# 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.3 Comparisons of Several Mean Vectors (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)*

# II.2 Inferences on Mean Vector (Chp 5) - 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 make inference about $\boldsymbol{\mu}$, such as estimating/testing on hypotheses about $\boldsymbol{\mu}$?

- *Point Estimation on $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$*
  - Method of Moments: The *MME* $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}} = \sum_{i=1}^{n} \mathbf{X}_i \big/ n$, the sample mean, and $\hat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})' = \frac{n-1}{n}\mathbf{S}$.
  - Maximum Likelihood Estimation: The *MLE* $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}$ and $\hat{\boldsymbol{\Sigma}} = \frac{n-1}{n}\mathbf{S}$.
- *Hypothesis Testing on $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$*
  - Wald-test: By $\mathbf{Z} = \sqrt{n}\boldsymbol{\Sigma}^{-1/2}\left(\bar{\mathbf{X}} - \boldsymbol{\mu}_0\right) \sim MN(\mathbf{0}, \mathbf{I})$ under $H_0$, when $\boldsymbol{\Sigma}$ is known, consider the *test statistic*:

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

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

# II.2 Inferences on Mean Vector (Chp 5) - 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 make inference about $\boldsymbol{\mu}$, such as estimating/testing on hypotheses about $\boldsymbol{\mu}$?

- *Point Estimation on $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$*
  - Method of Moments: The *MME* $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}} = \sum_{i=1}^n \mathbf{X}_i \big/ n$, the sample mean, and $\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^{'} = \frac{n-1}{n}\mathbf{S}$.
  - Maximum Likelihood Estimation: The *MLE* $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}$ and $\hat{\boldsymbol{\Sigma}} = \frac{n-1}{n}\mathbf{S}$.
- *Hypothesis Testing on $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$*
  - Hotelling's $T^2$-test: By $\mathbf{T} = \sqrt{n}\mathbf{S}^{-1/2}\left(\bar{\mathbf{X}} - \boldsymbol{\mu}_0\right) \sim MN(\mathbf{0}, \mathbf{I})$ under $H_0$ when $\boldsymbol{\Sigma}$ is unknown, consider the *test statistic*:

    $$W = \mathbf{T}^{'}\mathbf{T} = \left(\bar{\mathbf{X}} - \boldsymbol{\mu}_0\right)^{'}\left(\frac{\mathbf{S}}{n}\right)^{-1}\left(\bar{\mathbf{X}} - \boldsymbol{\mu}_0\right) \sim T^2(p, n-1) \text{ under } H_0$$

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

# II.2 Inferences on Mean Vector (Chp 5) - 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 make inference about $\boldsymbol{\mu}$, such as estimating/testing on hypotheses about $\boldsymbol{\mu}$?

- *Point Estimation on $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$*
  - Method of Moments: The *MME* $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}} = \sum_{i=1}^{n} \mathbf{X}_i / n$, the sample mean, and $\hat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})' = \frac{n-1}{n} \mathbf{S}$.
  - Maximum Likelihood Estimation: The *MLE* $\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}$ and $\hat{\boldsymbol{\Sigma}} = \frac{n-1}{n} \mathbf{S}$.
- *Hypothesis Testing on $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$*
  - Likelihood ratio test:

    Consider the *test statistic* $\Lambda = \frac{L(\boldsymbol{\mu}_0, \hat{\boldsymbol{\Sigma}}_0)}{L(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})} = \left( \frac{|\hat{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}_0|} \right)^{n/2}$ and $\Lambda^{2/n}$ is called *Wilks's lambda*.

    *rejection region.* With pre-determined $\alpha$, $\mathcal{R} = \{\lambda : \lambda < c\}$ with $c$ determined by $P_{H_0}(\Lambda < c) = \alpha$.

    *making decision.* Reject $H_0$ if $\Lambda_{obs} \in \mathcal{R}$; otherwise, accept $H_0$.

# II.2 Inferences on Mean Vector (Chp 5) - 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 make inference about $\boldsymbol{\mu}$, such as estimating/testing on hypotheses about $\boldsymbol{\mu}$?

