Inferences about a Mean Vector

Shyh-Kang Jeng
Department of Electrical Engineering/
Graduate Institute of Communication/
Graduate Institute of Networking and Multimedia

1

Outline

- Introduction
- Inferences about a Mean for Univariate Normal Distribution
- The probability of μ_0 as a value for a Normal Population Mean
- Hotelling's T^2 and Likelihood Ratio Tests
- Confidence Regions and Simultaneous Comparison of Component Means

Outline

- Large Sample Inferences about a Population Mean Vector
- Multivariate Quality Control Charts
- Inferences about Mean Vectors When Some Observations Are Missing
- Difficulties Due to Time Dependence in Multivariate Observations

Outline

- Introduction
- Inferences about a Mean for Univariate Normal Distribution
- The probability of μ_0 as a value for a Normal Population Mean
- Hotelling's T^2 and Likelihood Ratio Tests
- Confidence Regions and Simultaneous Comparison of Component Means

Questions

- What is a statistical inference?
- Give a sample scenario for making a statistical inference

Inference

- Reaching valid conclusions concerning a population on the basis of information from a sample

6

Scenarios

- To test if the following statements are plausible
  - A clam by a cram school that their course can increase the IQ of your children
  - A diuretic is effective
  - An MP3 compressor is with higher quality
  - A claim by a lady that she can distinguish whether the milk is added before making milk tea

7

Evaluating Normality of Univariate Marginal Distributions

Number of samples within an interval:

binomial distribution $\binom{n}{y} p^y q^{n-y}$

When n is large, $\binom{n}{y} p^y q^{n-y} \approx N(np, npq)$

The distribution of $\hat{p} = \frac{y}{n}$ is $N(p, \frac{pq}{n})$

8

Evaluating Normality of Univariate Marginal Distributions

After checking symmetry of data,

$\hat{p}_1$: portion of data lying in $(\bar{x} - \sqrt{s}, \bar{x} + \sqrt{s})$

$\hat{p}_2$: portion of data lying in $(\bar{x} - 2\sqrt{s}, \bar{x} + 2\sqrt{s})$

either $|\hat{p}_1 - 0.683| > 3\sqrt{\frac{(0.683)(0.317)}{n}} = \frac{1.396}{\sqrt{n}}$

or $|\hat{p}_2 - 0.954| > 3\sqrt{\frac{(0.954)(0.046)}{n}} = \frac{0.628}{\sqrt{n}}$

indicate departure from an assumed normal distribution

9

Outline

- Introduction
- Inferences about a Mean for Univariate Normal Distribution
- The probability of μ_0 as a value for a Normal Population Mean
- Hotelling's T^2 and Likelihood Ratio Tests
- Confidence Regions and Simultaneous Comparison of Component Means

Questions

- What are the two-sided and the one-sided tests of hypotheses?
- How to reject or accept a null hypothesis?
- What is the Student's t-statistics?
- What are the differences between the normal distribution and the Student's t-distribution?

Questions

- What is the meaning of the confidence interval for the population mean μ_0?

Tests of Hypotheses

- Developed by Fisher, Pearson, Neyman, etc.
- Two-sided
  - $H_0 : \mu = \mu_0$ (null hypothesis)
  - $H_1 : \mu \neq \mu_0$ (alternative hypothesis)
- One-sided
  - $H_0 : \mu > \mu_0$ (null hypothesis)
  - $H_1 : \mu \leq \mu_0$ (alternative hypothesis)

13

Assumption under Null Hypothesis

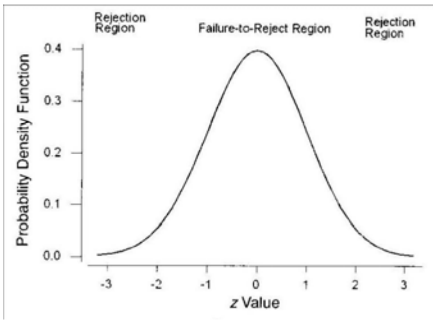
$$X : N(\mu_0, \sigma^2)$$

$$\overline{X} : N(\mu_0, \sigma^2 / n)$$

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} : N(0,1)$$

14

Rejection or Acceptance of Null Hypothesis



15

Student's *t*-Statistics

$$t = \frac{\left(\overline{X} - \mu_0\right)}{s / \sqrt{n}}, \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

16

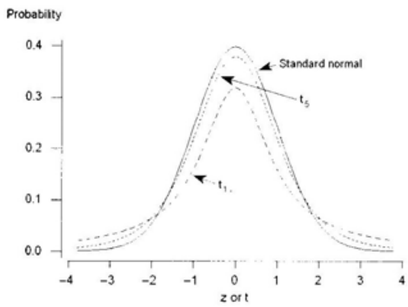
Student's *t*-distribution

$$t = \frac{Z}{\sqrt{\chi^2 / f}}$$

$$f(t) = \frac{\Gamma(\frac{f+1}{2})}{\sqrt{f\pi}\,\Gamma(\frac{f}{2})}\left(1 + \frac{t^2}{f}\right)^{-\frac{f+1}{2}}$$

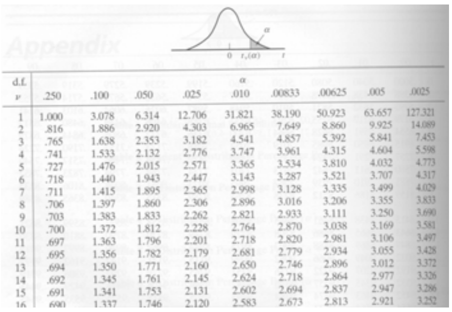
17

Student's *t*-distribution



18

Student's *t*-distribution



d.f. ν	.250	.100	.050	.025	.010	.00833	.00625	.005	.0025
1	1.000	3.078	6.314	12.706	31.821	38.190	50.923	63.657	127.321
2	.816	1.886	2.920	4.303	6.965	7.649	8.860	9.925	14.089
3	.765	1.638	2.353	3.182	4.541	4.857	5.392	5.841	7.453
4	.741	1.533	2.132	2.776	3.747	3.961	4.315	4.604	5.598
5	.727	1.476	2.015	2.571	3.365	3.534	3.810	4.032	4.773
6	.718	1.440	1.943	2.447	3.143	3.287	3.521	3.707	4.317
7	.711	1.415	1.895	2.365	2.998	3.128	3.335	3.499	4.029
8	.706	1.397	1.860	2.306	2.896	3.016	3.206	3.355	3.833
9	.703	1.383	1.833	2.262	2.821	2.933	3.111	3.250	3.690
10	.700	1.372	1.812	2.228	2.764	2.870	3.038	3.169	3.581
11	.697	1.363	1.796	2.201	2.718	2.820	2.981	3.106	3.497
12	.695	1.356	1.782	2.179	2.681	2.779	2.934	3.055	3.428
13	.694	1.350	1.771	2.160	2.650	2.746	2.896	3.012	3.372
14	.692	1.345	1.761	2.145	2.624	2.718	2.864	2.977	3.326
15	.691	1.341	1.753	2.131	2.602	2.694	2.837	2.947	3.286
16	.690	1.337	1.746	2.120	2.583	2.673	2.813	2.921	3.252

Origin of the Name "Student"

- Pseudonym of William Gossett at Guinness Brewery in Dublin around the turn of the 20th Century
- Gossett use pseudonym because all Guinness Brewery employees were forbidden to publish
- Too bad Guinness doesn't run universities

Test of Hypothesis

Reject H_0 in favor of H_1

at significance level α if

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

i.e.,

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

Selection of α

- Often chosen as 0.05, 0.01, or 0.1
- Actually, Fisher said in 1956:
  - No scientific worker has a fixed level of significance at which year to year, and in all circumstances, he rejects hypotheses; he rather gives his mind to each particular case in the light of his evidence and hid ideas

Confidence Interval for μ_0

$$\Pr\left(\left| \frac{\bar{X} - \mu_0}{s / \sqrt{n}} \right| \le t_{n-1}(0.025) \right) = 0.95$$

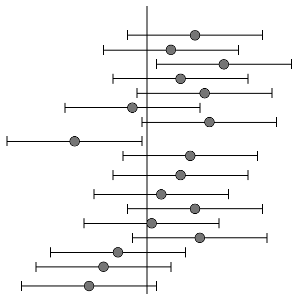
$$\Pr\left(t_{n-1}(0.025) s / \sqrt{n} \ge \bar{X} - \mu \ge -t_{n-1}(0.025) s / \sqrt{n} \right)$$
$$= 0.95$$

$$\Pr\left(-\bar{X} + t_{n-1}(0.025) s / \sqrt{n} \ge -\mu \ge -\bar{X} - t_{n-1}(0.025) s / \sqrt{n} \right)$$
$$= 0.95$$

$$\Pr\left(\bar{X} + t_{n-1}(0.025) s / \sqrt{n} \ge \mu \ge \bar{X} - t_{n-1}(0.025) s / \sqrt{n} \right)$$
$$= 0.95$$

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

Neyman's Interpretation



μ_0

Statistically Significant vs. Scientifically Significant

- The cram school claims that its course will increase the IQ of your child statistically significant at the 0.05 level
- Assume that 100 students took the courses were tested, and the population standard deviation is 15
- The actual IQ improvement to be statistically significant at 0.05 level is simply $(\sigma/\sqrt{n})\times z_{0.025} = 1.5\times 1.96 = 2.94$

25

Outline

- Introduction
- Inferences about a Mean for Univariate Normal Distribution
- The probability of $\boldsymbol{\mu}_0$ as a value for a Normal Population Mean
- Hotelling's T^2 and Likelihood Ratio Tests
- Confidence Regions and Simultaneous Comparison of Component Means

Questions

- How to test a null hypothesis for a multivariate normal distribution?
- How to convert Hotelling's T^2 distribution to the F distribution?
- Is the Hotelling's T^2 distribution invariant with linear transformation?

