N_p(\boldsymbol{\mu}, \Sigma)$ population.
- Then with $\overline{\boldsymbol{X}} = \frac{1}{n}\sum_{j=1}^n \boldsymbol{X}_j$ and $\mathbf{S} = \frac{1}{n-1}\sum_{j=1}^n \left(\boldsymbol{X}_j - \overline{\boldsymbol{X}}\right)\left(\boldsymbol{X}_j - \overline{\boldsymbol{X}}\right)^{\mathsf{T}}$,

$$\begin{aligned}\alpha &= P\left[T^2 > \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)\right] \\ &= P\left[n(\overline{\boldsymbol{X}} - \boldsymbol{\mu})^{\mathsf{T}} \mathbf{S}^{-1} (\overline{\boldsymbol{X}} - \boldsymbol{\mu}) > \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)\right]\end{aligned}$$

whatever the true $\boldsymbol{\mu}$ and Σ. Here $F_{p,n-p}(\alpha)$ is the upper (100α)th percentile of the $F_{p,n-p}$ distribution.

- Multivariate generalization (Hotelling's T^2):

$$T^2 = n\,(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^{\mathsf{T}} \quad \underset{\substack{\uparrow \\ \text{unbiased} \\ \text{estimate} \\ \text{of } \boldsymbol{\Sigma}}}{\mathbf{S}^{-1}} \quad (\bar{\mathbf{x}} - \boldsymbol{\mu}_0) \sim T^2_{p,n-1} \quad \text{under } H_0$$

$$= \underbrace{(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^{\mathsf{T}}}_{\substack{\uparrow \\ \bar{\mathbf{x}} \text{ and } \mathbf{S} \\ \text{are indep.}}} \underbrace{\left(\frac{\mathbf{S}}{n}\right)^{-1}}_{\substack{\text{inverse} \\ \text{sample cov matrix}}} (\bar{\mathbf{x}} - \boldsymbol{\mu}_0) \leftarrow \text{"characteristic form"}$$

$$= \underbrace{\sqrt{n}\,(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^{\mathsf{T}}}_{\substack{N_p(\mathbf{0},\,\boldsymbol{\Sigma})\\ \text{random vector}}} \quad \underbrace{\left[\frac{\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^{\mathsf{T}}}{n-1}\right]^{-1}}_{\substack{W_p(n-1,\,\boldsymbol{\Sigma})\\ \text{random matrix}\\ \text{divided by d.f.}}} \quad \underbrace{\sqrt{n}\,(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)}_{\substack{N_p(\mathbf{0},\,\boldsymbol{\Sigma})\\ \text{random vector}}}$$

Important properties of T^2

1. Sometimes we refer to the subscripts for $T^2_{p,\nu}$ distribution as "dimension" and "df" (e.g., $T^2_{\text{dim,df}}$)
2. Must have $n > p$
   - Otherwise $\mathbf{S}$ is singular and T^2 cannot be computed.
3. Degrees of freedom ν for T^2 is same as for analogous univariate t-test:
   - $\nu = n - 1$ for one-sample test
   - $\nu = n_1 + n_2 - 2$ for two-sample test

4. Alternative hypothesis is 2 -sided (no such thing as " $H_1 : \boldsymbol{\mu} > \boldsymbol{\mu}_0$ ")
   - Critical region is one-tailed (reject for large values) since test statistic is squared distance
5. $\frac{\nu-p+1}{\nu p} T^2_{p,\nu} \overset{q}{\equiv} F_{p,\nu-p+1}$

[Note: " $\equiv$ " is shorthand for the equivalence of the quantiles of two dist'ns]

- So, p-value for T^2 test is

$$p\text{-value } = \Pr\left\{F_{p,\nu-p+1} > \frac{\nu-p+1}{\nu p} T^2\right\}$$

- Critical value for T^2 test is

$$T^2_{\alpha,p,\nu} = \frac{\nu p}{\nu-p+1} F_{\alpha,p,\nu-p+1} \left(\text{or } \frac{(n-1)p}{n-p} F_{\alpha,p,n-p} \text{ when } \nu = n-1 \right)$$

