Middle East Technical University
Department of Statistics
2020-21 Fall Semester
STAT 467 Multivariate Analysis

December 10, 2020, Thursday

# RECITATION 6

## Comparing Several Multivariate Population Mean (MANOVA)

Assume that we have $g$ populations and would like to compare these,

$Pop1 = X_{11}, \dots\dots X_{1,n}$

..

$Pop\ g = X_{g1}, \dots\dots X_{g,n}$

The MANOVA model can be written as

$X_{lj} = \mu + \tau_l + \epsilon_{lj}$ for $l = 1, \dots g$ and $j = 1,2,\dots, n_l$

where

$$\epsilon_{lj}\ N_p(0, \Sigma)$$

$\sum_{l=1}^{g} n_l \tau_l = 0$ where $\tau_l$ indicates treatment effect.

## MANOVA model has the following assumptions

- The data from all groups have common variance covariance matrix.
- The subjects are drawn from independent samples.
- The data follows multivariate normal distribution.

In MANOVA, we test the following hypothesis

$H_0 = \tau_1 = \tau_2 = \dots = \tau_g = 0$ (No treatment effect) vs $H_1 =$ At least one differs.

The standard MANOVA table can be arranges as follow

MANOVA Table for Comparing Population Mean Vectors

| Source of variation | Matrix of sum of squares and cross products (SSP) | Degrees of freedom (d.f.) |
|---|---|---|
| Treatment | $\mathbf{B} = \sum_{\ell=1}^{g} n_\ell (\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})'$ | $g - 1$ |
| Residual (Error) | $\mathbf{W} = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'$ | $\sum_{\ell=1}^{g} n_\ell - g$ |
| Total (corrected for the mean) | $\mathbf{B} + \mathbf{W} = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})'$ | $\sum_{\ell=1}^{g} n_\ell - 1$ |

To make a final decision about the null hypothesis, we'll use

$$\Lambda^* = \frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|} = \frac{\left|\sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'\right|}{\left|\sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})'\right|}$$

Reject $H_0$ if $\Lambda^*$ is too small (i.e close to 0)

Fail to reject $H_0$ if $\Lambda^*$ is too large. (i.e close to 1)

The quantity $\Lambda^* = |\mathbf{W}|/|\mathbf{B} + \mathbf{W}|$, proposed originally by Wilks , corresponds to the equivalent form of the $F$ -test of $H_0$: no treatment effects in the univariate case. Wilks' lambda has the virtue of being convenient and related to the likelihood ratio criterion. The exact distribution of $\Lambda^*$ can be derived for the special cases. For other cases and large sample sizes, a modification of $\Lambda^*$ due to Bartlett can be used to test $H_0$.

**Multivariate Linear Regression Model**

Consider the problem of modelling the relationship between $m$ responses.

$(Y_1, \dots, Y_m)$ and set the predictors $(Z_1, \dots, Z_m)$

The multivariate linear regression model is

$$\underset{(n\times m)}{\mathbf{Y}} = \underset{(n\times(r+1))}{\mathbf{Z}} \underset{((r+1)\times m)}{\boldsymbol{\beta}} + \underset{(n\times m)}{\mathbf{E}}$$

with

$$E(\boldsymbol{\varepsilon}_{(i)}) = \mathbf{0} \quad \text{and} \quad \operatorname{Cov}(\boldsymbol{\varepsilon}_{(i)}, \boldsymbol{\varepsilon}_{(k)}) = \sigma_{ik}\mathbf{I} \quad i, k = 1, 2, \dots, m$$

The $m$ observations on the $j$ th trial have covariance matrix $\Sigma = \{\sigma_{ik}\}$, but observations from different trials are uncorrelated. Here $\beta$ and $\sigma_{ik}$ are unknown parameters; the design matrix $\mathbf{Z}$ has $j$ th row $[z_{j0}, z_{j1}, \dots, z_{jr}]$ In matrix notation, the design matrix is

$$\underset{(n\times(r+1))}{\mathbf{Z}} = \begin{bmatrix} z_{10} & z_{11} & \cdots & z_{1r} \\ z_{20} & z_{21} & \cdots & z_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n0} & z_{n1} & \cdots & z_{nr} \end{bmatrix}$$