Plausibility of $\boldsymbol{\mu}_0$ as a Multivariate Normal Population Mean

Null hypothesis $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$

(Two-sided) alternative hypothesis $H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$

$\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$: Random sample from a normal population

Hotelling's T^2 statistics:

$$T^2 = (\overline{\mathbf{X}} - \boldsymbol{\mu}_0)'\left(\frac{\mathbf{S}}{n}\right)^{-1}(\overline{\mathbf{X}} - \boldsymbol{\mu}_0)$$

$$= n(\overline{\mathbf{X}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\overline{\mathbf{X}} - \boldsymbol{\mu}_0)$$

$$\overline{\mathbf{X}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{X}_j, \mathbf{S} = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{X}_j - \overline{\mathbf{X}})(\mathbf{X}_j - \overline{\mathbf{X}})'$$

28

T^2 as an *F*-Distribution

$$T^2 : \frac{(n-1)p}{n-p}F_{p,n-p}$$

29

F-Distribution

χ_1^2, χ_2^2 : independent, with d.f. f_1 and f_2, respectively

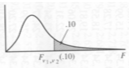
$$F = \frac{\chi_1^2 / f_1}{\chi_2^2 / f_2}, F > 0$$

$$f(F) = \frac{\Gamma\left(\frac{f_1+f_2}{2}\right)}{\Gamma\left(\frac{f_1}{2}\right)\Gamma\left(\frac{f_2}{2}\right)}\left(\frac{f_1}{f_2}\right)^{\frac{f_1}{2}}\frac{F^{\frac{f_1}{2}-1}}{\left(1+\frac{f_1F}{f_2}\right)^{(f_1+f_2)/2}}$$

$: F_{f_1,f_2}$

30

F-Distribution



ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.05	62.26	62.53	62.79
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.17	5.17	5.16	5.15
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.81	2.80	2.78	2.76
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.57	2.56	2.54	2.51
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.27	2.25	2.23	2.21
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.17	2.16	2.13	2.11
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03

Nature of T^2-Distribution

$$T^2=\sqrt{n}(\bar{\mathbf{X}}-\mu_0)'\left(\frac{\sum_{j=1}^{n}(\mathbf{X}_j-\bar{\mathbf{X}})(\mathbf{X}_j-\bar{\mathbf{X}})'}{n-1}\right)^{-1}\sqrt{n}(\bar{\mathbf{X}}-\mu_0)$$

$$T^2_{p,n-1}=\begin{pmatrix}\text{multivariate normal}\\ \text{random vector}\end{pmatrix}'\left(\frac{\text{Wishart random matrix}}{\text{d.f.}}\right)^{-1}\begin{pmatrix}\text{multivariate normal}\\ \text{random vector}\end{pmatrix}$$

$$=N_p(0,\Sigma)'\left[\frac{1}{n-1}W_{p,n-1}(\Sigma)\right]N_p(0,\Sigma)$$

$$\Rightarrow \text{by calculus, } \frac{\text{d.f.}\times p}{(\text{d.f.}-p+1)}F_{p,\text{d.f.}-p+1}=\frac{(n-1)p}{(n-p)}F_{p,n-p}$$

Test of Hypothesis

Reject H_0 in favor of H_1 at significance level α if

$$T^2=n(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)>\frac{(n-1)p}{n-p}F_{p,n-p}(\alpha)$$

Example 5.1 Evaluating T^2

$$\mathbf{X}=\begin{bmatrix}6&9\\10&6\\8&3\end{bmatrix},\boldsymbol{\mu}_0=\begin{bmatrix}9\\5\end{bmatrix}\Rightarrow\bar{\mathbf{x}}=\begin{bmatrix}8\\6\end{bmatrix},\mathbf{S}=\begin{bmatrix}4&-3\\-3&9\end{bmatrix}$$

$$\mathbf{S}^{-1}=\begin{bmatrix}1/3&1/9\\1/9&4/27\end{bmatrix}$$

$$T^2=3\begin{bmatrix}8-9&6-5\end{bmatrix}\begin{bmatrix}1/3&1/9\\1/9&4/27\end{bmatrix}\begin{bmatrix}8-9\\6-5\end{bmatrix}=\frac{7}{9}$$

$$T^2:\frac{(3-1)2}{(3-2)}F_{2,3-2}=4F_{2,1}$$

Example 5.2 Testing a Mean Vector

$H_0:\boldsymbol{\mu}'=[4\quad 50\quad 10]$, $H_1:\boldsymbol{\mu}'\neq[4\quad 50\quad 10]$

Test at a level $\alpha=0.10$. $n = 20$, check normality

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

$$\mathbf{S}^{-1}=\begin{bmatrix}0.586&-0.022&0.258\\-0.022&0.006&-0.002\\0.258&-0.002&0.402\end{bmatrix},T^2=9.74$$

$$\text{Critical value}:\frac{(n-1)p}{(n-p)}F_{p,n-p}(0.10)=\frac{19\times 3}{17}F_{3,17}(0.1)=8.18$$

$T^2=9.74>8.18\Rightarrow$ Reject H_0 at the 10% level

Invariance of T^2-Statistic

$\mathbf{Y}=\mathbf{CX}+\mathbf{d}$, $\mathbf{C}$: non-singular

$$\bar{\mathbf{y}}=\mathbf{C}\bar{\mathbf{x}}+\mathbf{d},\mathbf{S}_{\mathbf{y}}=\frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{y}_j-\bar{\mathbf{y}})(\mathbf{y}_j-\bar{\mathbf{y}})'=\mathbf{CSC}'$$

$$\boldsymbol{\mu}_{\mathbf{Y}}=E(\mathbf{Y})=\mathbf{C}\boldsymbol{\mu}+\mathbf{d},\ \boldsymbol{\mu}_{\mathbf{Y},0}=\mathbf{C}\boldsymbol{\mu}_0+\mathbf{d}$$

$$T^2=n(\bar{\mathbf{y}}-\boldsymbol{\mu}_{\mathbf{Y},0})'\mathbf{S}_{\mathbf{y}}^{-1}(\bar{\mathbf{y}}-\boldsymbol{\mu}_{\mathbf{Y},0})$$
$$=n(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\mathbf{C}'(\mathbf{CSC}')^{-1}\mathbf{C}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)$$
$$=n(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu}_0)$$

Outline

- Introduction
- Inferences about a Mean for Univariate Normal Distribution
- The probability of μ_0 as a value for a Normal Population Mean
- Hotelling's T^2 and Likelihood Ratio Tests
- Confidence Regions and Simultaneous Comparison of Component Means

Questions

- What is the likelihood ratio test?
- How to derive the Hotelling's T^2 distribution using the likelihood ratio test? (Result 5.1)
- What is the general likelihood ratio method?
- What is the behavior of the general likelihood ratio method when sample size n is large?