- *Confidence Regions of $\boldsymbol{\mu}$*

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

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

- *Simultaneous comparison of component means*

  $100(1-\alpha)\%$ 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 $1 - (\alpha_1 + \ldots + \alpha_m) = 1 - \alpha$ such as $\alpha_j = \alpha/m$.

# *II.2 Inferences on Mean Vector (Chp 5) - 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 make inference about $\boldsymbol{\mu}$, such as estimating/testing on hypotheses about $\boldsymbol{\mu}$?

- *Confidence Regions of* $\boldsymbol{\mu}$

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

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

- *Simultaneous comparison of component means*

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

  - Simultaneous confidence intervals:
    for all $\mathbf{a}$

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

# *Example: Bird Data (Tail length $x_1$, Wing length $x_2$)*

Table 5.12 from textbook.

- How to test on $H_0 : \boldsymbol{\mu} = \begin{pmatrix} 190 \\ 280 \end{pmatrix}$ ?

| $x_1$ | $x_2$ | $x_1$ | $x_2$ | $x_1$ | $x_2$ |
|---|---|---|---|---|---|
| 191 | 284 | 186 | 266 | 173 | 271 |
| 197 | 285 | 197 | 285 | 194 | 280 |
| 208 | 288 | 201 | 295 | 198 | 300 |
| 180 | 273 | 190 | 282 | 180 | 272 |
| 180 | 275 | 209 | 305 | 190 | 292 |
| 188 | 280 | 187 | 285 | 191 | 286 |
| 210 | 283 | 207 | 297 | 196 | 285 |
| 196 | 288 | 178 | 268 | 207 | 286 |
| 191 | 271 | 202 | 271 | 209 | 303 |
| 179 | 257 | 205 | 285 | 179 | 261 |
| 208 | 289 | 190 | 280 | 186 | 262 |
| 202 | 285 | 189 | 277 | 174 | 245 |
| 200 | 272 | 211 | 310 | 181 | 250 |
| 192 | 282 | 216 | 305 | 189 | 262 |
| 199 | 280 | 189 | 274 | 188 | 258 |

$$n = 45, \qquad \bar{\mathbf{x}} = \begin{pmatrix} 193.62 \\ 279.78 \end{pmatrix}, \qquad \mathbf{S} = \begin{pmatrix} 120.69 & 122.35 \\ 122.35 & 208.54 \end{pmatrix}, \qquad n = 45, \qquad p = 2.$$

## Example: Bird Data (Tail length $x_1$, Wing length $x_2$)

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \qquad \text{vs} \qquad H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0, \qquad \alpha = 0.05$$

- Compute the observed test statistic (given $\boldsymbol{\mu}_0$).

$$T^2_{\text{obs}} = n(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$$

$$= 45 \begin{pmatrix} 3.6222 \\ -0.2222 \end{pmatrix}' \begin{pmatrix} 120.69 & 122.34 \\ 122.34 & 208.54 \end{pmatrix}^{-1} \begin{pmatrix} 3.6222 \\ -0.2222 \end{pmatrix} = 12.96$$

- Critical value (given $\alpha$). Under normality assumption,

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

find critical value that:

$$c_{0.05} = \frac{2(45-1)}{45-2} F_{2,43}(0.95) = 6.578471.$$

$$\text{where} \quad P_{H_0}\left(T^2 > 6.578471\right) = 0.05.$$

- Is $T^2_{\text{obs}}$ in the rejection region?

## *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 $(n-p) \to \infty$.
  $n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \to \chi^2_p$ as $(n-p) \to \infty$.
- $n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \to \chi^2_p$ 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 $T^2 \geq \chi^2_p(\alpha)$. $\chi^2_p(\alpha)$ is the upper $100\alpha$-th percentile of $\chi^2_p$-distn.

