STAT 3690 Lecture Note

Week Six (Feb 13, 15, & 17, 2023)

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca, zhiyanggeezhou.github.io)

2023/Feb/16 23:46:35

Inference on $\boldsymbol{\mu}$ under the normality assumption (con'd)

Testing $\boldsymbol{\mu}$ (J&W Sec. 5.2 & 5.3, con'd)

- Sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $n > p$
  - $\mathcal{X} = \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n\}$
- $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ v.s. $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$
- Multivariate case (with unknown $\boldsymbol{\Sigma}$)
  - Name of approach: LRT
  - Test statistic: $T(\mathcal{X}) = n(\bar{\boldsymbol{X}} - \boldsymbol{\mu}_0)^\top \mathbf{S}^{-1}(\bar{\boldsymbol{X}} - \boldsymbol{\mu}_0)$ ($\sim T^2(p, n-1) = \frac{(n-1)p}{n-p} F(p, n-p)$ under H_0)
  - Level α rejection region (with respect to $T(\mathcal{X})$): $R_\alpha = \{T(\mathcal{X}) : \frac{n-p}{p(n-1)} T(\mathcal{X}) \geq F_{1-\alpha,p,n-p}\}$, i.e., reject H_0 if $T(\mathcal{X}) \geq \frac{p(n-1)}{n-p} F_{1-\alpha,p,n-p}$
    - $F_{1-\alpha,p,n-p}$: the $(1-\alpha)$-quantile of $F(p, n-p)$
  - p-value: $p(\mathcal{X}) = 1 - F_{F(p,n-p)}\{\frac{n-p}{p(n-1)} T(\mathcal{X})\}$
    - $F_{F(p,n-p)}$: the cdf of $F(p, n-p)$

```
options(digits = 4)
install.packages(c("dslabs"))
library(dslabs)
data("gapminder")
dataset = as.matrix(gapminder[
  !is.na(gapminder$infant_mortality),
  c("infant_mortality", "life_expectancy", "fertility")])

(mu_hat <- colMeans(dataset))

# Test mu = mu_0
mu_0 <- c(25, 50, 3)
n = nrow(dataset)
p = ncol(dataset)
(test.stat <- drop(
  n * t(mu_hat - mu_0) %*% solve(cov(dataset)) %*% (mu_hat - mu_0)
))
(cri.point = (n-1)*p/(n-p)*qf(.95, p, n-p))
```

```r
test.stat >= cri.point
(p.val = 1-pf((n-p)/(n-1)/p*test.stat, p, n-p))
```

- Report: Testing hypotheses $H_0: \boldsymbol{\mu} = [25, 50, 3]^\top$ v.s. $H_1: \boldsymbol{\mu} \neq [25, 50, 3]^\top$, we carried on the LRT and obtained 249718 as the value of test statistic with $[7.819, \infty)$ as the corresponding level .05 rejection region. In addition, the p-value was almost 0. So, at the .05 level, there was a strong statistical evidence implying the rejection of H_0, i.e., we believed that the population mean vector was not $[25, 50, 3]^\top$.

$(1-\alpha) \times 100\%$ confidence region (CR) for $\boldsymbol{\mu}$ (J&W Sec. 5.4)

- $\Pr\{(1-\alpha) \times 100\%$ CR covers $\boldsymbol{\mu}\} \geq 1-\alpha$
  - CR is a set made of observations and is hence random
  - $\boldsymbol{\mu}$ is fixed
  - $(1-\alpha) \times 100\%$ CR covers $\boldsymbol{\mu}$ with probability at least $(1-\alpha) \times 100\%$
- Inverted from the level α rejection region for $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$ v.s. $H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$. Specifically,
  1. Take the rejection region as a function of $\boldsymbol{\mu}_0$;
  2. Replace $\boldsymbol{\mu}_0$ with $\boldsymbol{\mu}$;
  3. Take the complement.
- Eventually, $(1-\alpha) \times 100\%$ CR
  - $= \{\boldsymbol{\mu} : n(\bar{\boldsymbol{x}} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\bar{\boldsymbol{x}} - \boldsymbol{\mu}) < \chi^2_{1-\alpha,p}\}$ if $\boldsymbol{\Sigma}$ is known
  - $= \{\boldsymbol{\mu} : \frac{n(n-p)}{p(n-1)} (\bar{\boldsymbol{x}} - \boldsymbol{\mu})^\top \mathbf{S}^{-1} (\bar{\boldsymbol{x}} - \boldsymbol{\mu}) < F_{1-\alpha,p,n-p}\}$ if $\boldsymbol{\Sigma}$ is not known

