STAT 3690 Lecture 14

zhiyanggeezhou.github.io

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca)

Mar 02, 2022

2-way MANOVA (J&W Sec. 6.7)

- Model: $\mathbf{X}_{ijk} = \boldsymbol{\mu} + \boldsymbol{\tau}_i + \boldsymbol{\beta}_j + \boldsymbol{\gamma}_{ij} + \mathbf{E}_{ijk}$ with $\mathbf{E}_{ijk} \overset{\text{iid}}{\sim} MVN_p(\mathbf{0}, \boldsymbol{\Sigma})$, $i = 1, \ldots, m$, $j = 1, \ldots, b$, $k = 1, \ldots, n$
  - $\boldsymbol{\tau}_i$: the main effect of factor 1 at level i
  - $\boldsymbol{\beta}_j$: the main effect of factor 2 at level j
  - $\boldsymbol{\gamma}_{ij}$: the interaction of factors 1 and 2 whose levels are i and j, respectively
  - Identifiability: $\sum_i \boldsymbol{\tau}_i = \sum_j \boldsymbol{\beta}_j = \sum_i \boldsymbol{\gamma}_{ij} = \sum_j \boldsymbol{\gamma}_{ij} = \mathbf{0}$

- Sum of squares and cross products matrix (SSP)
  - Total corrected SSP
  $$\mathbf{SSP}_{\text{cor}} = \sum_{i=1}^{m}\sum_{j=1}^{b}\sum_{k=1}^{n}(\mathbf{X}_{ijk} - \bar{\mathbf{X}})(\mathbf{X}_{ijk} - \bar{\mathbf{X}})^\top$$
    - $\bar{\mathbf{X}} = (mbn)^{-1}\sum_{i,j,k}\mathbf{X}_{ijk}$
  - SSP for factor 1
  $$\mathbf{SSP}_{\text{f1}} = \sum_{i=1}^{m} bn(\bar{\mathbf{X}}_{i\cdot} - \bar{\mathbf{X}})(\bar{\mathbf{X}}_{i\cdot} - \bar{\mathbf{X}})^\top$$
    - $\bar{\mathbf{X}}_{i\cdot} = (bn)^{-1}\sum_{j,k}\mathbf{X}_{ijk}$
  - SSP for factor 2
  $$\mathbf{SSP}_{\text{f2}} = \sum_{j=1}^{b} mn(\bar{\mathbf{X}}_{\cdot j} - \bar{\mathbf{X}})(\bar{\mathbf{X}}_{\cdot j} - \bar{\mathbf{X}})^\top$$
    - $\bar{\mathbf{X}}_{\cdot j} = (mn)^{-1}\sum_{i,k}\mathbf{X}_{ijk}$
  - SSP for interaction
  $$\mathbf{SSP}_{\text{int}} = \sum_{i=1}^{m}\sum_{j=1}^{b} n(\bar{\mathbf{X}}_{ij} - \bar{\mathbf{X}}_{i\cdot} - \bar{\mathbf{X}}_{\cdot j} + \bar{\mathbf{X}})(\bar{\mathbf{X}}_{ij} - \bar{\mathbf{X}}_{i\cdot} - \bar{\mathbf{X}}_{\cdot j} + \bar{\mathbf{X}})^\top$$
    - $\bar{\mathbf{X}}_{ij} = n^{-1}\sum_{k}\mathbf{X}_{ijk}$
  - SSP for residual
  $$\mathbf{SSP}_{\text{res}} = \sum_{i=1}^{m}\sum_{j=1}^{b}\sum_{k=1}^{n}(\mathbf{X}_{ijk} - \bar{\mathbf{X}}_{ij})(\mathbf{X}_{ijk} - \bar{\mathbf{X}}_{ij})^\top$$
  - $\mathbf{SSP}_{\text{cor}} = \mathbf{SSP}_{\text{f1}} + \mathbf{SSP}_{\text{f2}} + \mathbf{SSP}_{\text{int}} + \mathbf{SSP}_{\text{res}}$

$$x_{ijk} - \bar{x} = \underbrace{(x_{ijk} - \bar{x}_{ij})}_{a_{ijk}} + \underbrace{(\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x})}_{b_{ij}} + \underbrace{(\bar{x}_{\cdot j} - \bar{x})}_{c_j} + \underbrace{(\bar{x}_{i\cdot} - \bar{x})}_{d_i}$$