T^2-Statistic from Likelihood Ratio Test

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

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{j=1}^{n}(\mathbf{x}_j-\bar{\mathbf{x}})(\mathbf{x}_j-\bar{\mathbf{x}})', \quad \hat{\boldsymbol{\mu}} = \bar{\mathbf{x}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{x}_j$$

$$L(\boldsymbol{\mu}_0,\boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{np/2}|\boldsymbol{\Sigma}|^{n/2}} \exp\left(-\frac{1}{2}\sum_{j=1}^{n}(\mathbf{x}_j-\boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_j-\boldsymbol{\mu}_0)\right)$$

$$\text{Likelihood ratio} = \Lambda = \frac{\max_{\boldsymbol{\Sigma}} L(\boldsymbol{\mu}_0,\boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu},\boldsymbol{\Sigma}} L(\boldsymbol{\mu},\boldsymbol{\Sigma})}$$

T^2-Statistic from Likelihood Ratio Test

$$\sum_{j=1}^{n}(\mathbf{x}_j-\boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\mathbf{x}_j-\boldsymbol{\mu}_0) = \sum_{j=1}^{n}\mathrm{tr}\left[\boldsymbol{\Sigma}^{-1}(\mathbf{x}_j-\boldsymbol{\mu}_0)(\mathbf{x}_j-\boldsymbol{\mu}_0)'\right]$$

$$= \mathrm{tr}\left[\boldsymbol{\Sigma}^{-1}\sum_{j=1}^{n}(\mathbf{x}_j-\boldsymbol{\mu}_0)(\mathbf{x}_j-\boldsymbol{\mu}_0)'\right]$$

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

$$\hat{\boldsymbol{\Sigma}}_0 = \frac{1}{n}\sum_{j=1}^{n}(\mathbf{x}_j-\boldsymbol{\mu}_0)(\mathbf{x}_j-\boldsymbol{\mu}_0)'$$

Result 4.10

$\mathbf{B}$: $p\times p$ symmetric positive definite matrix

b : positive scalar

$$\frac{1}{|\boldsymbol{\Sigma}|^b} e^{-\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{B})/2} \le \frac{1}{|\mathbf{B}|^b}(2b)^{pb} e^{-bp}$$

for all positive definite $\underset{(p\times p)}{\boldsymbol{\Sigma}}$, with equality holding only for $\boldsymbol{\Sigma} = (1/2b)\mathbf{B}$

Likelihood Ratio Test

$\Lambda^{2/n} = |\hat{\boldsymbol{\Sigma}}|/|\hat{\boldsymbol{\Sigma}}_0|$: Wilks' lambda

Likelihood ratio test of

$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ against $H_1 : \boldsymbol{\mu} \ne \boldsymbol{\mu}_0$

Reject H_0 at the level α

$$\text{if } \Lambda = \left(\frac{|\hat{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}_0|}\right)^{n/2} = \left(\frac{\left|\sum_{j=1}^{n}(\mathbf{x}_j-\bar{\mathbf{x}})(\mathbf{x}_j-\bar{\mathbf{x}})'\right|}{\left|\sum_{j=1}^{n}(\mathbf{x}_j-\boldsymbol{\mu}_0)(\mathbf{x}_j-\boldsymbol{\mu}_0)'\right|}\right)^{n/2} < c_\alpha$$

Result 5.1

$\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$: random sample from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$\Rightarrow T^2$ test is equivalent to the likelihood test of $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ vs. $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$ because

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

43

Proof of Result 5.1

$$\mathbf{A} = \left[\begin{array}{c|c} \sum_{j=1}^{n}(\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})' & \sqrt{n}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0) \\ \hline \sqrt{n}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)' & -1 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right]$$

$$|\mathbf{A}| = |\mathbf{A}_{22}||\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}| = |\mathbf{A}_{11}||\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}|$$

44

Proof of Result 5.1

$$(-1)\left|\sum_{j=1}^{n}(\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})' + n(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\right|$$

$$= \left|\sum_{j=1}^{n}(\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})'\right|\left|-1 - n(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\left(\sum_{j=1}^{n}(\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})'\right)^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)\right|$$

$$\sum_{j=1}^{n}(\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})' + n(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)' = \sum_{j=1}^{n}(\mathbf{x}_j - \boldsymbol{\mu}_0)(\mathbf{x}_j - \boldsymbol{\mu}_0)'$$

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

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

45

Computing T^2 from Determinants

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

$$= \frac{(n-1)\left|\sum_{j=1}^{n}(\mathbf{x}_j - \boldsymbol{\mu}_0)(\mathbf{x}_j - \boldsymbol{\mu}_0)'\right|}{\left|\sum_{j=1}^{n}(\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})'\right|} - (n-1)$$

46

General Likelihood Ratio Method

$\boldsymbol{\theta}$: unknown population parameters, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$

$L(\boldsymbol{\theta})$: likelihood function by random sample

$H_0 : \boldsymbol{\theta} \in \boldsymbol{\Theta}_0$

Rejects H_0 in favor of $H_1 : \boldsymbol{\theta} \notin \boldsymbol{\Theta}_0$ if

$$\Lambda = \frac{\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} L(\boldsymbol{\theta})}{\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} L(\boldsymbol{\theta})} < c$$

47

Result 5.2

when sample size n is large,

$$-2\ln \Lambda = -2\ln\left(\frac{\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} L(\boldsymbol{\theta})}{\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} L(\boldsymbol{\theta})}\right)$$

is approximately a $\chi^2_{\nu-\nu_0}$ random variable, where

$\nu - \nu_0 = (\text{dimension of } \boldsymbol{\Theta}) - (\text{dimension of } \boldsymbol{\Theta}_0)$

48

Outline

- Introduction
- Inferences about a Mean for Univariate Normal Distribution
- The probability of μ_0 as a value for a Normal Population Mean
- Hotelling's T^2 and Likelihood Ratio Tests
- Confidence Regions and Simultaneous Comparison of Component Means

Questions

- How to find the confidence region for the population mean vector of a multivariate normal distribution?
- What are the axes of the confidence ellipsoid?
- What are the simultaneous confidence statements?

Questions

- How to find the confidence interval for a linear combination of multivariate normal random variables?
- How to find the maximum of the t value for all linear combination coefficients?
- How to determine the T^2 intervals? (Result 5.2)

Questions

- What are the difference of trends for t and T^2 intervals as the sample size n increases?
- How to determine simultaneous T^2 intervals?
- How to determine one-at-a-time intervals?

Questions

- What is the Bonferroni inequality?
- How to find the simultaneous Bonferroni intervals?
- What is the trend of the ratio between the length of Bonferroni interval to the length of T^2-Interval with increasing sample size n under different m?

100(1-α)% Confidence Region

$\boldsymbol{\theta}$: unknown population parameters, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$

$R(\mathbf{X})$: region of likely $\boldsymbol{\theta}$ values determined by data array $\mathbf{X}$

100(1-α)% confidence region : R($\mathbf{X}$) where

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

(Region consisting of all $\boldsymbol{\theta}_0$ for which the test will not reject $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ in favor of H_1 at significance level α)

54

100(1-α)% Confidence Region

Univariate Normal Case :

The interval of μ

$$n(\bar{x}-\mu)(s^2)^{-1}(\bar{x}-\mu)\le t_{n-1}^2(\alpha)$$

Multivariate Normal Case :

The ellipsoid determined by all $\boldsymbol{\mu}$ such that

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

55

Axes of the Confidence Ellipsoid

beginning at the center $\bar{\mathbf{x}}$, the axes are

$$\pm\sqrt{\lambda_i}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)}\mathbf{e}_i$$

where $\mathbf{S}\mathbf{e}_i=\lambda_i\mathbf{e}_i$

56

Example 5.3 : Microwave Oven Radiation

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

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

$$\bar{\mathbf{x}}=\begin{bmatrix}0.564\\0.603\end{bmatrix},\quad \mathbf{S}=\begin{bmatrix}0.0144 & 0.0117\\0.0117 & 0.0146\end{bmatrix}$$

$$\mathbf{S}^{-1}=\begin{bmatrix}203.018 & -163.391\\-163.391 & 200.228\end{bmatrix}$$

$\lambda_1=0.026,\quad \mathbf{e}_1'=[0.704\quad 0.710]$

$\lambda_2=0.002,\quad \mathbf{e}_2'=[-0.710\quad 0.704]$

57

Example 5.3 : 95% Confidence Region

$$42[0.564-\mu_1\quad 0.603-\mu_2]\begin{bmatrix}203.018 & -163.391\\-163.391 & 200.228\end{bmatrix}\begin{bmatrix}0.564-\mu_1\\0.603-\mu_2\end{bmatrix}$$

$$\le\frac{2(41)}{40}F_{2,40}(0.05)=6.62$$

$\boldsymbol{\mu}'=[0.562\quad 0.589]$

$$42[0.564-0.562\quad 0.603-0.589]\begin{bmatrix}203.018 & -163.391\\-163.391 & 200.228\end{bmatrix}\begin{bmatrix}0.564-0.562\\0.603-0.589\end{bmatrix}$$

$=1.30\le 6.62$

$\therefore\ \boldsymbol{\mu}$ is in the 95% confidence region.