6. T^2 invariant under transformations of the form $\underset{p\times 1}{\ddot{\mathbf{x}}} = \underset{p\times p}{\mathbf{C}}\underset{p\times 1}{\mathbf{x}} + \mathbf{d}$, where $\mathbf{C}$ is nonsingular
7. T^2 is the likelihood ratio test (LRT) of $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$

- Under H_0 the likelihood is

$$L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{np/2}|\boldsymbol{\Sigma}|^{n/2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu}_0)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_0)\right\}$$
$$= \frac{1}{(2\pi)^{np/2}|\boldsymbol{\Sigma}|^{n/2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left[\boldsymbol{\Sigma}^{-1}\left(\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu}_0)(\mathbf{x}_i - \boldsymbol{\mu}_0)^{\mathrm{T}}\right)\right]\right\}$$

Using Result 4.10 (again), we obtain

$$\max_{\boldsymbol{\Sigma}} L(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{np}{2}}\left|\hat{\boldsymbol{\Sigma}}_0\right|^{\frac{n}{2}}} \exp\left\{\frac{-np}{2}\right\}$$

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

- Recall

$$\max_{\boldsymbol{\mu},\boldsymbol{\Sigma}} L(\boldsymbol{\mu},\boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{np}{2}}|\hat{\boldsymbol{\Sigma}}|^{\frac{n}{2}}} \exp\left\{\frac{-np}{2}\right\}$$

where $\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^{\mathsf{T}}$

Likelihood Ratio: $\lambda = \frac{\max_{\boldsymbol{\Sigma}} L(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu}_0,\boldsymbol{\Sigma}} L(\boldsymbol{\mu},\boldsymbol{\Sigma})} = \left(\frac{|\widehat{\boldsymbol{\Sigma}}|}{|\widehat{\boldsymbol{\Sigma}}_0|}\right)^{\frac{n}{2}}$

$$\Lambda = \lambda^{\frac{2}{n}} = \frac{|\hat{\boldsymbol{\Sigma}}|}{\left|\hat{\boldsymbol{\Sigma}}_0\right|} = \frac{1}{1 + \frac{1}{n-1} T^2}$$

"Wilks' Lambda" is rejected for small Λ or large T^2

$* T^2 = (n-1)\frac{\left|\hat{\Sigma}_0\right|}{|\hat{\boldsymbol{\Sigma}}|} - (n-1)$

- $-2 \ln \lambda \sim \chi^2_{\nu-\nu_0}$

where $\nu = \#$ of unrestricted parameters and $\nu_0 = \#$ of parameters under H_0

Example 4.1(Evaluating T^2) Let the data matrix for a random sample of size $n = 3$ from a bivariate normal population be

$$\boldsymbol{X} = \begin{bmatrix} 6 & 9 \\ 10 & 6 \\ 8 & 3 \end{bmatrix}$$

Evaluate the observed T^2 for $\boldsymbol{\mu}_0^\top = [9, 5]$. What is the sampling distribution of T^2 in this case?

Example 4.2 (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 $H_0 : \boldsymbol{\mu}^{\mathsf{T}} = [4, 50, 10]$ against $H_1 : \boldsymbol{\mu}^{\mathsf{T}} \neq [4, 50, 10]$ at level of significance $\alpha = .10$.

Table 5.1 Sweat Data

Individual	X_1 (Sweat rate)	X_2 (Sodium)	X_3 (Potassium)
1	3.7	48.5	9.3
2	5.7	65.1	8.0
3	3.8	47.2	10.9
4	3.2	53.2	12.0
5	3.1	55.5	9.7
6	4.6	36.1	7.9
7	2.4	24.8	14.0
8	7.2	33.1	7.6
9	6.7	47.4	8.5
10	5.4	54.1	11.3
11	3.9	36.9	12.7
12	4.5	58.8	12.3
13	3.5	27.8	9.8
14	4.5	40.2	8.4
15	1.5	13.5	10.1
16	8.5	56.4	7.1
17	4.5	71.6	8.2
18	6.5	52.8	10.9
19	4.1	44.1	11.2
20	5.5	40.9	9.4

Source: Courtesy of Dr. Gerald Bargman.

Confidence Regions and Simultaneous Comparisons of Component Means

- The region $R(\boldsymbol{X})$ is said to be a $100(1-\alpha)\%$ confidence region if, before the sample is selected,

$$P[(\boldsymbol{X}) \text{ will cover the true } \theta] = 1-\alpha$$

This probability is calculated under the true, but unknown value of $\boldsymbol{\theta}$.