$$\begin{aligned}
SSP_{cor} &= \sum_{i,j,k} (x_{ijk} - \bar{x})(x_{ijk} - \bar{x})^T \\
&= \sum_{i,j,k} (a_{ijk} + b_{ij} + c_j + d_i)(a_{ijk} + b_{ij} + c_j + d_i)^T \\
&= \sum_{i,j,k} \left(a_{ijk} a_{ijk}^T + b_{ij} b_{ij}^T + c_j c_j^T + d_i d_i^T \right) \\
&= \underbrace{\sum_{ijk} a_{ijk} a_{ijk}^T}_{SSP_{res}} + \underbrace{\sum_{ijk} b_{ij} b_{ij}^T}_{SSP_{int}} + \underbrace{\sum_{ijk} c_j c_j^T}_{SSP_{f2}} + \underbrace{\sum_{ijk} d_i d_i^T}_{SSP_{f1}}
\end{aligned}$$

because cross products are all zero matrces. i.e.

$$\begin{aligned}
\sum_{ijk} a_{ijk} b_{ij}^T &= \sum_{ijk} (x_{ijk} - \bar{x}_{ij})(\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x})^T \\
&= \sum_{ij} \Big\{ \underbrace{\sum_k (x_{ijk} - \bar{x}_{ij})}_{= 0_{p\times 1}} \Big\} (\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x})^T = 0_{p\times p}
\end{aligned}$$

$$\sum_{ijk} a_{ijk} c_j^T = \sum_{ij} \Big\{ \underbrace{\sum_k (x_{ijk} - \bar{x}_{ij})}_{= 0_{p\times 1}} \Big\} (\bar{x}_{\cdot j} - \bar{x})^T = 0_{p\times p}$$

$$\sum_{ijk} a_{ijk} d_i^T = \sum_{ij} \Big\{ \sum_k (x_{ijk} - \bar{x}_{ij}) \Big\} (\bar{x}_{i\cdot} - \bar{x})^T = 0_{p\times p}.$$

$$\begin{aligned}
\sum_{ijk} b_{ij} c_j^T &= \sum_{ijk} (\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x})(\bar{x}_{\cdot j} - \bar{x})^T \\
&= \sum_{jk} \Big\{ \underbrace{\sum_i (\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x})}_{= 0_{p\times 1}} \Big\} (\bar{x}_{\cdot j} - \bar{x})^T \quad \left(\because \begin{array}{l} \sum_i \bar{x}_{ij} = n^{-1} \sum_{ik} x_{ijk}, \ \sum_i \bar{x}_{i\cdot} = (bn)^{-1} \sum_{ijk} x_{ijk} \\ \sum_i \bar{x}_{\cdot j} = n^{-1} \sum_{ik} x_{ijk}, \ \sum_i \bar{x} = (bn)^{-1} \sum_{ijk} x_{ijk} \end{array} \right) \\
&= 0_{p\times p}.
\end{aligned}$$

$$\begin{aligned}
\sum_{ijk} b_{ij} d_i^T &= \sum_{ijk} (\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x})(\bar{x}_{i\cdot} - \bar{x})^T \\
&= \sum_{ik} \Big\{ \underbrace{\sum_j (\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x})}_{= 0_{p\times 1}} \Big\} (\bar{x}_{i\cdot} - \bar{x})^T \quad \left(\because \sum_j \bar{x}_{ij} = n^{-1} \sum_{jk} x_{ijk} = \sum_j \bar{x}_{i\cdot}, \ \sum_j \bar{x}_{\cdot j} = (mn)^{-1} \sum_{ijk} x_{ijk} = \sum_j \bar{x} \right) \\
&= 0_{p\times p}.
\end{aligned}$$

$$\begin{aligned}
\sum_{ijk} c_j d_i^T &= \sum_{ijk} (\bar{x}_{\cdot j} - \bar{x})(\bar{x}_{i\cdot} - \bar{x})^T \\
&= \sum_{ik} \Big\{ \underbrace{\sum_j (\bar{x}_{\cdot j} - \bar{x})}_{= 0_{p\times 1}} \Big\} (\bar{x}_{i\cdot} - \bar{x})^T = 0_{p\times p}
\end{aligned}$$

