What to do today?

Part I. Introduction and Preparation

Part II. Inference under Multivariate Normal Distribution (Chp 4-7)

II.1 Multivariate Normal Distribution (Chp 4)

II.2 Inferences on Mean Vector (Chp 5)

II.2.1 Introduction (Chp5.1-2a)

II.2.2 Hotelling's T^2 and Likelihood Ratio Test (Chp5.2b-3)

II.2.3 Confidence Regions and Simultaneous Comparisons of Component Means (Chp5.4)

II.2.4 Large Sample Inference on Population Mean (Chp5.5)

II.2.5 Some Related Topics* (Chp5.6-8)

II.3 Comparisons of Several Mean Vector (Chp 6)

II.4 Multivariate Linear Regression (Chp 7)

Part II.1.1 Multivariate Normal Distribution, Review

Multivariate normal distribution A p-dim r.v. $\mathbf{X}$ has a *normal* distribution if its pdf

$$f(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \frac{1}{\sqrt{|2\pi\boldsymbol{\Sigma}|}} \exp\Big\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{'}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Big\}, \quad -\infty < \mathbf{x} < \infty,$$

where $\boldsymbol{\Sigma}$ is positive definite. Denote it by $\mathbf{X} \sim MN_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$.

Shape of Multivariate Normal Density If $\mathbf{X} \sim BN(\boldsymbol{\mu},\boldsymbol{\Sigma})$, its pdf $f(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma})$ is a constant for all $\mathbf{x} = (x_1, x_2)^{'}$ satisfy $(\mathbf{x}-\boldsymbol{\mu})^{'}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = c^2$. (It defines an ellipse centered at $\boldsymbol{\mu} = (\mu_1, \mu_2)^{'}$, and with axes $\pm c\sqrt{\lambda_j}\mathbf{e}_j$.)

$(\mathbf{X}-\boldsymbol{\mu})^{'}\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}) \sim \chi^2_p$, the Chi-square distribution with degree of freedom p.

II.1.2 Estimation of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, Review

Consider $\mathbf{X} \sim MN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$: what are $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$?

Suppose $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are iid observations on $\mathbf{X}$, how to use the data to estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$?

II.1.2A. By Method of Moments: The *MME* $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}} = \sum_{i=1}^{n} \mathbf{X}_i \big/ n$, the sample mean, and $\hat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})' = \frac{n-1}{n} \mathbf{S}$.

II.1.2B. By Maximum Likelihood Estimation The *MLE* $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}$ and $\hat{\boldsymbol{\Sigma}} = \frac{n-1}{n} \mathbf{S}$ are the same as the *MME*.

II.2.2 Hotelling's T^2 and Likelihood Ratio Test (Chp5.2b-3), Review

Suppose $\mathbf{X} \sim MN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with iid observations $\mathbf{X}_1, \ldots, \mathbf{X}_n$

To test on $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ vs H_1 : otherwise when $\boldsymbol{\Sigma}$ is known

- *Test statistic.* Note that $\mathbf{Z} = \sqrt{n}\boldsymbol{\Sigma}^{-1/2}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0) \sim MN(\mathbf{0}, \mathbf{I})$ under H_0. Consider to use

$$W = \mathbf{Z}'\mathbf{Z} = (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)' \left(\frac{\boldsymbol{\Sigma}}{n}\right)^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0) \sim \chi^2(p) \quad \text{under} \quad H_0$$

- *Rejection region.* With pre-determined α, $\mathcal{R} = \left\{w : w > \chi^2_\alpha(p)\right\}$
- *Making decision.* Reject H_0 if $W_{obs} \in \mathcal{R}$; otherwise, accept H_0.

Wald-test procedure.

Suppose $\mathbf{X} \sim MN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with iid observations $\mathbf{X}_1, \ldots, \mathbf{X}_n$

To test on $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ vs H_1 : otherwise when $\boldsymbol{\Sigma}$ is unknown

- *Test statistic.* Consider $\mathbf{T} = \sqrt{n}\mathbf{S}^{-1/2}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)$ and under H_0

$$W = \mathbf{T}'\mathbf{T} = (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)' \left(\frac{\mathbf{S}}{n}\right)^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0) \sim T^2(p, n-1)$$

- *Rejection region.* With pre-determined α, $\mathcal{R} = \left\{w : w > T^2_\alpha(p, n-1)\right\}$
- *Making decision.* Reject H_0 if $W_{obs} \in \mathcal{R}$; otherwise, accept H_0.