**Approximate** $100(1-\alpha)\%$ **confidence region.**
$T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})$.

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

# II.3 Comparisons of Several Mean Vectors (Chp 6)

Consider $g$ populations: $\mathbf{X}_1, \ldots, \mathbf{X}_g$. Suppose $\mathbf{X}_j \sim MN_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ for $j = 1, \ldots, g$.

- how to compare $\mu_1, \ldots, \mu_g$? (Chp 6.4-6)
- how about to compare $\mu_1, \mu_2$ (i.e. $g = 2$)? (Chp 6.2-3)
- what if the $g$ groups may be looked by two ways: $(l, k)$ for $l = 1, \ldots, a$ and $k = 1, \ldots, b$? (Chp 6.7)

# II.3.2 Comparing Mean Vectors from Two Populations (Chp 6.2-3)

Consider 2 populations: $\mathbf{X}_1, \mathbf{X}_2$. Suppose $\mathbf{X}_j \sim MN_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ for $j = 1, 2$.

**Goal.** to compare $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$

- (i) to estm $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$?
- (ii) to test on $H_0 : \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \boldsymbol{\delta}_0$?

**Data.** $\mathbf{X}_{11}, \ldots, \mathbf{X}_{1n_1}$ are iid observations on $\mathbf{X}_1$, and $\mathbf{X}_{21}, \ldots, \mathbf{X}_{2n_2}$ are iid observations on $\mathbf{X}_2$.

The *key idea* is to use $\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2$.

- $E(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$
- $Var(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) = Var(\bar{\mathbf{X}}_1) + Var(\bar{\mathbf{X}}_2) - 2Cov(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2)$

# II.3.2A Comparing Mean Vectors from Two Populations (Chp 6.2-3)

<u>**Scenario A.**</u> $\mathbf{X}_1 \perp \mathbf{X}_2$ and $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$

- $E(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$
- $Var(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) = Var(\bar{\mathbf{X}}_1) + Var(\bar{\mathbf{X}}_2) - 2Cov(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2) = \mathbf{\Sigma}/n_1 + \mathbf{\Sigma}/n_2$
- Often $\mathbf{S}_{pooled} = \frac{n_1-1}{n_1+n_2-2}\mathbf{S}_1 + \frac{n_2-1}{n_1+n_2-2}\mathbf{S}_2$ is used to estimate $\mathbf{\Sigma}$.
- The $T^2$ statistic follows the Hotelling's $T^2$-distn $\frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_{p,n_1+n_2-p-1}$:

$$T^2 = \left[(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\right]' \left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\mathbf{S}_{pooled}\right]^{-1} \left[(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\right]$$

# II.3.2A Comparing Mean Vectors from Two Populations (Chp 6.2-3)

- Hotelling's $T^2$-test on $H_0 : \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \boldsymbol{\delta}_0$
  *Test statistic* Under $H_0$,

$$\left[(\bar{\mathbf{X}}_1-\bar{\mathbf{X}}_2)-\boldsymbol{\delta}_0\right]' \left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\mathbf{S}_{pooled}\right]^{-1}\left[(\bar{\mathbf{X}}_1-\bar{\mathbf{X}}_2)-\boldsymbol{\delta}_0\right] \sim \frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_{p,n_1+n_2-p-1}$$

  - e.g. $\boldsymbol{\delta}_0 = \mathbf{0}$
- **Example (see R code).** Suppose $p = 2$, where the two components represent some clinical markers (e.g. blood pressure, heart rate). Population 1 is the control group, and population 2 is the treatment group.
  - The mean difference vector $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ describes how the treatment affects these two outcomes jointly.
  - Given the data, compute $\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2$ and $\mathbf{S}_{\text{pooled}}$. Compute the observed test statistic $T^2_{\text{obs}}$ under $H_0$ (i.e. given $\boldsymbol{\delta}_0$), and compare it with the critical value. (R code: `((n1+n2-2)*p)/(n1+n2-p-1) * qf(1-alpha, p, n1+n2-p-1)`)

*What if $n_1 + n_2 - p$ is large?* $T^2 \sim \chi^2_p$ approximately.