By this test $H_0:\boldsymbol{\mu}=\begin{bmatrix}0.562\\0.589\end{bmatrix}$ would not be rejected

in favor of $H_1:\boldsymbol{\mu}\ne\begin{bmatrix}0.562\\0.589\end{bmatrix}$ at the significance level $\alpha=0.05$

58

Example 5.3 : 95% Confidence Ellipse for μ

center :

$\bar{\mathbf{x}}'=[0.564\quad 0.603]$

semi - major and semi - minor axes :

$$\sqrt{\lambda_1}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)}=\sqrt{0.026}\sqrt{\frac{2(41)}{42(40)}(3.23)}$$

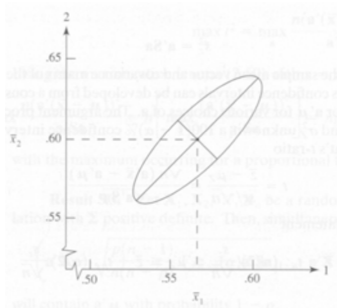
$=0.064$

$$\sqrt{\lambda_2}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)}=\sqrt{0.002}\sqrt{\frac{2(41)}{42(40)}(3.23)}$$

$=0.018$

59

Example 5.3 : 95% Confidence Ellipse for μ



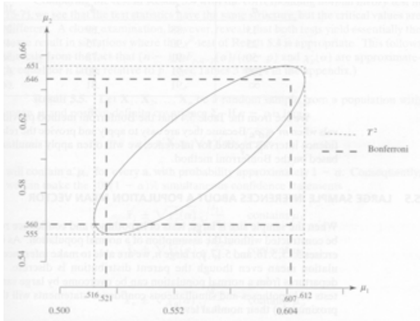
60

Simultaneous Confidence Statements

- Sometimes we need confidence statements about the individual component means
- All if the separate confidence statements should hold simultaneously with a specified high probability

61

Concept of Simultaneously Confidence Statements



62

Confidence Interval of Linear Combination of Variables

$$\mathbf{X} : N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad Z = \mathbf{a}'\mathbf{X}$$

$$\mu_Z = \mathbf{a}'\boldsymbol{\mu}, \quad \sigma_z^2 = \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}, \quad Z : N(\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})$$

$$\bar{z} = \mathbf{a}'\bar{\mathbf{x}}, \quad s_z^2 = \mathbf{a}'\mathbf{S}\mathbf{a}$$

$$t = \frac{\bar{z} - \mu_Z}{s_z / \sqrt{n}} = \frac{\sqrt{n}(\mathbf{a}'\bar{\mathbf{x}} - \mathbf{a}'\boldsymbol{\mu})}{\sqrt{\mathbf{a}'\mathbf{S}\mathbf{a}}}$$

$$|t| \le t_{n-1}^2(\alpha)$$

$$\mathbf{a}'\bar{\mathbf{x}} - t_{n-1}(\alpha)\frac{\sqrt{\mathbf{a}'\mathbf{S}\mathbf{a}}}{\sqrt{n}} \le \mathbf{a}'\boldsymbol{\mu} \le \mathbf{a}'\bar{\mathbf{x}} + t_{n-1}(\alpha)\frac{\sqrt{\mathbf{a}'\mathbf{S}\mathbf{a}}}{\sqrt{n}}$$

63

Maximum t^2 Value for All $\mathbf{a}$

$$\max_a t^2 = \max_a \frac{n(\mathbf{a}'(\bar{\mathbf{x}} - \boldsymbol{\mu}))^2}{\mathbf{a}'\mathbf{S}\mathbf{a}}$$

$$\max_{\mathbf{a}} \frac{n(\mathbf{a}'(\bar{\mathbf{x}} - \boldsymbol{\mu}))^2}{\mathbf{a}'\mathbf{S}\mathbf{a}} = n\left[\max_{\mathbf{a}} \frac{(\mathbf{a}'(\bar{\mathbf{x}} - \boldsymbol{\mu}))^2}{\mathbf{a}'\mathbf{S}\mathbf{a}}\right]$$

$$= n(\bar{\mathbf{x}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}) = T^2$$

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

64

Maximization Lemma

$\mathbf{B}$ positive definite matrix, $\mathbf{d}$ given vector

$$\max_{\mathbf{x} \ne 0} \frac{(\mathbf{x}'\mathbf{d})^2}{\mathbf{x}'\mathbf{B}\mathbf{x}} = \mathbf{d}'\mathbf{B}^{-1}\mathbf{d}$$

maximum attained when $\mathbf{x} = c\mathbf{B}^{-1}\mathbf{d}$ for $c \ne 0$

Proof :

$$(\mathbf{x}'\mathbf{d})^2 \le (\mathbf{x}'\mathbf{B}\mathbf{x})(\mathbf{d}'\mathbf{B}^{-1}\mathbf{d})$$

$$\mathbf{x}'\mathbf{B}\mathbf{x} > 0$$

$$\frac{(\mathbf{x}'\mathbf{d})^2}{\mathbf{x}'\mathbf{B}\mathbf{x}} \le \mathbf{d}'\mathbf{B}^{-1}\mathbf{d}$$

65

Result 5.3: T^2 Interval

$\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$: random sample from $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Simultaneously for all $\mathbf{a}$, the interval (T^2 interval) determined by

$$\mathbf{a}'\bar{\mathbf{x}} - \sqrt{\frac{p(n-1)}{n(n-p)} F_{p,n-p}(\alpha) \mathbf{a}'\mathbf{S}\mathbf{a}} \text{ and}$$

$$\mathbf{a}'\bar{\mathbf{x}} + \sqrt{\frac{p(n-1)}{n(n-p)} F_{p,n-p}(\alpha) \mathbf{a}'\mathbf{S}\mathbf{a}}$$

will contain $\mathbf{a}'\boldsymbol{\mu}$ with probability at least $1 - \alpha$

66

Comparison of t- and T^2-Intervals

n	$t_{n-1}(.025)$	$\sqrt{\frac{(n-1)p}{(n-p)}F_{p,n-p}(.05)}$	
		$p = 4$	$p = 10$
15	2.145	4.14	11.52
25	2.064	3.60	6.39
50	2.010	3.31	5.05
100	1.970	3.19	4.61
∞	1.960	3.08	4.28

67

Simultaneous T^2-Intervals

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

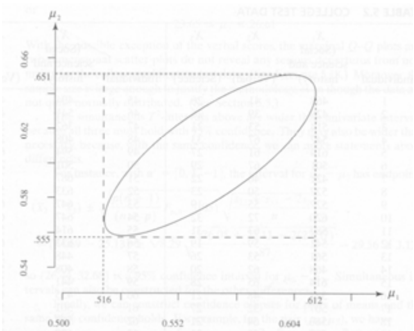
$$\bar{x}_2 - \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{22}}{n}} \le \mu_2 \le \bar{x}_2 + \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{22}}{n}}$$

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

$$\bar{x}_p - \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{pp}}{n}} \le \mu_p \le \bar{x}_p + \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{pp}}{n}}$$

68

Example 5.4: Shadows of the Confidence Ellipsoid



69

Example 5.5

X_1: CLEP score for social science and history

X_2: CQT score for verbal

X_3: CQT score for science

$n = 87$

$$\bar{\mathbf{x}} = \begin{bmatrix} 526.59 \\ 54.69 \\ 25.13 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 5691.34 & 600.51 & 217.25 \\ 600.51 & 126.05 & 23.37 \\ 217.25 & 23.37 & 23.11 \end{bmatrix}$$

70

Example 5.5

$$\frac{p(n-1)}{n-p}F_{p,n-p}(\alpha) = \frac{3(87-1)}{87-3}F_{p,n-p}(0.05) = 8.29$$

$$526.59 - \sqrt{8.29}\sqrt{\frac{5691.34}{87}} \le \mu_1 \le 526.59 + \sqrt{8.29}\sqrt{\frac{5691.34}{87}}$$

$$503.30 \le \mu_1 \le 549.88, \quad 51.22 \le \mu_2 \le 58.16, \quad 23.65 \le \mu_3 \le 26.61$$

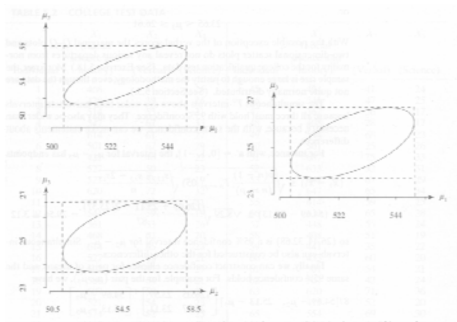
$\mathbf{a}' = [0, 1, -1]$ for $\mu_2 - \mu_3$, end points of its confidence interval are

$$(\bar{x}_2 - \bar{x}_3) \pm \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(0.05)}\sqrt{\frac{s_{22} + s_{33} - 2s_{23}}{n}}$$

i.e., 29.56 ± 3.12 is an at-least 95% confidence interval for $\mu_2 - \mu_3$

71

Example 5.5: Confidence Ellipses for Pairs of Means



72

One-at-a-Time Intervals

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

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

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

$$\bar{x}_p - t_{n-1}(\alpha/2)\sqrt{\frac{s_{pp}}{n}} \le \mu_p \le \bar{x}_p + t_{n-1}(\alpha/2)\sqrt{\frac{s_{pp}}{n}}$$