Hotelling's T^2-test procedure.

Suppose $\mathbf{X} \sim MN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with iid observations $\mathbf{X}_1, \ldots, \mathbf{X}_n$
To test on $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ vs H_1 : otherwise when $\boldsymbol{\Sigma}$ is unknown by the *likelihood ratio test*

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

- *Test statistic.*
$$\Lambda = \frac{L(\boldsymbol{\mu}_0, \hat{\boldsymbol{\Sigma}}_0)}{L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})} = \left(\frac{|\hat{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}_0|}\right)^{n/2}$$
  $\Lambda^{2/n}$ is called *Wilks's lambda* and
$$\Lambda^{2/n} = \left(1 + \frac{T^2}{n-1}\right)^{-1}$$
  under H_0.
- *Rejection region.* With pre-determined α, $\mathcal{R} = \{\lambda : \lambda < c\}$ with c determined by $P_{H_0}(\Lambda < c) = \alpha$.
- *Making decision.* Reject H_0 if $\Lambda_{obs} \in \mathcal{R}$; otherwise, accept H_0.

Likelihood ratio test procedure

Why not just perform multiple univariate t-tests?

Suppose $\mathbf{X} \sim MN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and we want to test

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0.$$

- A natural idea: test each component μ_j using a univariate t-test.
- Reject H_0 if *any* of the component-wise tests rejects.

A simulation example

- Generate repeated random samples from a multivariate normal distribution under $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$.
- For each sample:
  - perform p separate univariate t-tests for $\mu_1, \ldots, \mu_p$, and reject if any test rejects;
  - perform a single Hotelling's T^2 test for $\boldsymbol{\mu}$.
- Compare how often each procedure rejects H_0.

II.2.3A Confidence Regions (Chp5.4)

Consider $\mathbf{X} \sim MN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$: what are $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$?

Suppose $\mathbf{X}_1, \dots, \mathbf{X}_n$ are iid observations on $\mathbf{X}$, how to use the data to estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$?

The $100(1-\alpha)\%$ **confidence region** for $\boldsymbol{\mu}$ is

$$R(\mathbf{x}_1, \dots, \mathbf{x}_n) = \Big\{\boldsymbol{\mu} : \frac{(n-p)t^2}{(n-1)p} \leq F_{p,n-p}(\alpha)\Big\}, \ \ t^2 = (\bar{\mathbf{x}} - \boldsymbol{\mu})' \Big(\frac{\mathbf{S}}{n}\Big)^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu})$$

since

$$P\Big\{R(\mathbf{x}_1, \dots, \mathbf{x}_n) \ni \text{ the true } \boldsymbol{\mu}\Big\} = 1 - \alpha$$

$$\Longleftrightarrow P\Big\{(\bar{\mathbf{X}} - \boldsymbol{\mu})' \Big(\frac{\mathbf{S}}{n}\Big)^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \leq \frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha)\Big\} = 1 - \alpha.$$

Remarks.

- The $100(1-\alpha)\%$ confidence region is the collection of $\boldsymbol{\mu}_0$:

$\{\boldsymbol{\mu}_0$: "$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$" cannot be rejected by T^2-*test* with type I error rate of $\alpha\}$.

- $R(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is an ellipsoid centered at $\bar{\mathbf{x}}$, the axes of the confidence ellipsoid are

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

with $(\lambda_i, \mathbf{e}_i)$ are the eigenvalues and eigenvectors of $\mathbf{S}\big/n$.

- To determine whether a given $\boldsymbol{\mu}$ is in the region or not, compute the *squared generalized distance* between the value and $\bar{\mathbf{x}}$: $(\bar{\mathbf{x}}-\boldsymbol{\mu})'\left(\frac{\mathbf{S}}{n}\right)^{-1}(\bar{\mathbf{x}}-\boldsymbol{\mu})$ and then compare it with $\frac{(n-1)p}{(n-p)}F_{p,n-p}(\alpha)$.