# II.3.2B Comparing Mean Vectors from Two Populations (Chp 6.2-3)

**Scenario B.** $\mathbf{X}_1 \perp \mathbf{X}_2$ and $\mathbf{\Sigma}_1 \neq \mathbf{\Sigma}_2$

- Use $\mathbf{S}_1, \mathbf{S}_2$ to estimate $\mathbf{\Sigma}_1, \mathbf{\Sigma}_2$ correspondingly.

$$\left[(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) - \boldsymbol{\delta}_0\right]' \left[\frac{1}{n_1}\mathbf{S}_1 + \frac{1}{n_2}\mathbf{S}_2\right]^{-1} \left[(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) - \boldsymbol{\delta}_0\right] \sim \chi^2_p, \qquad (n_1 - p, n_2 - p \to \infty).$$

- $T^2 \sim \chi^2_p$ approximately if $n_1, n_2$ are large.
- $T^2$'s distribution is complicate if $n_1, n_2$ are not large.*

**Example** (Example 6.5, page 293) Electrical-consumption data. $X_1, X_2$ are two different measurements of electrical usage.

The following summary statistics are given:

$$\bar{\mathbf{x}}_1 = \begin{pmatrix} 204.4 \\ 556.6 \end{pmatrix}, \qquad \mathbf{S}_1 = \begin{pmatrix} 13825.3 & 23823.4 \\ 23823.4 & 73107.4 \end{pmatrix}, \qquad n_1 = 45,$$

$$\bar{\mathbf{x}}_2 = \begin{pmatrix} 130.0 \\ 355.0 \end{pmatrix}, \qquad \mathbf{S}_2 = \begin{pmatrix} 8632.0 & 19616.7 \\ 19616.7 & 55964.5 \end{pmatrix}, \qquad n_2 = 55.$$

To test $H_0 : \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0}$

# II.3.2C Comparing Mean Vectors from Two Populations (Chp 6.2-3)

<u>**Scenario C.**</u> $\mathbf{X}_1 \not\perp \mathbf{X}_2$

- If there is a good estimator for $Var(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) = Var(\bar{\mathbf{X}}_1) + Var(\bar{\mathbf{X}}_2) - 2Cov(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2)$, denoted by $\hat{\mathbf{\Pi}}_n$,

  consider $T^2 = \left[(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) - \delta_0\right]' \left[\hat{\mathbf{\Pi}}_n\right]^{-1} \left[(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) - \delta_0\right]$?

- If observations on the two populations $\mathbf{X}_1$ and $\mathbf{X}_2$ are in pairs: $(\mathbf{X}_{1i}, \mathbf{X}_{2i})$ for $i = 1, \ldots, n$,

  change the two-population problem into a one-population problem: $\mathbf{D} = \mathbf{X}_1 - \mathbf{X}_2$ with iid observations $\mathbf{D}_i = \mathbf{X}_{1i} - \mathbf{X}_{2i}$ for $i = 1, \ldots, n$.

**Example** (Example 6.1, page 276) Measures of biochemical oxygen demand (BOD) and suspended solids (SS) of $n = 11$ sample splits from two labs. Do the two labs' analyses agree? To test $H_0 : \boldsymbol{\mu}_d = \boldsymbol{\mu}_C - \boldsymbol{\mu}_S = \mathbf{0}$.

## II.3.3 Comparing Several Mean Vectors and Related (Chp 6.4, 6.6)

Consider $g$ populations: $\mathbf{X}_1, \ldots, \mathbf{X}_g$. Suppose $\mathbf{X}_j \sim MN_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ for $j = 1, \ldots, g$.

**Goal.** to compare $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_g$

**Data.** $\mathbf{X}_{11}, \ldots, \mathbf{X}_{1n_1}$ are iid observations on $\mathbf{X}_1$; $\ldots$; $\mathbf{X}_{g1}, \ldots, \mathbf{X}_{gn_g}$ are iid observations on $\mathbf{X}_g$.