- The confidence region for the mean $\boldsymbol{\mu}$ of a p-dimension normal population is available. Before the sample is selected,

$$P\left[n(\overline{\boldsymbol{X}}-\boldsymbol{\mu})^{\mathsf{T}}\mathbf{S}^{-1}(\overline{\boldsymbol{X}}-\boldsymbol{\mu}) \leq \frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha)\right] = 1-\alpha$$

whatever the values of the unknown $\boldsymbol{\mu}$ and Σ. In words, $\overline{\boldsymbol{X}}$ will be within $[(n-1)pF_{p,n-p}(\alpha)/(n-p)]^{1/2}$ of $\boldsymbol{\mu}$, with probability $1-\alpha$, provided that distance is defined in terms of $n\mathbf{S}^{-1}$.

- For $P \geq 4$,we cannot graph the joint confidence region for $\boldsymbol{\mu}$. However we can calculate the axes of the confidence ellipsoid and their relative lengths.
- These are determined from the eigenvalues λ_i and eigenvectors $\mathbf{e}_i$ of $\mathbf{S}$. The direction and lengths of the axes of

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

  are determined by going

$$\sqrt{\lambda} c / \sqrt{n} = \sqrt{\lambda_i} \sqrt{p(n-1) F_{p,n-p}(\alpha) / n(n-p)}$$

units along the eigenvectors $\mathbf{e}_i$.

Beginning at the center $\bar{\mathbf{x}}$, the axes of the confidence ellipsoid are

$$\pm\sqrt{\lambda_i}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)}\mathbf{e}_i \quad \text{where} \quad \mathbf{S}\mathbf{e}_i = \lambda_i\mathbf{e}_i, i = 1, 2, \ldots, p$$

The ratio of the λ_i 's will help identify relative amount of elongation along pair of axes

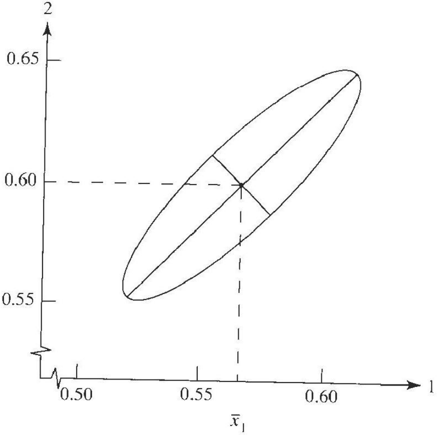
Example 4.3 (Constructing a confidence ellipse for μ) Data for radiation from microwave oven were introduced in Example 3.10 and 3.17. Let

$$x_1 = (\text{ measured ration with door closed })^{\frac{1}{4}}$$

and

$$x_2 = (\text{ measured ration with door open })^{\frac{1}{4}}$$

Construct the confidence ellipse for $\boldsymbol{\mu} = [\mu_1, \mu_2]$



95% ellipse for μ based on microwave radiation data.

Simultaneous Confidence Statements

- While the confidence region $n(\bar{\mathbf{x}} - \boldsymbol{\mu})^\mathsf{T}\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}) \leq c^2$, for c a constant, correctly assesses the joint knowledge concerning plausible value of $\boldsymbol{\mu}$, any summary of conclusions ordinarily includes confidence statement about the individual component means.
- In so doing, we adopt the attitude that all of the separate confidence statements should hold simultaneously with specified high probability.
- It is the guarantee of a specified probability against any statement being incorrect that motivates the term simultaneous confidence intervals

- Let $\boldsymbol{X}$ have an $N_p(\boldsymbol{\mu}, \Sigma)$ distribution and form the linear combination

$$Z = a_1 X_1 + a_2 X_2 + \cdots + a_p X_p = \mathbf{a}^\top \boldsymbol{X}$$

$$\mu_Z = \mathrm{E}(Z) = \mathbf{a}^\top \boldsymbol{\mu} \quad \text{and} \quad \sigma_Z^2 = \mathrm{Var}(Z) = \mathbf{a}^\top \Sigma \mathbf{a}$$