Testing $\mathbf{A}\boldsymbol{\mu}$ (J&W pp. 279)

- $\mathbf{A}$ is of $q \times p$ and $\mathrm{rk}(\mathbf{A}) = q$, i.e., $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top > 0$
- Known: iid $\mathbf{A}\boldsymbol{X}_i \sim \mathrm{MVN}_q(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\Sigma\mathbf{A}^\top)$.
- LRT for $H_0: \mathbf{A}\boldsymbol{\mu} = \boldsymbol{\nu}_0$ v.s. $H_1: \mathbf{A}\boldsymbol{\mu} \neq \boldsymbol{\nu}_0$
  - Test statistic: $T(\mathcal{X}) = n(\mathbf{A}\bar{\boldsymbol{X}} - \boldsymbol{\nu}_0)^\top (\mathbf{A}\mathbf{S}\mathbf{A}^\top)^{-1} (\mathbf{A}\bar{\boldsymbol{X}} - \boldsymbol{\nu}_0)$ $(\sim T^2(q, n-1) = \frac{(n-1)q}{n-q} F(q, n-q)$ under $H_0)$
  - Level α rejection region (with respect to $T(\mathcal{X})$): $R_\alpha = \{T(\mathcal{X}) : \frac{n-q}{q(n-1)} T(\mathcal{X}) \geq F_{1-\alpha,q,n-q}\}$
  - p-value: $p(\mathcal{X}) = 1 - F_{F(q,n-q)}\{\frac{n-q}{q(n-1)} T(\mathcal{X})\}$
- Multiple comparison
  - Interested in $H_0: \mu_1 = \cdots = \mu_p$ v.s. H_1 : Not all entries of $\boldsymbol{\mu}$ are equal.
    - μ_k: the kth entry of $\boldsymbol{\mu}$
  - Take
$$\boldsymbol{\nu}_0 = \mathbf{0}_{(p-1)\times 1}, \quad \mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix}_{(p-1)\times p}.$$
  - $p = 2$ (i.e., $\mathbf{A} = [1, -1]$): the case of A/B testing

---

```r
options(digits = 4)
install.packages(c("dslabs",'tidyverse'))
library(dslabs)
library(tidyverse)
data("gapminder")
dataset = gapminder[
  !is.na(gapminder$infant_mortality) &
    gapminder$region == 'South America' &
```

```
    gapminder$year %in% 2000:2008,
  c('country', 'year', "life_expectancy")] %>%
  spread(year, life_expectancy)
(dataset = as.matrix(dataset[, -1]))
n = nrow(dataset); p = ncol(dataset)
(mu_hat <- colMeans(dataset))

# Test H0:A %*% mu = nu_0
(nu_0 <- as.matrix(rep(0, p-1)))
(A = cbind(rep(1, p-1), -diag(p-1)))

(test.stat <- drop(
  n * t(A %*% mu_hat - nu_0) %*%
    solve(A %*% cov(dataset) %*% t(A)) %*%
    (A %*% mu_hat - nu_0)
))
(cri.point = (n-1)*(p-1)/(n-p+1)*qf(.95, p-1, n-p+1))
test.stat >= cri.point
(p.val = 1-pf((n-p+1)/(n-1)/(p-1)*test.stat, p-1, n-p+1))
```

- Report: Testing hypotheses H_0: the average life expectancy over south american countries doesn't vary with time v.s. H_1: otherwise, we carried on the LRT and obtained 628.5 as the value of test statistic and $[132.9, \infty)$ as the corresponding level .05 rejection region. In addition, the p-value was .002858. So, at the .05 level, there was a strong statistical evidence against H_0, i.e., we believed that the average life expectancy over south american countries does vary with time.