73

Bonferroni Inequality

C_i : confidence statement about $a_i'\mu$

$$P[C_i \text{true}] = 1 - \alpha_i, \quad i = 1, 2, \cdots, m$$

$$P[\text{all } C_i \text{ true}] = 1 - P[\text{at least one } C_i \text{ false}]$$

$$\ge 1 - \sum_{i=1}^{m} P[C_i \text{ false}] = 1 - \sum_{i=1}^{m}\left(1 - P[C_i \text{ true}]\right)$$

$$= 1 - (\alpha_1 + \alpha_2 + \cdots + \alpha_m)$$

74

Bonferroni Method of Multiple Comparisons

$$\alpha_i = \alpha/m, \quad m = p$$

$$P\left[\bar{x}_i \pm t_{n-1}\left(\frac{\alpha}{2m}\right)\sqrt{\frac{s_{ii}}{n}} \text{ contains } \mu_i, \text{all } i\right] \ge$$

$$1 - \underbrace{\left(\frac{\alpha}{m} + \frac{\alpha}{m} + \cdots + \frac{\alpha}{m}\right)}_{m \text{ terms}} = 1 - \alpha$$

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

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

$$\bar{x}_p - t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{pp}}{n}} \le \mu_p \le \bar{x}_p + t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{pp}}{n}}$$

75

Example 5.6

$$p = 2, \quad \alpha_i = 0.05/2 = 0.025$$

$$t_{41}\left(\frac{0.025}{2}\right) = 2.327$$

$$\bar{x}_1 \pm t_{41}(0.0125)\sqrt{\frac{s_{11}}{n}} = 0.564 \pm 2.327\sqrt{\frac{0.0144}{42}}$$

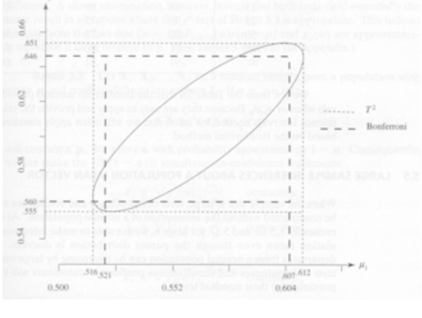
or $0.521 \le \mu_1 \le 0.607$

$$\bar{x}_2 \pm t_{41}(0.0125)\sqrt{\frac{s_{22}}{n}} = 0.603 \pm 2.327\sqrt{\frac{0.0146}{42}}$$

or $0.560 \le \mu_2 \le 0.646$

76

Example 5.6



77

(Length of Bonferroni Interval)/ (Length of T^2-Interval)

	$m = p$		
n	2	4	10
15	.88	.69	.29
25	.90	.75	.48
50	.91	.78	.58
100	.91	.80	.62
∞	.91	.81	.66

78

Outline

- Large Sample Inferences about a Population Mean Vector
- Multivariate Quality Control Charts
- Inferences about Mean Vectors When Some Observations Are Missing
- Difficulties Due to Time Dependence in Multivariate Observations

Questions

- What is the limit distribution of the square of the statistical distance?
- How to reject or accept a null hypothesis when *n-p* is large?
- How to find the confidence interval and the simultaneous confidence statements for large *n-p*? (Result 5.5)

Limit Distribution of the Square of Statistical Distance

$\overline{\mathbf{X}}$: nearly $N_p(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma})$ for large sample size $n >> p$

$n(\overline{\mathbf{X}}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{X}}-\boldsymbol{\mu})$: approximately χ_p^2 for large *n-p*

$\mathbf{S}$ close to $\boldsymbol{\Sigma}$ with high probability when n is large

$\therefore n(\overline{\mathbf{X}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\overline{\mathbf{X}}-\boldsymbol{\mu})$: approximately χ_p^2 for large *n-p*

81

Result 5.4

$\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$: random sample from a population with mean $\boldsymbol{\mu}$ and positive definite covariance $\boldsymbol{\Sigma}$

$n-p$ large

$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ is rejected in favor of $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$, at a level of significance approximately α, if

$n(\overline{\mathbf{x}}-\boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\overline{\mathbf{x}}-\boldsymbol{\mu}_0) > \chi_p^2(\alpha)$

82

Result 5.5

$\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$: random sample from a population with mean $\boldsymbol{\mu}$ and positive definite covariance $\boldsymbol{\Sigma}$

$n-p$ large

$$\mathbf{a}'\overline{\mathbf{x}} \pm \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{\mathbf{a}'\mathbf{S}\mathbf{a}}{n}}$$

will contain $\mathbf{a}'\boldsymbol{\mu}$, for every $\mathbf{a}$, with probability approximately $1-\alpha$

83

Result 5.5

100(1-α)% simultaneous confidence statements

$$\bar{x}_1 - \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{11}}{n}} \le \mu_1 \le \bar{x}_1 + \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{11}}{n}}$$

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

$$\bar{x}_p - \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{pp}}{n}} \le \mu_p \le \bar{x}_p + \sqrt{\chi_p^2(\alpha)}\sqrt{\frac{s_{pp}}{n}}$$

for all pairs (μ_i, μ_k)

$$n\begin{bmatrix}\bar{x}_i-\mu_i & \bar{x}_k-\mu_k\end{bmatrix}\begin{bmatrix}s_{ii} & s_{ik}\\ s_{ik} & s_{kk}\end{bmatrix}^{-1}\begin{bmatrix}\bar{x}_i-\mu_i\\ \bar{x}_k-\mu_k\end{bmatrix} \le \chi_p^2(\alpha)$$

contain (μ_i, μ_k) with confidence (1-α)

84

Example 5.7: Musical Aptitude Profile for 96 Finish Students

Variable	Raw score: Mean ($\bar{x}_i$)	Raw score: 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.

85

Example 5.7: Simultaneous 90% Confidence Limits

$$\bar{x}_i \pm \sqrt{\chi_7^2(0.10)}\sqrt{\frac{s_{ii}}{n}}, \quad \chi_7^2(0.10) = 12.02$$

$$26.06 \le \mu_1 \le 30.14, \quad 24.53 \le \mu_2 \le 28.67$$

$$34.05 \le \mu_3 \le 36.75, \quad 32.39 \le \mu_4 \le 36.01$$

$$22.27 \le \mu_5 \le 24.93, \quad 20.61 \le \mu_6 \le 23.39$$

$$21.27 \le \mu_7 \le 24.13$$

Profile of American students

$$\mu_0' = [31 \quad 27 \quad 34 \quad 31 \quad 23 \quad 22 \quad 22]$$

melody, tempo, meter components are not plausible

86

One-at-a-Time and Bonferroni Confidence Intervals

One-at-a-time confidence intervals

$$\bar{x}_i - z\left(\frac{\alpha}{2}\right)\sqrt{\frac{s_{ii}}{n}} \le \mu_i \le \bar{x}_i + z\left(\frac{\alpha}{2}\right)\sqrt{\frac{s_{ii}}{n}}$$

Bonferroni confidence intervals

$$\bar{x}_i - z\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{ii}}{n}} \le \mu_i \le \bar{x}_i + z\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{ii}}{n}}$$

87

Large-Sample 95% Intervals for Example 5.7

The one-at-a-time confidence intervals use $z(.025) = 1.96$.
The simultaneous Bonferroni intervals use $z(.025/7) = 2.69$.
The simultaneous T^2, or shadows of the ellipsoid, use $\chi_7^2(.05) = 14.07$.

Variable	One-at-a-time Lower	One-at-a-time Upper	Bonferroni Intervals Lower	Bonferroni Intervals Upper	Shadow of Ellipsoid Lower	Shadow of Ellipsoid Upper
X_1 = melody	26.95	29.25	26.52	29.68	25.90	30.30
X_2 = harmony	25.43	27.77	24.99	28.21	24.36	28.84
X_3 = tempo	34.64	36.16	34.35	36.45	33.94	36.86
X_4 = meter	33.18	35.22	32.79	35.61	32.24	36.16
X_5 = phrasing	22.85	24.35	22.57	24.63	22.16	25.04
X_6 = balance	21.21	22.79	20.92	23.08	20.50	23.50
X_7 = style	21.89	23.51	21.59	23.81	21.16	24.24

88

95% Intervals for Example 5.7

The one-at-a-time confidence intervals use $t_{95}(.025) = 1.99$.
The simultaneous Bonferroni intervals use $t_{95}(.025/7) = 2.75$.
The simultaneous T^2, or shadows of the ellipsoid, use $F_{7,89}(.05) = 2.11$.