- Moreover, by Result 3.2, Z has an $N\left(\mathbf{a}^\top \boldsymbol{\mu}, \mathbf{a}^\top \Sigma \mathbf{a}\right)$ distribution. If a random sample $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ from the $N_p(\boldsymbol{\mu}, \Sigma)$ and $Z_j = \mathbf{a}^\top \boldsymbol{X}_j, j = 1, 2, \ldots, n$. Then the sample mean and variance of the observed values $z_1, z_2, \ldots, z_n$ are

$$\bar{z} = \mathbf{a}^\top \overline{\mathbf{x}} \quad \text{and} \quad s_z^2 = \mathbf{a}^\top \text{ Sa}$$

where $\hat{\mathbf{x}}$ and $\mathbf{S}$ are the sample mean vector and covariance matrix of the $\mathbf{x}_j$'s respectively.

- For a fixed and σ_Z^2 unknown, a $100(1-\alpha)\%$ confidence interval for $\mu_Z = \mathbf{a}^\mathsf{T}\boldsymbol{\mu}$ is based on student's t-ratio

$$t = \frac{\bar{z} - \mu_Z}{s_Z/\sqrt{n}} = \frac{\sqrt{n}\left(\mathbf{a}^\mathsf{T}\overline{\mathbf{x}} - \mathbf{a}^\mathsf{T}\boldsymbol{\mu}\right)}{\sqrt{\mathbf{a}^\mathsf{T}\mathbf{S}\mathbf{a}}}$$

- and leads to the statement

$$\bar{z} - t_{n-1}(\alpha/2)\frac{s_z}{\sqrt{n}} \le \mu_Z \le \bar{z} + t_{n-1}(\alpha/2)\frac{s_z}{\sqrt{n}}$$

or

$$\mathbf{a}^\mathsf{T}\overline{\mathbf{x}} - t_{n-1}(\alpha/2)\frac{\sqrt{\mathbf{a}^\mathsf{T}\mathbf{S}\mathbf{a}}}{\sqrt{n}} \le \mathbf{a}^\mathsf{T}\boldsymbol{\mu} \le \mathbf{a}^\mathsf{T}\overline{\mathbf{x}} + t_{n-1}(\alpha/2)\frac{\sqrt{\mathbf{a}^\mathsf{T}\mathbf{S}\mathbf{a}}}{\sqrt{n}}$$

where $t_{n-1}(\alpha/2)$ is the upper $100(\alpha/2)$ th percentile of a t-distribution with $n-1$ d.f.

- Clearly, we could make serval confidence statements about the components of $\boldsymbol{\mu}$, each with associated confidence coefficient $1-\alpha$, by choosing different coefficient vector a. However, the confidence associated with all of the statements taken together is not $1-\alpha$.
- Intuitively, it would be desirable to associate a "collective" confidence coefficient of $1-\alpha$ with the confidence intervals that can be generated by all choices of a. However, a price must be paid for the convenience of a large simultaneous confidence coefficient: intervals that are wider than the interval for a specific choice of **a**.
- Given a data set $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ and a particular **a**, the confidence interval is that set of $\mathbf{a}^\top\boldsymbol{\mu}$ values for which

$$|t| = \left| \frac{\sqrt{n}\left(\mathbf{a}^{\mathsf{T}}\bar{\mathbf{x}} - \mathbf{a}^{\mathsf{T}}\boldsymbol{\mu}\right)}{\sqrt{\mathbf{a}^{\mathsf{T}}\mathbf{S}\mathbf{a}}} \right| \leq t_{n-1}(\alpha/2)$$

- A simultaneous confidence region is given by the set $\mathbf{a}^{\mathsf{T}}\boldsymbol{\mu}$ values such that t^2 is relatively small for all choice of $\mathbf{a}$. It seems reasonable to expect that the constant $t^2(\alpha/2)$ will be replaced by a large value c^2, when statements are developed for many choices of $\mathbf{a}$.
- Considering the values of $\mathbf{a}$ for which $t^2 \leq c^2$, we are naturally led to the determination of

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

with the maximum occurring for a proportional to $\mathbf{S}^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu})$.