$(1-\alpha)\times 100\%$ CR for $\boldsymbol{\nu} = \mathbf{A}\boldsymbol{\mu}$

- $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with unknown $\boldsymbol{\Sigma}$ and $n > p$
- $\mathbf{A}$ is of $q \times p$ and $\text{rk}(\mathbf{A}) = q$, i.e., $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top > 0$
- Then iid $\mathbf{A}\boldsymbol{X}_i \sim \text{MVN}_q(\boldsymbol{\nu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$
- $(1-\alpha)\times 100\%$ CR for $\boldsymbol{\nu}$ is $\{\boldsymbol{\nu} : \frac{n(n-q)}{q(n-1)}(\mathbf{A}\bar{\boldsymbol{x}} - \boldsymbol{\nu})^\top(\mathbf{A}\mathbf{S}\mathbf{A}^\top)^{-1}(\mathbf{A}\bar{\boldsymbol{x}} - \boldsymbol{\nu}) < F_{1-\alpha,q,n-q}\}$
- Special case: $\mathbf{A} = \boldsymbol{a}^\top \in \mathbb{R}^{1\times p}$, i.e., $\mathbf{A}$ is a row vector. Then
  - $(1-\alpha)\times 100\%$ confidence interval (CI) for scalar $\nu = \boldsymbol{a}^\top\boldsymbol{\mu}$ is $\{\nu : n(\boldsymbol{a}^\top\bar{\boldsymbol{x}} - \nu)^2/(\boldsymbol{a}^\top\mathbf{S}\boldsymbol{a}) < F_{1-\alpha,1,n-1}\}$, i.e.,

$$\left(\boldsymbol{a}^\top\bar{\boldsymbol{x}} - t_{1-\alpha/2,n-1}\sqrt{\boldsymbol{a}^\top\mathbf{S}\boldsymbol{a}/n},\quad \boldsymbol{a}^\top\bar{\boldsymbol{x}} + t_{1-\alpha/2,n-1}\sqrt{\boldsymbol{a}^\top\mathbf{S}\boldsymbol{a}/n}\right)$$

    * E.g., when $\mathbf{A} = [1, 0, \ldots, 0]$, it is the CI for the first entry of $\boldsymbol{\mu}$, say μ_1
- Checking the coverage probability of the previous CI for each μ_k

```
options(digits = 4)
install.packages(c("MASS"))
set.seed(1)
B = 5e3L
n = 5e2L
Mu = (1:10)^2; (p = length(Mu))
(Sigma = diag(p)+.5)
alpha <- .05
(A = diag(p))
```

```
cover = matrix(0, ncol = p, nrow = B)
for (b in 1:B){
  sample = MASS::mvrnorm(n, Mu, Sigma)
  mu_hat = colMeans(sample)
  sample_cov = cov(sample)
  LB = A %*% mu_hat - qt(1-alpha/2, n-1)* sqrt(diag(A %*% sample_cov %*% t(A))/n)
  RB = A %*% mu_hat + qt(1-alpha/2, n-1)* sqrt(diag(A %*% sample_cov %*% t(A))/n)
  cover[b,] = (LB < Mu) * (Mu < RB)
}
(cover_prob_indiv = colMeans(cover))
(cover_prob_simul = mean(apply(cover, 1, prod)))
```

Simultaneous confidence intervals

- Interested in $(1-\alpha_k)\times 100\%$ CIs for scalars $\boldsymbol{a}_k^\top\boldsymbol{\mu}$, say CR_k, $k=1,\ldots,m$, simultaneously
- Make sure $\Pr(\bigcap_k\{\boldsymbol{a}_k^\top\boldsymbol{\mu}\in\mathrm{CR}_k\})\geq 1-\alpha$
- Bonferroni correction
  - Bonferroni inequality (optional) :

$$\Pr(\bigcap_{k=1}^m\{\boldsymbol{a}_k^\top\boldsymbol{\mu}\in\mathrm{CR}_k\}) = 1-\Pr(\bigcup_{k=1}^m\{\boldsymbol{a}_k^\top\boldsymbol{\mu}\notin\mathrm{CR}_k\})\geq 1-\sum_{k=1}^m\Pr(\boldsymbol{a}_k^\top\boldsymbol{\mu}\notin\mathrm{CR}_k) = 1-\sum_{k=1}^m\alpha_k$$