Variable	One-at-a-time Lower	One-at-a-time Upper	Bonferroni Intervals Lower	Bonferroni Intervals Upper	Shadow of Ellipsoid Lower	Shadow of Ellipsoid Upper
X_1 = melody	26.93	29.27	26.48	29.72	25.76	30.44
X_2 = harmony	25.41	27.79	24.96	28.24	24.23	28.97
X_3 = tempo	34.63	36.17	34.33	36.47	33.85	36.95
X_4 = meter	33.16	35.24	32.76	35.64	32.12	36.28
X_5 = phrasing	22.84	24.36	22.54	24.66	22.07	25.13
X_6 = balance	21.20	22.80	20.90	23.10	20.41	23.59
X_7 = style	21.88	23.52	21.57	23.83	21.07	24.33

89

Outline

- Large Sample Inferences about a Population Mean Vector
- Multivariate Quality Control Charts
- Inferences about Mean Vectors When Some Observations Are Missing
- Difficulties Due to Time Dependence in Multivariate Observations

Questions

- What is the control chart?
- How to monitor a sample for stability?
- How to draw an quality control ellipse?
- How to draw an $\bar{X}$ chart?
- How to draw a T^2 chart

Questions

- What is the distribution of T^2 for an individual future observation? (Result 5.6)
- How to find the control region for an individual future observation?
- How to draw T^2-chart for future observations?
- How to draw control chart based on subsample means?

Control Chart

- Represents collected data to evaluate the capabilities and stability of the process
- Identify occurrences of special causes of variation that come from outside of the usual process

93

Example 5.8: Overtime Hours for a Police Department

x_1 Legal Appearances Hours	x_2 Extraordinary Event Hours	x_3 Holdover Hours	x_4 COA[1] Hours	x_5 Meeting Hours
3387	2200	1181	14,861	236
3109	875	3532	11,367	310
2670	957	2502	13,329	1182
3125	1758	4510	12,328	1208
3469	868	3032	12,847	1385
3120	398	2130	13,979	1053
3671	1603	1982	13,528	1046
4531	523	4675	12,699	1100
3678	2034	2354	13,534	1349
3238	1136	4606	11,609	1150
3135	5326	3044	14,189	1216
5217	1658	3340	15,052	660
3728	1945	2111	12,236	299
3506	344	1291	15,482	206
3824	807	1365	14,900	239
3516	1223	1175	15,078	161

[1]Compensatory overtime allowed.

94

Example 5.8 Univariate Control Chart

95

Monitoring a Sample for Stability

$\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$: independently distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$$E(\mathbf{X}_j - \bar{\mathbf{X}}) = 0$$

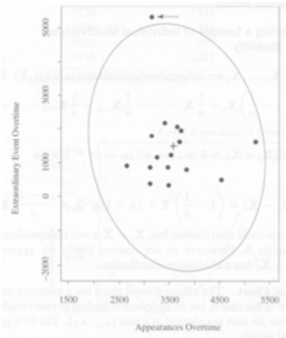
$$\mathbf{Cov}(\mathbf{X}_j - \bar{\mathbf{X}}) = \frac{n-1}{n}\boldsymbol{\Sigma}$$

$\mathbf{X}_j - \bar{\mathbf{X}}$ is normal, but is not independent of $\mathbf{S}$

Approximate $(\mathbf{X}_j - \bar{\mathbf{X}})'\mathbf{S}^{-1}(\mathbf{X}_j - \bar{\mathbf{X}})$ as a chi-square distribution

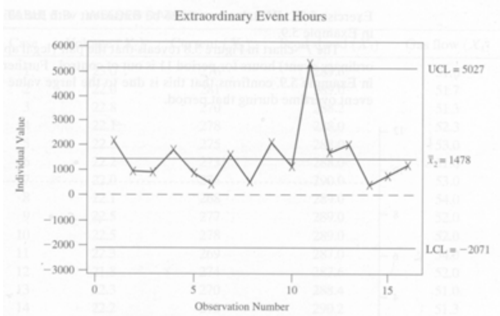
96

Example 5.9: 99% Ellipse Format Chart



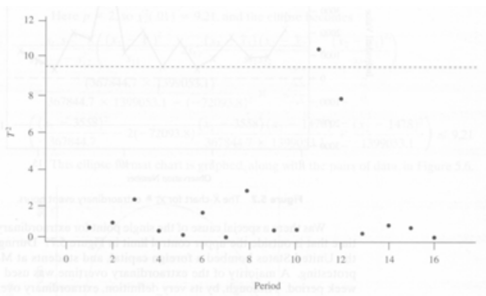
97

Example 5.9: $\overline{X}$-Chart for X_2



98

Example 5.10: T^2 Chart for X_1 and X_2



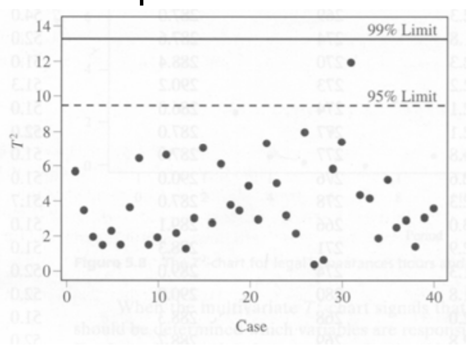
99

Example 5.11: Robotic Welders

X_1 : Voltage (volts)
X_2 : Current (amps)
X_3 : Feed speed (in/min)
X_4 : (inert) Gas flow (cfm)
Normal assumption is reasonable
No appreciable serial correlation for successive observations on each variable

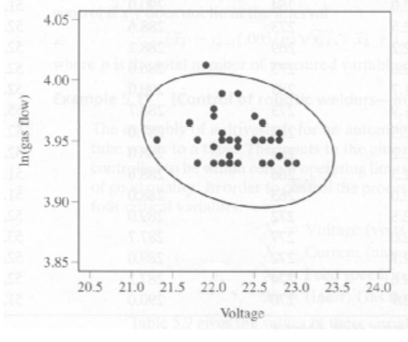
100

Example 5.11: T^2 Chart



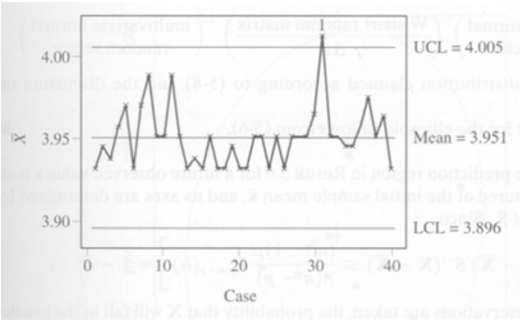
101

Example 5.11: 99% Quality Control Ellipse for ln(Gas flow) and voltage



102

Example 5.11: $\bar{X}$ -Chart for ln(Gas flow)



103

Control Regions for Future Individual Observations

- Set for future observations from collected data when process is stable
- Forecast or prediction region
  - in which a future observation is expected to lie

104

Result 5.6

$\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$: independently as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$\mathbf{X}$: future observation from the same distribution

$$T^2 = \frac{n}{n+1}\left(\mathbf{X}-\overline{\mathbf{X}}\right)'\mathbf{S}^{-1}\left(\mathbf{X}-\overline{\mathbf{X}}\right)$$

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

105

Proof of Result 5.6

$$E(\mathbf{X}-\overline{\mathbf{X}}) = 0$$

$$\mathbf{Cov}(\mathbf{X}-\overline{\mathbf{X}}) = \mathbf{Cov}(\mathbf{X}) + \mathbf{Cov}(\overline{\mathbf{X}}) = \boldsymbol{\Sigma} + \frac{1}{n}\boldsymbol{\Sigma}$$

$$= \frac{n+1}{n}\boldsymbol{\Sigma}$$

$$\sqrt{\frac{n}{n+1}}\left(\mathbf{X}-\overline{\mathbf{X}}\right) : N_p(0, \boldsymbol{\Sigma}), \quad \mathbf{S} : W_{p,n-1}(\boldsymbol{\Sigma})$$

$$\Rightarrow \frac{n}{n+1}\left(\mathbf{X}-\overline{\mathbf{X}}\right)'\mathbf{S}^{-1}\left(\mathbf{X}-\overline{\mathbf{X}}\right) : \frac{(n-1)p}{n-p} F_{p,n-p}$$

106

Result 4.8

$\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n$: mutually independent

$\mathbf{X}_j : N_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma})$

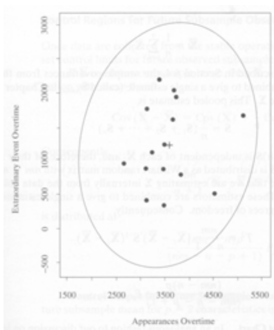
$$\mathbf{V}_1 = c_1\mathbf{X}_1 + c_2\mathbf{X}_2 + \cdots + c_n\mathbf{X}_n : N_p\left(\sum_{j=1}^{n} c_j\boldsymbol{\mu}_j, (\sum_{j=1}^{n} c_j^2)\boldsymbol{\Sigma}\right)$$