- Testing interaction
  - Hypotheses $H_0 : \boldsymbol{\gamma}_{11} = \cdots = \boldsymbol{\gamma}_{mb} = \mathbf{0}$ v.s. H_1 : otherwise
  - Wilk's lambda test statistic
    $$\Lambda = \frac{\det \mathbf{SSP}_{\text{res}}}{\det(\mathbf{SSP}_{\text{res}} + \mathbf{SSP}_{\text{int}})}$$
    - Under H_0, by Bartlett's approximation
      $$[\{p + 1 - (m-1)(b-1)\}/2 - mb(n-1)] \ln \Lambda \approx \chi^2((m-1)(b-1))$$
  - Rejection H_0 at level α when
    $$[\{p + 1 - (m-1)(b-1)\}/2 - mb(n-1)] \ln \Lambda \geq \chi^2_{1-\alpha,(m-1)(b-1)}$$
  - p-value
    $$1 - F_{\chi^2((m-1)(b-1))}([\{p + 1 - (m-1)(b-1)\}/2 - mb(n-1)] \ln \Lambda)$$

- Testing main effects
  - Testing factor 1 main effects
    - Hypotheses $H_0 : \boldsymbol{\tau}_1 = \cdots = \boldsymbol{\tau}_m = \mathbf{0}$ v.s. H_1 : otherwise
    - Wilk's lambda test statistic
      $$\Lambda = \frac{\det \mathbf{SSP}_{\text{res}}}{\det(\mathbf{SSP}_{\text{res}} + \mathbf{SSP}_{\text{f1}})}$$
      - Under H_0, by Bartlett's approximation
        $$[\{p + 1 - (m-1)\}/2 - mb(n-1)] \ln \Lambda \approx \chi^2(m-1)$$

* Rejection H_0 at level α when

$$[\{p+1-(m-1)\}/2 - mb(n-1)] \ln \Lambda \geq \chi^2_{1-\alpha, m-1}$$

* p-value

$$1 - F_{\chi^2(m-1)}([\{p+1-(m-1)\}/2 - mb(n-1)] \ln \Lambda)$$

- Testing factor 2 main effects
  * Hypotheses $H_0 : \boldsymbol{\beta}_1 = \cdots = \boldsymbol{\beta}_b = \mathbf{0}$ v.s. H_1 : otherwise
  * Wilk's lambda test statistic

$$\Lambda = \frac{\det \mathbf{SSP}_{\text{res}}}{\det(\mathbf{SSP}_{\text{res}} + \mathbf{SSP}_{\text{f2}})}$$

    · Under H_0, by Bartlett's approximation

$$[\{p+1-(b-1)\}/2 - mb(n-1)] \ln \Lambda \approx \chi^2(b-1)$$

  * Rejection H_0 at level α when

$$[\{p+1-(b-1)\}/2 - mb(n-1)] \ln \Lambda \geq \chi^2_{1-\alpha, b-1}$$

  * p-value

$$1 - F_{\chi^2(b-1)}([\{p+1-(b-1)\}/2 - mb(n-1)] \ln \Lambda)$$

- Exercise: factors in producing plastic film (continued)
  - One more factor `ADDITIVE` (amount of an additive, 2-level, low or high) in the production test

```
options(digits = 4)
tear <- c(
  6.5, 6.2, 5.8, 6.5, 6.5, 6.9, 7.2, 6.9, 6.1, 6.3,
  6.7, 6.6, 7.2, 7.1, 6.8, 7.1, 7.0, 7.2, 7.5, 7.6
)
gloss <- c(
  9.5, 9.9, 9.6, 9.6, 9.2, 9.1, 10.0, 9.9, 9.5, 9.4,
  9.1, 9.3, 8.3, 8.4, 8.5, 9.2, 8.8, 9.7, 10.1, 9.2
)
opacity <- c(
  4.4, 6.4, 3.0, 4.1, 0.8, 5.7, 2.0, 3.9, 1.9, 5.7,
  2.8, 4.1, 3.8, 1.6, 3.4, 8.4, 5.2, 6.9, 2.7, 1.9
)
(X <- cbind(tear, gloss, opacity))
(rate <- factor(gl(2,10,length=nrow(X)), labels=c("Low", "High")))
(additive <- factor(gl(2,5,length=nrow(X)), labels=c("Low", "High")))

summary(manova(X ~ rate*additive), test = 'Wilks')
summary(car::Manova(lm(X ~ rate*additive)), test.statistic='Wilks')
```