  - Taking α_k such that $\alpha=\sum_{k=1}^m\alpha_k$, e.g., $\alpha_k=\alpha/m$, i.e.,

$$(\boldsymbol{a}_k^\top\bar{\boldsymbol{x}}-t_{1-\alpha/(2m),n-1}\sqrt{\boldsymbol{a}_k^\top\mathbf{S}\boldsymbol{a}_k/n},\quad \boldsymbol{a}_k^\top\bar{\boldsymbol{x}}+t_{1-\alpha/(2m),n-1}\sqrt{\boldsymbol{a}_k^\top\mathbf{S}\boldsymbol{a}_k/n})$$

  - Appropriate for small m
- Scheffé's method
  - Let $\mathrm{CI}_{\boldsymbol{a}}=(\boldsymbol{a}^\top\bar{\boldsymbol{x}}-c\sqrt{\boldsymbol{a}^\top\mathbf{S}\boldsymbol{a}/n},\boldsymbol{a}^\top\bar{\boldsymbol{x}}+c\sqrt{\boldsymbol{a}^\top\mathbf{S}\boldsymbol{a}/n})$ for all $\boldsymbol{a}\in\mathbb{R}^p$. Then we may find that $c=\sqrt{p(n-1)(n-p)^{-1}F_{1-\alpha,p,n-p}}$.
  - Derivation by Cauchy-Schwarz inequality (optional): $\{\boldsymbol{a}^\top(\bar{\boldsymbol{x}}-\boldsymbol{\mu})\}^2=[(\mathbf{S}^{1/2}\boldsymbol{a})^\top\{\mathbf{S}^{-1/2}(\bar{\boldsymbol{x}}-\boldsymbol{\mu})\}]^2\leq\{(\boldsymbol{a}^\top\mathbf{S}\boldsymbol{a})^\top/n\}\{n(\bar{\boldsymbol{x}}-\boldsymbol{\mu})^\top\mathbf{S}^{-1}(\bar{\boldsymbol{x}}-\boldsymbol{\mu})\}\Rightarrow$

$$\begin{aligned}\Pr(\bigcap_{k=1}^m\{\boldsymbol{a}_k^\top\boldsymbol{\mu}\in\mathrm{CI}_k\}) &\geq \Pr(\bigcap_{\boldsymbol{a}\in\mathbb{R}^p}\{\boldsymbol{a}^\top\boldsymbol{\mu}\in\mathrm{CI}_{\boldsymbol{a}}\}) = 1-\Pr(\bigcup_{\boldsymbol{a}\in\mathbb{R}^p}\{\boldsymbol{a}^\top\boldsymbol{\mu}\notin\mathrm{CI}_{\boldsymbol{a}}\})\\ &= 1-\Pr(\bigcup_{\boldsymbol{a}\in\mathbb{R}^p}[\{\boldsymbol{a}^\top(\bar{\boldsymbol{X}}-\boldsymbol{\mu})\}^2/\{(\boldsymbol{a}^\top\mathbf{S}\boldsymbol{a})^\top/n\}>c^2])\\ &\geq 1-\Pr(\{n(\bar{\boldsymbol{X}}-\boldsymbol{\mu})^\top\mathbf{S}^{-1}(\bar{\boldsymbol{X}}-\boldsymbol{\mu})>c^2\})\end{aligned}$$