**Test on** $H_0 : \boldsymbol{\mu}_1 = \ldots = \boldsymbol{\mu}_g = \boldsymbol{\mu}$ with type I error $\alpha$?

- When $g = 2$, by the procedures in Part II.3.2 (Chp6.2-3), compare the two population means in Scenarios A-C.

For $g > 2$, consider pairwise comparisons?

e.g. $g = 3$, consider $H_{01} : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, $H_{02} : \boldsymbol{\mu}_2 = \boldsymbol{\mu}_3$, and $H_{03} : \boldsymbol{\mu}_3 = \boldsymbol{\mu}_1$, each with the adjusted type I error by the Bonferroni correction of $\alpha/3$?

## II.3.3 Comparing Several Mean Vectors and Related (Chp 6.4, 6.6)

Consider $g$ populations: $\mathbf{X}_1, \ldots, \mathbf{X}_g$. Suppose $\mathbf{X}_j \sim MN_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ for $j = 1, \ldots, g$.

**Goal.** to compare $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_g$

**Data.** $\mathbf{X}_{11}, \ldots, \mathbf{X}_{1n_1}$ are iid observations on $\mathbf{X}_1$; $\ldots$; $\mathbf{X}_{g1}, \ldots, \mathbf{X}_{gn_g}$ are iid observations on $\mathbf{X}_g$.

**Test on** $H_0 : \boldsymbol{\mu}_1 = \ldots = \boldsymbol{\mu}_g = \boldsymbol{\mu}$ with type I error $\alpha$?

*Have we studied anything on how to deal with a related problem?*

- When $p = 1$, by the (univariate) ANOVA (analysis of variance), compare the $g$ population means

Consider the ANOVA model: for $l = 1, \ldots, g$,

$$X_{li} = \mu + [\mu_l - \mu] + \epsilon_{li}, \quad \epsilon_{li} \sim N(0, \sigma^2) \;\; iid \;\; i = 1, \ldots, n_l.$$

An analogous decomposition of the observations:

$$\begin{array}{ccccccc} x_{li} & = & \bar{x} & + & [\bar{x}_l - \bar{x}] & + & [x_{li} - \bar{x}_l] \\ \text{(obstn)} & & \begin{pmatrix}\text{overall} \\ \text{sample mean}\end{pmatrix} & & \begin{pmatrix}\text{estm} \\ \text{trt effect}\end{pmatrix} & & \text{(residual)} \end{array}$$

$$\begin{array}{ccccc} \sum_{l=1}^{g}\sum_{i=1}^{n_l}(x_{li} - \bar{x})^2 & = & \sum_{l=1}^{g} n_l(\bar{x}_l - \bar{x})^2 & + & \sum_{l=1}^{g}\sum_{i=1}^{n_l}(x_{li} - \bar{x}_l)^2 \\ (SS_{cor}) & & (SS_{tr}) & & (SS_{res}) \end{array}$$

(Univariate) ANOVA Table ($n_T = \sum_{l=1}^{g} n_l$)

| Source of Variation | df | SS | MSS | F-value |
|---|---|---|---|---|
| treatment | g-1 | $SS_{trt}$ | $\frac{SS_{trt}}{(g-1)}$ | $F = \frac{MSS_{trt}}{MSS_{res}}$ |
| error | $n_T - g$ | $SS_{res}$ | $\frac{SS_{res}}{(n_T-g)}$ | |
| total | $n_T - 1$ | $SS_{cor}$ | $\frac{SS_{cor}}{(n_T-1)}$ | |

To test on $H_0 : \mu_1 = \ldots = \mu_g$ at level $\alpha$ using

$$F = \frac{SS_{trt}/(g-1)}{SS_{res}/(n_T - g)} \sim F(g-1, n_T - g)$$

under $H_0$.

*Can it be extended to multivariate settings?*

## II.3.3 Comparing Several Mean Vectors and Related (Chp 6.4, 6.6)

Consider $g$ populations: $\mathbf{X}_1, \ldots, \mathbf{X}_g$. Suppose $\mathbf{X}_j \sim MN_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ for $j = 1, \ldots, g$.