$\mathbf{V}_2 = b_1\mathbf{X}_1 + b_2\mathbf{X}_2 + \cdots + b_n\mathbf{X}_n$ and $\mathbf{V}_1$ are joint normal

with covariance matrix $\begin{bmatrix} (\sum_{j=1}^{n} c_j^2)\boldsymbol{\Sigma} & (\mathbf{b}'\mathbf{c})\boldsymbol{\Sigma} \\ (\mathbf{b}'\mathbf{c})\boldsymbol{\Sigma} & (\sum_{j=1}^{n} b_j^2)\boldsymbol{\Sigma} \end{bmatrix}$

107

Example 5.12 Control Ellipse



108

T^2-Chart for Future Observations

Plot

$$T^2 = \frac{n}{n+1}(\mathbf{x}-\bar{\mathbf{x}})'\mathbf{S}^{-1}(\mathbf{x}-\bar{\mathbf{x}})$$

in time order

$$\text{LCL} = 0$$

$$\text{UCL} = \frac{(n-1)p}{(n-p)} F_{p,n-p}(0.05)$$

109

Control Chart Based on Subsample Means

Process : $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $m > 1$ units be sampled at the same time

$\bar{\mathbf{X}}_j$: subsample mean at time j, $\bar{\bar{\mathbf{X}}} = \frac{1}{n}\sum_{j=1}^{n}\bar{\mathbf{X}}_j$

$$\bar{\mathbf{X}}_j - \bar{\bar{\mathbf{X}}} : N_p(0, \frac{(n-1)}{nm}\Sigma)$$

$$\because \text{Cov}(\bar{\mathbf{X}}_j - \bar{\bar{\mathbf{X}}})$$

$$= \text{Cov}\left\{(1-\frac{1}{n})\bar{\mathbf{X}}_j + \frac{1}{n}\bar{\mathbf{X}}_1 + \cdots + \frac{1}{n}\bar{\mathbf{X}}_{j-1} + \frac{1}{n}\bar{\mathbf{X}}_{j+1} + \cdots + \frac{1}{n}\bar{\mathbf{X}}_n\right\}$$

$$= \left(1-\frac{1}{n}\right)^2 \text{Cov}(\bar{\mathbf{X}}_j) + \frac{n-1}{n^2}\text{Cov}(\bar{\mathbf{X}}_1)$$

$$= \left\{\left(1-\frac{1}{n}\right)^2 + \frac{n-1}{n^2}\right\}\frac{1}{m}\boldsymbol{\Sigma} = \frac{(n-1)}{nm}\boldsymbol{\Sigma}$$

110

Control Chart Based on Subsample Means

$$\mathbf{S} = \frac{1}{n}(\mathbf{S}_1 + \mathbf{S}_2 + \cdots + \mathbf{S}_n) : W_{p,nm-n}(\Sigma)$$

$$\sqrt{\frac{nm}{n-1}}(\bar{\mathbf{X}}_j - \bar{\bar{\mathbf{X}}}) : N_p(0, \boldsymbol{\Sigma})$$

$$\Rightarrow T^2 = \frac{nm}{n-1}(\bar{\mathbf{X}}_j - \bar{\bar{\mathbf{X}}})\mathbf{S}^{-1}(\bar{\mathbf{X}}_j - \bar{\bar{\mathbf{X}}}) :$$

$$\frac{(nm-n)p}{(nm-n-p+1)} F_{p,nm-n-p+1}$$

111

Control Regions for Future Subsample Observations

Process : $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $m > 1$ units be sampled at the same time

$\bar{\mathbf{X}}$: future subsample mean, $\bar{\bar{\mathbf{X}}} = \frac{1}{n}\sum_{j=1}^{n}\bar{\mathbf{X}}_j$

$$\bar{\mathbf{X}}_j - \bar{\bar{\mathbf{X}}} : N_p(\mathbf{0}, \frac{(n+1)}{nm}\boldsymbol{\Sigma})$$

$$\because \text{Cov}(\bar{\mathbf{X}} - \bar{\bar{\mathbf{X}}}) = \text{Cov}(\bar{\mathbf{X}}) + \text{Cov}\left\{\frac{1}{n}\bar{\mathbf{X}}_1 + \cdots + \frac{1}{n}\bar{\mathbf{X}}_n\right\}$$

$$= \text{Cov}(\bar{\mathbf{X}}) + \frac{1}{n}\text{Cov}(\bar{\mathbf{X}}_1) = \frac{n+1}{nm}\boldsymbol{\Sigma}$$

$$\frac{nm}{n+1}(\bar{\mathbf{X}} - \bar{\bar{\mathbf{X}}})\mathbf{S}^{-1}(\bar{\mathbf{X}} - \bar{\bar{\mathbf{X}}}) : \frac{(nm-n)p}{(nm-n-p+1)} F_{p,nm-n-p+1}$$

112

Control Chart Based on Subsample Means

$$\mathbf{S} = \frac{1}{n}(\mathbf{S}_1 + \mathbf{S}_2 + \cdots + \mathbf{S}_n) : W_{p,nm-n}(\Sigma)$$

$$\sqrt{\frac{nm}{n-1}}(\bar{\mathbf{X}}_j - \bar{\bar{\mathbf{X}}}) : N_p(0, \boldsymbol{\Sigma})$$

$$\Rightarrow T^2 = \frac{nm}{n-1}(\bar{\mathbf{X}}_j - \bar{\bar{\mathbf{X}}})\mathbf{S}^{-1}(\bar{\mathbf{X}}_j - \bar{\bar{\mathbf{X}}}) :$$

$$\frac{(nm-n)p}{(nm-n-p+1)} F_{p,nm-n-p+1}$$

113

Outline

- Large Sample Inferences about a Population Mean Vector
- Multivariate Quality Control Charts
- Inferences about Mean Vectors When Some Observations Are Missing
- Difficulties Due to Time Dependence in Multivariate Observations

Questions

- What is the EM algorithm?
- What are sufficient statistics?
- What are sufficient statistics for multivariate normal distribution?
- How to estimate the mean and variance-covariance matrix for multivariate normal distribution when some observations are missing?

EM Algorithm

- Prediction step
  - Given some estimate of the unknown parameters, predict the contribution of the missing observations to the sufficient statistics
- Estimation step
  - Use the predicted statistics to compute a revised estimate of the parameters
- Cycle from one step to the other

116

Complete-Data Sufficient Statistics

$$\mathbf{T}_1 = \sum_{j=1}^{n} \mathbf{X}_j = n\overline{\mathbf{X}}$$

$$\mathbf{T}_2 = \sum_{j=1}^{n} \mathbf{X}_j \mathbf{X}_j' = (n-1)\mathbf{S} + n\overline{\mathbf{X}}\overline{\mathbf{X}}'$$

117

Prediction Step for Multivariate Normal Distribution

$\mathbf{x}_j^{(1)}$: missing components of $\mathbf{x}_j$

$\mathbf{x}_j^{(2)}$: available components of $\mathbf{x}_j$

Given estimates $\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}$

$$\tilde{\mathbf{x}}_j^{(1)} = E\left(\mathbf{X}_j^{(1)} \mid \mathbf{x}_j^{(2)}; \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}\right) = \tilde{\boldsymbol{\mu}}^{(1)} + \tilde{\boldsymbol{\Sigma}}_{12}\tilde{\boldsymbol{\Sigma}}_{22}^{-1}\left(\mathbf{x}_j^{(2)} - \tilde{\boldsymbol{\mu}}^{(2)}\right)$$

$$\widetilde{\mathbf{x}_j^{(1)}\mathbf{x}_j^{(1)'}} = E\left(\mathbf{X}_j^{(1)}\mathbf{X}_j^{(1)'} \mid \mathbf{x}_j^{(2)}; \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}\right)$$

$$= \tilde{\boldsymbol{\Sigma}}_{11} - \tilde{\boldsymbol{\Sigma}}_{12}\tilde{\boldsymbol{\Sigma}}_{22}^{-1}\tilde{\boldsymbol{\Sigma}}_{21} + \tilde{\mathbf{x}}_j^{(1)}\tilde{\mathbf{x}}_j^{(1)'}$$

$$\widetilde{\mathbf{x}_j^{(1)}\mathbf{x}_j^{(2)'}} = E\left(\mathbf{X}_j^{(1)}\mathbf{X}_j^{(2)'} \mid \mathbf{x}_j^{(2)}; \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}\right) = \tilde{\mathbf{x}}_j^{(1)}\tilde{\mathbf{x}}_j^{(2)'}$$