    Assume $\Pr(\{n(\bar{\boldsymbol{X}}-\boldsymbol{\mu})^\top\mathbf{S}^{-1}(\bar{\boldsymbol{X}}-\boldsymbol{\mu})>c^2\})=\alpha$ and obtain $c=\sqrt{p(n-1)(n-p)^{-1}F_{1-\alpha,p,n-p}}$.
  - Appropriate for large even infinite m

```
options(digits = 4)
install.packages(c("dslabs"))
library(dslabs)
data("gapminder")
dataset = gapminder[
  !is.na(gapminder$infant_mortality) &
    gapminder$year == 2012,
```

```
  c('infant_mortality', "life_expectancy")]
dataset = as.matrix(dataset)

n = nrow(dataset); p = ncol(dataset)

alpha <- .05
a1 = c(1,0); a2 = c(0,1)
A = rbind(a1, a2)
(mu_hat <- colMeans(dataset))
(sample_cov <- cov(dataset))

# Simultaneous CIs without correction
c = qt(1-alpha/2, n-1)
(NOcorrection <- cbind(
    A %*% mu_hat - c * sqrt(diag(A %*% sample_cov %*% t(A))/n),
    A %*% mu_hat + c * sqrt(diag(A %*% sample_cov %*% t(A))/n)
))

# Simultaneous CIs with Bonferroni correction
m = nrow(A)
c = qt(1-alpha/2/m, n-1)
(Bonferroni <- cbind(
    A %*% mu_hat - c * sqrt(diag(A %*% sample_cov %*% t(A))/n),
    A %*% mu_hat + c * sqrt(diag(A %*% sample_cov %*% t(A))/n)
))

# Simultaneous CIs with Scheffe correction
c = sqrt(p*(n-1)/(n-p) * qf(1-alpha, p, n-p))
(Scheffe <- cbind(
    A %*% mu_hat - c * sqrt(diag(A %*% sample_cov %*% t(A))/n),
    A %*% mu_hat + c * sqrt(diag(A %*% sample_cov %*% t(A))/n)
))
```

- Report: After the Bonferroni correction, the resulting CIs (21.82, 29.82) and (69.92, 72.70) cover the mean infant mortality and mean life expectancy, simultaneously, with probability at least 95%.

The confidence region for $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_p]^\top$ vs. simultaneous confidence intervals for $\mu_1, \ldots, \mu_p$

- $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with known $\boldsymbol{\Sigma}$ and $n > p$
- $(1-\alpha) \times 100\%$ CR for $\boldsymbol{\mu}$: $\{\boldsymbol{\mu} : n(\bar{\boldsymbol{x}} - \boldsymbol{\mu})^\top \mathbf{S}^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}) < \frac{p(n-1)}{n-p} F_{1-\alpha,p,n-p}\}$
  - CR covering $\boldsymbol{\mu}$ with a probability at least $1-\alpha$
  - With a coverage probability closer to $(1-\alpha) \times 100\%$
- $(1-\alpha) \times 100\%$ simultaneous CI_k for μ_k: $\left(\bar{x}_k - c\sqrt{S_{kk}/n}, \quad \bar{x}_k + c\sqrt{S_{kk}/n}\right)$ with $\bar{x}_k$ the kth entry of $\bar{\boldsymbol{x}}$ and S_{kk} the (k,k)-th entry of $\mathbf{S}$
  - $c = \sqrt{\frac{p(n-1)}{n-p} F_{1-\alpha,p,n-p}}$ (Scheffé) and $t_{1-\alpha/(2p),n-1}$ (Bonferroni)
  - $\text{CI}_1 \times \cdots \times \text{CI}_p$ covering $\boldsymbol{\mu}$ with a probability at least $1-\alpha$
  - Clearly indicating the range for each μ_k

Comparing two population mean vectors (J&W Sec. 6.3)

- Two independent samples following two distributions with equal covariance
  - $\boldsymbol{X}_{11}, \ldots, \boldsymbol{X}_{1n_1} \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$
  - $\boldsymbol{X}_{21}, \ldots, \boldsymbol{X}_{2n_2} \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$
- Let $\bar{\boldsymbol{X}}_i$ and $\mathbf{S}_i$ be the sample mean and sample covariance for the ith sample, $i = 1, 2$.
- Hypotheses $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ v.s. $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$
- Test statistic following LRT

$$T(\mathcal{X}) = (\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2)^\top \{(n_1^{-1} + n_2^{-1})\mathbf{S}_{\text{pool}}\}^{-1}(\bar{\boldsymbol{X}}_1 - \bar{\boldsymbol{X}}_2) \sim \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} F(p, n_1 + n_2 - p - 1)$$

  - $\mathbf{S}_{\text{pool}} = \frac{(n_1-1)\mathbf{S}_1 + (n_2-1)\mathbf{S}_2}{n_1+n_2-2}$
- Level α rejection region

$$\left\{ T(\mathcal{X}) : T(\mathcal{X}) \geq \frac{p(n_1 + n_2 - 2)}{n_1 + n_2 - p - 1} F_{1-\alpha, p, n_1+n_2-p-1} \right\}$$