**Goal.** to compare $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_g$

**Data.** $\mathbf{X}_{11}, \ldots, \mathbf{X}_{1n_1}$ are iid observations on $\mathbf{X}_1$; $\ldots$; $\mathbf{X}_{g1}, \ldots, \mathbf{X}_{gn_g}$ are iid observations on $\mathbf{X}_g$.

**Test on** $H_0 : \boldsymbol{\mu}_1 = \ldots = \boldsymbol{\mu}_g = \boldsymbol{\mu}$ with type I error $\alpha$?

Consider the ANOVA model: for $l = 1, \ldots, g$,

$$\mathbf{X}_{li} = \boldsymbol{\mu} + [\boldsymbol{\mu}_l - \boldsymbol{\mu}] + \boldsymbol{\epsilon}_{li}, \quad \boldsymbol{\epsilon}_{li} \sim MN_p(0, \boldsymbol{\Sigma}) \quad iid \quad i = 1, \ldots, n_l.$$

An analogous decomposition of the observations:

$$\underset{\text{(obstn)}}{\mathbf{x}_{li}} = \underset{\begin{pmatrix}\text{overall} \\ \text{sample mean}\end{pmatrix}}{\bar{\mathbf{x}}} + \underset{\begin{pmatrix}\text{estm} \\ \text{trt effect}\end{pmatrix}}{[\bar{\mathbf{x}}_l - \bar{\mathbf{x}}]} + \underset{\text{(residual)}}{[\mathbf{x}_{li} - \bar{\mathbf{x}}_l]}$$

$$\underset{(\mathbf{SS}_{cor})}{\sum_{l=1}^{g}\sum_{i=1}^{n_l}(\mathbf{x}_{li} - \bar{\mathbf{x}})(\mathbf{x}_{li} - \bar{\mathbf{x}})'} = \underset{(\mathbf{SS}_{tr})}{\sum_{l=1}^{g} n_l(\bar{\mathbf{x}}_l - \bar{\mathbf{x}})(\bar{\mathbf{x}}_l - \bar{\mathbf{x}})'} + \underset{(\mathbf{SS}_{res})}{\sum_{l=1}^{g}\sum_{i=1}^{n_l}(\mathbf{x}_{li} - \bar{\mathbf{x}}_l)(\mathbf{x}_{li} - \bar{\mathbf{x}}_l)'}$$

## Multivariate ANOVA Table ($n_T = \sum_{l=1}^{g} n_l$)

| Source of Variation | df | SS |
|---|---|---|
| treatment | g-1 | $\mathbf{SS}_{trt} = \sum_{l=1}^{g} n_l(\bar{\mathbf{x}}_l - \bar{\mathbf{x}})(\bar{\mathbf{x}}_l - \bar{\mathbf{x}})'$ |
| error | $n_T - g$ | $\mathbf{SS}_{res} = \sum_{l=1}^{g} \sum_{i=1}^{n_l} (\mathbf{x}_{li} - \bar{\mathbf{x}}_l)(\mathbf{x}_{li} - \bar{\mathbf{x}}_l)'$ |
| total | $n_T - 1$ | $\mathbf{SS}_{cor} = \sum_{l=1}^{g} \sum_{i=1}^{n_l} (\mathbf{x}_{li} - \bar{\mathbf{x}})(\mathbf{x}_{li} - \bar{\mathbf{x}})'$ |

To test on $H_0 : \boldsymbol{\mu}_1 = \ldots = \boldsymbol{\mu}_g$ using the Wilks' lambda statistic: $\Lambda^* = \frac{|\mathbf{SS}_{res}|}{|\mathbf{SS}_{cor}|}$.

- Reject $H_0$ if $\Lambda^*_{obs}$ is small.
- Textbook Table 6.3 presents the distn of $\Lambda^*$.
- We use software to implement the test (e.g. *manova()* function in R).

## 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 Vectors (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.3.4 Two-Way Multivariate Analysis of Variance (Chp 6.7)**
  - *II.4 Multivariate Linear Regression (Chp 7)*