118

Result 4.6

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ --- \\ \mathbf{X}_2 \end{bmatrix} : N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ --- \\ \boldsymbol{\mu}_2 \end{bmatrix},$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & | & \boldsymbol{\Sigma}_{12} \\ --- & + & --- \\ \boldsymbol{\Sigma}_{21} & | & \boldsymbol{\Sigma}_{22} \end{bmatrix}, \quad |\boldsymbol{\Sigma}_{22}| > 0 \Rightarrow$$

conditional distribution of $\mathbf{X}_1$ given $\mathbf{X}_2 = \mathbf{x}_2$ is normal with mean $= \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{1}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$ and covariance $= \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$

119

Estimation Step for Multivariate Normal Distribution

Compute the revised maximum likelihood estimates

$$\tilde{\boldsymbol{\mu}} = \frac{\tilde{\mathbf{T}}_1}{n}$$

$$\tilde{\boldsymbol{\Sigma}} = \frac{\tilde{\mathbf{T}}_2}{n} - \tilde{\boldsymbol{\mu}}\tilde{\boldsymbol{\mu}}'$$

120

Example 5.13: EM Algorithm

$$X=\begin{bmatrix} - & 0 & 3 \\ 7 & 2 & 6 \\ 5 & 1 & 2 \\ - & - & 5 \end{bmatrix},\quad \tilde{\mu}_1=6,\quad \tilde{\mu}_2=1,\quad \tilde{\mu}_3=4$$

$$\tilde{\sigma}_{11}=\frac{(6-6)^2+(7-6)^2+(5-6)^2+(6-6)^2}{4}=\frac{1}{2}$$

$$\tilde{\sigma}_{22}=\frac{1}{2},\quad \tilde{\sigma}_{33}=\frac{5}{2},\quad \tilde{\sigma}_{12}=\frac{1}{4},\quad \tilde{\sigma}_{23}=\frac{3}{4},\quad \tilde{\sigma}_{13}=1$$

Example 5.13: Prediction Step

$$\tilde{\mu}=\begin{bmatrix} \tilde{\mu}_1 \\ -- \\ \tilde{\mu}_2 \\ \tilde{\mu}_3 \end{bmatrix}=\begin{bmatrix} \tilde{\boldsymbol{\mu}}^{(1)} \\ -- \\ \tilde{\boldsymbol{\mu}}^{(2)} \end{bmatrix}\quad \tilde{\boldsymbol{\Sigma}}=\begin{bmatrix} \tilde{\sigma}_{11} & | & \tilde{\sigma}_{12} & \tilde{\sigma}_{13} \\ -- & + & -- & -- \\ \tilde{\sigma}_{12} & | & \tilde{\sigma}_{22} & \tilde{\sigma}_{23} \\ \tilde{\sigma}_{13} & | & \tilde{\sigma}_{23} & \tilde{\sigma}_{33} \end{bmatrix}=\begin{bmatrix} \tilde{\boldsymbol{\Sigma}}_{11} & | & \tilde{\boldsymbol{\Sigma}}_{12} \\ -- & + & -- \\ \tilde{\boldsymbol{\Sigma}}_{21} & | & \tilde{\boldsymbol{\Sigma}}_{22} \end{bmatrix}$$

$$\tilde{x}_{11}=\tilde{\mu}_1+\tilde{\boldsymbol{\Sigma}}_{12}\tilde{\boldsymbol{\Sigma}}_{22}^{-1}[x_{12}-\tilde{\mu}_2 \quad x_{13}-\tilde{\mu}_3]'=5.73$$

$$\widetilde{x_{11}^2}=\tilde{\sigma}_{11}-\tilde{\boldsymbol{\Sigma}}_{12}\tilde{\boldsymbol{\Sigma}}_{22}^{-1}\tilde{\boldsymbol{\Sigma}}_{21}+\tilde{x}_{11}^2=32.99$$

$$\widetilde{x_{11}[x_{12} \quad x_{13}]}=\tilde{x}_{11}[x_{12} \quad x_{13}]=[0 \quad 17.18]$$

Example 5.13: Prediction Step

$$\tilde{\mu}=\begin{bmatrix} \tilde{\mu}_1 \\ \tilde{\mu}_2 \\ -- \\ \tilde{\mu}_3 \end{bmatrix}=\begin{bmatrix} \tilde{\boldsymbol{\mu}}^{(1)} \\ -- \\ \tilde{\boldsymbol{\mu}}^{(2)} \end{bmatrix}\quad \tilde{\boldsymbol{\Sigma}}=\begin{bmatrix} \tilde{\sigma}_{11} & \tilde{\sigma}_{12} & | & \tilde{\sigma}_{13} \\ \tilde{\sigma}_{12} & \tilde{\sigma}_{22} & | & \tilde{\sigma}_{23} \\ -- & -- & + & -- \\ \tilde{\sigma}_{13} & \tilde{\sigma}_{23} & | & \tilde{\sigma}_{33} \end{bmatrix}=\begin{bmatrix} \tilde{\boldsymbol{\Sigma}}_{11} & | & \tilde{\boldsymbol{\Sigma}}_{12} \\ -- & + & -- \\ \tilde{\boldsymbol{\Sigma}}_{21} & | & \tilde{\boldsymbol{\Sigma}}_{22} \end{bmatrix}$$

$$\widetilde{\begin{bmatrix} x_{41} \\ x_{42} \end{bmatrix}}=\begin{bmatrix} \tilde{\mu}_1 \\ \tilde{\mu}_2 \end{bmatrix}+\tilde{\boldsymbol{\Sigma}}_{12}\tilde{\boldsymbol{\Sigma}}_{22}^{-1}(x_{43}-\tilde{\mu}_3)=\begin{bmatrix} 6.4 \\ 1.3 \end{bmatrix}$$

$$\begin{bmatrix} \widetilde{x_{41}^2} & \widetilde{x_{41}x_{42}} \\ \widetilde{x_{41}x_{42}} & \widetilde{x_{42}^2} \end{bmatrix}=\begin{bmatrix} 41.06 & 8.27 \\ 8.27 & 1.97 \end{bmatrix},\quad \widetilde{\begin{bmatrix} x_{41} \\ x_{42} \end{bmatrix} x_{43}}=\begin{bmatrix} 32.0 \\ 6.5 \end{bmatrix}$$

Example 5.13: Prediction Step

$$\tilde{\mathbf{T}}_1=\begin{bmatrix} \tilde{x}_{11}+x_{21}+x_{31}+\tilde{x}_{41} \\ x_{12}+x_{22}+x_{32}+\tilde{x}_{42} \\ x_{13}+x_{23}+x_{33}+x_{43} \end{bmatrix}=\begin{bmatrix} 24.13 \\ 4.30 \\ 16.00 \end{bmatrix}$$

$$\tilde{\mathbf{T}}_2=\begin{bmatrix} 148.05 & 27.27 & 101.18 \\ 27.27 & 6.97 & 20.50 \\ 101.18 & 20.50 & 74.00 \end{bmatrix}$$

Example 5.13: Estimation Step

$$\tilde{\boldsymbol{\mu}}=\frac{1}{n}\tilde{\mathbf{T}}_1=\begin{bmatrix} 6.03 \\ 1.08 \\ 4.00 \end{bmatrix}$$

$$\tilde{\boldsymbol{\Sigma}}=\frac{1}{n}\tilde{\mathbf{T}}_2-\tilde{\boldsymbol{\mu}}\tilde{\boldsymbol{\mu}}'=\begin{bmatrix} 0.61 & 0.33 & 1.17 \\ 0.33 & 0.59 & 0.83 \\ 1.17 & 0.83 & 2.50 \end{bmatrix}$$

Outline

- Large Sample Inferences about a Population Mean Vector
- Multivariate Quality Control Charts
- Inferences about Mean Vectors When Some Observations Are Missing
- Difficulties Due to Time Dependence in Multivariate Observations

Question

- What is the confidence region for a time-varying multivariate normal distribution following AR(1) model?

Time Dependence in Observations

$\mathbf{X}_t - \boldsymbol{\mu} = \boldsymbol{\Phi}(\mathbf{X}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon}_t \quad$: AR(1) model

$\boldsymbol{\Phi} = \phi \mathbf{I}, \quad |\phi| < 1$

nominal 95% confidence interval

$\{\text{all } \boldsymbol{\mu} \text{ such that } n(\overline{\mathbf{X}} - \boldsymbol{\mu})' \mathbf{S}^{-1} (\overline{\mathbf{X}} - \boldsymbol{\mu}) \le \chi_p^2(0.05)\}$

actual coverage probability

$P[\chi_p^2 \le (1-\phi)(1+\phi)^{-1} \chi_p^2(0.05)]$

128

Coverage Probability of the 95% Confidence Ellipsoid

		ϕ			
		−.25	0	.25	.5
p	1	.989	.950	.871	.742
	2	.993	.950	.834	.632
	5	.998	.950	.751	.405
	10	.999	.950	.641	.193
	15	1.000	.950	.548	.090

129