- p-value

$$1 - F_{F_{1-\alpha,p,n_1+n_2-p-1}} \left[\frac{n_1 + n_2 - p - 1}{p(n_1 + n_2 - 2)} T(\mathcal{X}) \right]$$

- Report: Testing hypotheses H_0: in 2012 Asia and Africa shared the identical mean value in both infant mortality and life expectancy v.s. H_1: otherwise, we carried on the LRT and obtained 87.65 as the value of test statistic and $[6.255, \infty)$ as the corresponding rejection region. In addition, the p-value was 4.952e-14. So, at the .05 level, there was a strong statistical evidence against H_0, i.e., we rejected H_0 and believed that in 2012 Asia and Africa didn't share the identical mean value in either infant mortality or life expectancy.

Testing for equality of population means (one-way multivariate analysis of variance (1-way MANOVA), J&W Sec. 6.4)

- Generalization of two-sample problem
  - Model: m independent samples, where
    - $\boldsymbol{X}_{11}, \ldots, \boldsymbol{X}_{1n_1} \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$
    - $\vdots$
    - $\boldsymbol{X}_{m1}, \ldots, \boldsymbol{X}_{mn_m} \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu}_m, \boldsymbol{\Sigma})$
  - Hypotheses $H_0 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_m$ v.s. H_1 : otherwise
- Alternatively
  - Model: m independent samples, where
    - $\boldsymbol{X}_{11}, \ldots, \boldsymbol{X}_{1n_1} \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu} + \boldsymbol{\tau}_1, \boldsymbol{\Sigma})$
    - $\vdots$
    - $\boldsymbol{X}_{m1}, \ldots, \boldsymbol{X}_{mn_m} \overset{\text{iid}}{\sim} \text{MVN}_p(\boldsymbol{\mu} + \boldsymbol{\tau}_m, \boldsymbol{\Sigma})$
      - Identifiability: $\sum_i \boldsymbol{\tau}_i = 0$ otherwise there are infinitely many models that lead to the same data-generating mechanism.
  - Hypotheses $H_0 : \boldsymbol{\tau}_1 = \cdots = \boldsymbol{\tau}_m = 0$ v.s. H_1 : otherwise
- Alternatively
  - Model: $\boldsymbol{X}_{ij} = \boldsymbol{\mu} + \boldsymbol{\tau}_i + \mathbf{E}_{ij}$ with $\mathbf{E}_{ij} \overset{\text{iid}}{\sim} \text{MVN}_p(\mathbf{0}, \boldsymbol{\Sigma})$
    - Identifiability: $\sum_i \boldsymbol{\tau}_i = 0$

- Hypotheses $H_0 : \boldsymbol{\tau}_1 = \cdots = \boldsymbol{\tau}_m = 0$ v.s. H_1 : otherwise

---

- Sample means and sample covariances
  - Sample mean for the ith sample $\bar{\boldsymbol{X}}_i = n_i^{-1} \sum_j \boldsymbol{X}_{ij}$
  - Sample covariance for the ith sample $\mathbf{S}_i = (n_i - 1)^{-1} \sum_j (\boldsymbol{X}_{ij} - \bar{\boldsymbol{X}}_i)(\boldsymbol{X}_{ij} - \bar{\boldsymbol{X}}_i)^\top$
  - Grand mean $\bar{\boldsymbol{X}} = \sum_i n_i \bar{\boldsymbol{X}}_i / \sum_i n_i = \sum_{ij} \boldsymbol{X}_{ij} / \sum_i n_i$
  - Sum of squares and cross products matrix (SSP)
    * Within-group SSP

$$\mathbf{SSP}_{\mathrm{w}} = \sum_i (n_i - 1)\mathbf{S}_i = \sum_{ij} (\boldsymbol{X}_{ij} - \bar{\boldsymbol{X}}_i)(\boldsymbol{X}_{ij} - \bar{\boldsymbol{X}}_i)^\top$$