Example. (Data of textbook p215, Table 5.1, Perspiration from 20 healthy females). Consider first two variables:

$$x_1 = (\text{Sweat rate}), \quad x_2 = (\text{Sodium})$$

II.2.3B Simultaneous Comparison of Component Means

Consider $\mathbf{X} \sim MN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Suppose $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are iid observations on $\mathbf{X}$, how to use the data to estimate $\mathbf{a}^{'}\boldsymbol{\mu}$?

Various $\mathbf{a}^{'}\boldsymbol{\mu}$:

- $\mu_1 = \mathbf{a}^{'}\boldsymbol{\mu}$ with $\mathbf{a}^{'} = (1, 0, \ldots, 0)$
- $\mu_1 - \mu_2 = \mathbf{a}^{'}\boldsymbol{\mu}$ with $\mathbf{a}^{'} = (1, -1, \ldots, 0)$

Note that

$$Z = \mathbf{a}^{'}\bar{\mathbf{X}} \sim N(\mathbf{a}^{'}\boldsymbol{\mu}, \frac{1}{n}\mathbf{a}^{'}\boldsymbol{\Sigma}\mathbf{a}), \quad s_Z^2 = \mathbf{a}^{'}\mathbf{S}\mathbf{a}.$$

The t-statistic:

$$\frac{\mathbf{a}^{'}\bar{\mathbf{X}} - \mathbf{a}^{'}\boldsymbol{\mu}}{\sqrt{\mathbf{a}'\mathbf{S}\mathbf{a}/n}} \sim t_{n-1}$$

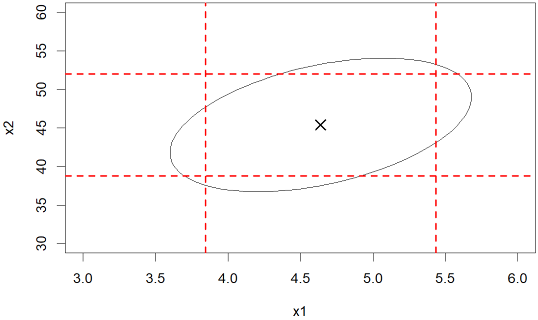
Thus, a $100(1-\alpha)\%$ CI for $\mathbf{a}^{'}\boldsymbol{\mu}$ is

$$\mathbf{a}^{'}\bar{\mathbf{x}} \pm t_{n-1}(\alpha/2)\sqrt{\mathbf{a}'\mathbf{S}\mathbf{a}/n}.$$

Example. (cont'd) Consider $\mathbf{a}_1 = (1,0)'$ and $\mathbf{a}_2 = (0,1)'$. They CIs are

$$\bar{x}_1 \pm t_{20-1}(0.05/2)\sqrt{\frac{2.87}{20}} = (3.846,\ 5.434).$$

$$\bar{x}_2 \pm t_{20-1}(0.05/2)\sqrt{\frac{199.78}{20}} = (38.785,\ 52.015).$$



Why the confidence region and the two confidence intervals are different?

Denote the CIs by μ_1, μ_2 are (L_1, R_1) and (L_2, R_2).
Let $A =$ "$\mu_1 \in (L_1, R_1)$" and $B =$ "$\mu_2 \in (L_2, R_2)$"

- $P(A) = .95$ and $P(B) = .95$.
- $P(A \bigcap B) =?$
  If A and B are indpt,
  $P(A \bigcap B) = P(A)P(B) = (.95)^2 = .9025 < .95$.
- In general, if $P(A) = 1 - \alpha$ and $P(B) = 1 - \alpha$.

$$P(A \bigcap B) = 1 - P(A^c \bigcup B^c) \geq 1 - \left[P(A^c) + P(B^c)\right] = 1 - 2\alpha$$

Correction?

How to construct $100(1-\alpha)\%$ simultaneous confidence intervals for m quantities, $\mathbf{a}_1^{'}\boldsymbol{\mu}, \ldots, \mathbf{a}_m^{'}\boldsymbol{\mu}$?