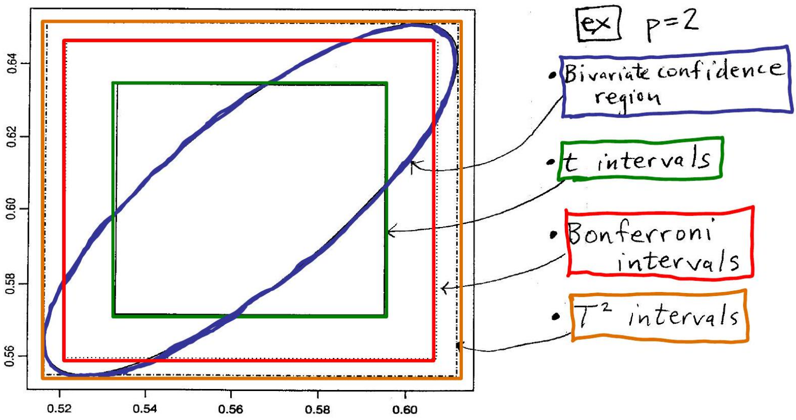
Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ be a random sample from an $N_p(\boldsymbol{\mu}, \Sigma)$ population with Σ positive definite. Then simultaneously for all a, the interval

$$\left(\mathbf{a}^\mathsf{T}\overline{\boldsymbol{X}} - \sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)\mathbf{a}^\mathsf{T}\mathbf{S}\mathbf{a}}, \quad \mathbf{a}^\mathsf{T}\overline{\boldsymbol{X}} + \sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)\mathbf{a}^\mathsf{T}\mathbf{S}\mathbf{a}}\right)$$

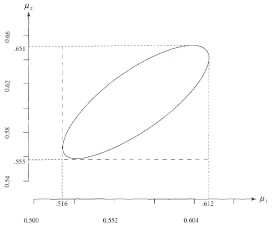
will contain $\mathbf{a}^\mathsf{T}\boldsymbol{\mu}$ with probability $1-\alpha$.

Example 4.4 (Simultaneous confidence intervals as shadows of the confidence ellipsoid)

In Example 4.3, we obtain the 95% confidence ellipse for the means of the four roots of the door-closed and door-open microwave radiation measurements. Obtain the 95% simultaneous T^2 intervals for the two component means.


0.64
0.62
0.60
0.58
0.56
0.52
0.54
0.56
0.58
0.60
ex p=2
Bivariate confidence region
t intervals
Bonferroni intervals
T² intervals

Example 4.5 (Constructing simultaneous confidence intervals and ellipse) 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 5.3 for $X_1 =$ social science and history, $X_2 =$ verbal, and $X_3 =$ science. Construct simultaneous confidence intervals for μ_1, μ_2 and μ_3.



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

Individual	X_1 (Social science and history)	(Verbal)	(Science)	Individual	X_1 (Social science and history)	(Verbal)	X_3 (Science)
1	468	41	26	45	494	41	24
2	428	39	26	46	541	47	25
3	514	53	21	47	362	36	17
4	547	67	33	48	408	28	17
5	614	61	27	49	594	68	23
6	501	67	29	50	501	25	26
7	421	46	22	51	687	75	33
8	527	50	23	52	633	52	31
9	527	55	19	53	647	67	29
10	620	72	32	54	647	65	34
11	587	63	31	55	614	59	25
12 ...	541 ..	59 ..	19 ..	56 ..	633 ..	65 ..	28 ..

Control regions for future individual observations

Let $\mathbf{X}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, i.i.d., $i = 1, \ldots, n$. Let $\mathbf{X}$ be a future observation from the same distribution. Then

$$T^2 = \frac{n}{n+1}\left(\mathbf{X} - \bar{\mathbf{X}}\right)^{\mathsf{T}} \mathbf{S}^{-1}\left(\mathbf{X} - \bar{\mathbf{X}}\right) \sim \frac{(n-1)p}{n-p} F_{p,n-p}$$

and a $100(1-\alpha)\%$ p-dim prediction ellipsoid is given by all $\mathbf{x}$ satisfying

$$\frac{n}{n+1}\left(\mathbf{x} - \bar{\mathbf{X}}\right)^{\mathsf{T}} \mathbf{S}^{-1}\left(\mathbf{x} - \bar{\mathbf{X}}\right) \le \frac{(n^2-1)}{n(n-p)} F_{p,n-p}(\alpha)$$