    * Between-group SSP

$$\mathbf{SSP}_{\mathrm{b}} = \sum_i n_i (\bar{\boldsymbol{X}}_i - \bar{\boldsymbol{X}})(\bar{\boldsymbol{X}}_i - \bar{\boldsymbol{X}})^\top$$

    * Total (corrected) SSP

$$\mathbf{SSP}_{\mathrm{cor}} = \sum_{ij} (\boldsymbol{X}_{ij} - \bar{\boldsymbol{X}})(\boldsymbol{X}_{ij} - \bar{\boldsymbol{X}})^\top = \mathbf{SSP}_{\mathrm{w}} + \mathbf{SSP}_{\mathrm{b}}$$

- Exercise: verify the decomposition $\mathbf{SSP}_{\mathrm{cor}} = \mathbf{SSP}_{\mathrm{w}} + \mathbf{SSP}_{\mathrm{b}}$.

- ML estimator of $(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_m, \boldsymbol{\Sigma})$
  - Unconstrained
    * $\hat{\boldsymbol{\mu}}_i = \bar{\boldsymbol{X}}_i = n_i^{-1} \sum_j \boldsymbol{X}_{ij}$
    * $\widehat{\boldsymbol{\Sigma}} = (\sum_i n_i)^{-1} \mathbf{SSP}_{\mathrm{w}}$
  - Under H_0
    * $\hat{\boldsymbol{\mu}}_i = \bar{\boldsymbol{X}}$ for each i
    * $\widehat{\boldsymbol{\Sigma}} = (\sum_i n_i)^{-1} \mathbf{SSP}_{\mathrm{cor}}$
- Likelihood ratio

$$\lambda = \left\{ \frac{\det(\mathbf{SSP}_{\mathrm{w}})}{\det(\mathbf{SSP}_{\mathrm{cor}})} \right\}^{\sum_i n_i / 2}$$

  monotonic with respect to the Wilk's lambda test statistic

$$\Lambda = \frac{\det(\mathbf{SSP}_{\mathrm{w}})}{\det(\mathbf{SSP}_{\mathrm{cor}})}$$

  - Under H_0, $\Lambda \sim$ Wilk's lambda distribution $\Lambda(\boldsymbol{\Sigma}, \sum_i n_i - m, m - 1)$
    * Since $\mathbf{SSP}_{\mathrm{w}} \sim W_p(\boldsymbol{\Sigma}, \sum_i n_i - m)$ and $\mathbf{SSP}_{\mathrm{b}} \sim W_p(\boldsymbol{\Sigma}, m - 1)$
    * When $\sum_i n_i - m$ is large (i.e., $(p + m)/2 - \sum_i n_i + 1 \ll 0$), Bartlett's approximation

$$\{(p + m)/2 - \sum_i n_i + 1\} \ln \Lambda \approx \chi^2(p(m - 1))$$

- Level α rejection region (with respect to Λ)

$$\begin{aligned} &\left\{ \Lambda : \{(p + m)/2 - \sum_i n_i + 1\} \ln \Lambda \geq \chi^2_{1-\alpha, p(m-1)} \right\} \\ &= \left\{ \Lambda : \Lambda \leq \exp\left\{ \frac{\chi^2_{1-\alpha, p(m-1)}}{(p + m)/2 - \sum_i n_i + 1} \right\} \right\} \end{aligned}$$

- p-value

$$1 - F_{\chi^2(p(m-1))}\left[\{(p+m)/2 - \sum_i n_i + 1\} \ln \Lambda\right]$$

- Exercise: factors in producing plastic film (see W. J. Krzanowski (1988) *Principles of Multivariate Analysis*. A User's Perspective. Oxford UP, pp. 381.)
  - Three response variables (tear, gloss and opacity) describing measured characteristics of the resultant film
  - A total of 20 runs
  - One factor `RATE` (rate of extrusion, 2-level, low or high) in the production test
- Report: Testing hypotheses H_0: no `RATE` effect on film characteristics v.s. H_1: otherwise, we carried on the Wilk's lambda test and obtained 0.4136 as the value of test statistic and $(-\infty, 0.6227]$ as the level .05 rejection region. In addition, the p-value was 0.002227. So, at the .05 level, there was statistical evidence against H_0, i.e., we rejected H_0 and believed that there was an effect from `RATE` on film characteristics.