Bonferroni intervals

- For $\mathbf{a}_j^{'}\boldsymbol{\mu}$,

$$\mathbf{a}_j^{'}\bar{\mathbf{x}} \pm t_{n-1}(\alpha_j/2)\sqrt{\mathbf{a}_j^{'}\mathbf{S}\mathbf{a}_j/n}$$

  with $j = 1, \ldots, m$ and to have $1-(\alpha_1 + \ldots + \alpha_m) = 1-\alpha$
- Often to use $\alpha_j = \alpha/m$.
- Make sure overall confidence level $\geq 1-\alpha$.

How to construct $100(1-\alpha)\%$ simultaneous confidence intervals for m quantities, $\mathbf{a}_1^{'}\boldsymbol{\mu}, \ldots, \mathbf{a}_m^{'}\boldsymbol{\mu}$?

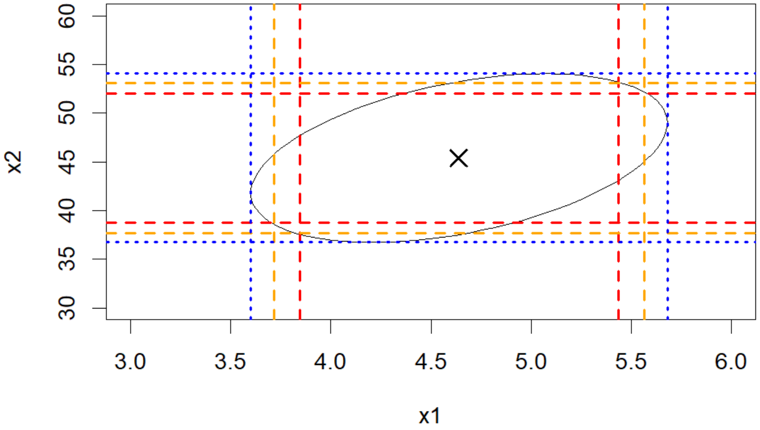
Simultaneous confidence intervals

- $\Longrightarrow$ for all $\mathbf{a}$

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

- Make sure even the worst-case direction has confidence level $\geq 1-\alpha$.

Example. (cont'd)



II.2.4 Large Sample Inference on Population Mean (Chp5.5)

Consider iid p-dim r.v.s. $\mathbf{X}_1, \ldots, \mathbf{X}_n$ with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$. How to make inference about $\boldsymbol{\mu}$?

- $\sqrt{n}\boldsymbol{\Sigma}^{-1/2}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \to MN_p(\mathbf{0}, \mathbf{I})$ by CLT as $(n-p) \to \infty$. Thus $n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \to \chi_p^2$ as $(n-p) \to \infty$.
- $n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \to \chi_p^2$ as $(n-p) \to \infty$.

Hypothesis Testing $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ vs $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$.

- Use the test statistic $T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)$.
- Reject H_0 if $T^2 \geq \chi_p^2(\alpha)$. $\chi_p^2(\alpha)$ is the upper 100α-th percentile of χ_p^2-distn.

Approximate confidence region. $T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})$.

$$R(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \{\boldsymbol{\mu} : t^2 \leq \chi_p^2(\alpha)\}$$

II.2.5 Some Related Topics* (Chp5.6-8)

- *Multivariate Quality Control Charts (Chp 5.6)*
  system is under control or not?
  - extensions of univariate quality control charts
  - Ellipse Format Chart; T^2-Chart; ...
- *Inference about Mean Vectors when Some Observations are Missing (Chp 5.7)*
  - prediction?
  - missing mechanism?
- *Dependence in Multivariate Observations*
  - time series
  - spatial correlated data
  - clustered data

What will we study next?

- *Part I. Introduction and Preparation*
- **Part II. Inference under Multivariate Normal Distribution (Chp 4-7)**
  - *II.1 Multivariate Normal Distribution (Chp 4)*
  - *II.2 Inferences on Mean Vector (Chp 5)*
  - **II.3 Comparisons of Several Mean Vector (Chp 6)**
    - **II.3.1 Introduction (Chp 6.1)**
    - **II.3.2 Comparing Mean Vectors from Two Populations (Chp 6.2-3)**
    - *II.3.3 Comparing Several Mean Vectors and Related (Chp 6.4-6)*
    - *II.3.3 Two-Way Multivariate Analysis of Variance (Chp 6.7)*
  - *II.4 Multivariate Linear Regression (Chp 7)*