The other quantities are,

$$\underset{(n\times m)}{\mathbf{Y}} = \begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1m} \\ Y_{21} & Y_{22} & \cdots & Y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{nm} \end{bmatrix} = [\mathbf{Y}_{(1)} \mid \mathbf{Y}_{(2)} \mid \cdots \mid \mathbf{Y}_{(m)}]$$

$$\underset{((r+1)\times m)}{\boldsymbol{\beta}} = \begin{bmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0m} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{r1} & \beta_{r2} & \cdots & \beta_{rm} \end{bmatrix} = [\boldsymbol{\beta}_{(1)} \mid \boldsymbol{\beta}_{(2)} \mid \cdots \mid \boldsymbol{\beta}_{(m)}]$$

$$\underset{(n\times m)}{\boldsymbol{\varepsilon}} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1m} \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{nm} \end{bmatrix} = [\boldsymbol{\varepsilon}_{(1)} \mid \boldsymbol{\varepsilon}_{(2)} \mid \cdots \mid \boldsymbol{\varepsilon}_{(m)}]$$

$$= \begin{bmatrix} \boldsymbol{\varepsilon}_1' \\ \hline \boldsymbol{\varepsilon}_2' \\ \hline \vdots \\ \hline \boldsymbol{\varepsilon}_n' \end{bmatrix}$$

The least square estimation of coefficient matrix is

$$\widehat{\boldsymbol{\beta}} = [\widehat{\boldsymbol{\beta}}_{(1)} \vdots \widehat{\boldsymbol{\beta}}_{(2)} \quad \cdots \quad \widehat{\boldsymbol{\beta}}_{(m)}] = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'[\mathbf{Y}_{(1)} \quad \mathbf{Y}_{(2)} \vdots \cdots \vdots \mathbf{Y}_{(m)}]$$

or

$$\widehat{\boldsymbol{\beta}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}$$

**Properties of Estimators**

- $E(\hat{\beta}) = \beta$
- $Cov(\hat{\beta}_{(i)}, \hat{\beta}_{(k)}) = \sigma_{ik}(\mathbf{Z}'\mathbf{Z})^{-1}$
- $Cov(\hat{\beta}_{(i)}, \hat{\epsilon}_{(k)}) = 0$

**Recall**

In simple linear regression, we have one $X$ and one $Y$.

In multiple linear regression, we have more than one $X$ and one $Y$.

In multivariate linear regression, we have more than one $Y$

**QUESTIONS**

**1.** Observations on two responses are collected for three treatments. The observation vector $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ are;

Treatment 1: $\begin{bmatrix} 6 \\ 7 \end{bmatrix}, \begin{bmatrix} 5 \\ 9 \end{bmatrix}, \begin{bmatrix} 8 \\ 6 \end{bmatrix}, \begin{bmatrix} 4 \\ 9 \end{bmatrix}, \begin{bmatrix} 7 \\ 9 \end{bmatrix}$
Treatment 2: $\begin{bmatrix} 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$
Treatment 3: $\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

**a)** Test if there is any significant difference between treatments. State your assumptions for this test.

**b)** If you found a significant treatment effect, find out which treatment this difference is due to

**2.** Consider the following data on one predictor variable and two responses:

| | | | | | |
|---|---|---|---|---|---|
| $\boldsymbol{Z_1}$ | -2 | -1 | 0 | 1 | 2 |
| $\boldsymbol{Y_1}$ | 5 | 3 | 4 | 2 | 1 |
| $\boldsymbol{Y_2}$ | -3 | -1 | -1 | 2 | 3 |

**a)** Determine the least squares estimates of the parameters in the straight line regression model.

**b)** Caculate the matrices of fitted values $\hat{Y}$ and residuals $\hat{\varepsilon}$ with $Y$:[ $\boldsymbol{y_1} : \boldsymbol{y_2}$].Verify the sum of squares and cross products decompositon. $Y'Y = \hat{Y}'\hat{Y} + \hat{\varepsilon}'\hat{\varepsilon}$

STAT 467/ Recitation 6

Answer Key

1) One way MANOVA

| Group | # of obs | Sample mean Vectors | Sample Cov Matrix | $\bar{\bar{X}}$ |
|---|---|---|---|---|
| $l_1$ = trt 1 | $n_1 = 5$ | $\bar{x}_1 = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$ | $S_1 = \begin{bmatrix} 2.5 & -1.5 \\ -1.5 & 2 \end{bmatrix}$ | $\begin{bmatrix} 4 \\ 5 \end{bmatrix}$ |
| $l_2$ = trt 2 | $n_2 = 3$ | $\bar{x}_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ | $S_2 = \begin{bmatrix} 1 & -1.5 \\ & 3 \end{bmatrix}$ | |
| $l_3$ = trt 3 | $n_3 = 4$ | $\bar{x}_3 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ | $S_3 = \begin{bmatrix} 2 & -1.33 \\ & 1.33 \end{bmatrix}$ | $\frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3}{N}$ |

$N = \sum n_l = 12$

$H_0 = \tau_1 = \tau_2 = \tau_3 = 0$ vs $H_1$ = at least one differs

| SOV | matrix of SS | df |
|---|---|---|
| Treatment | B | $g-1$ |
| Residual | W | $n-g$ |
| Total | B+W | $n-1$ |

Trt effect: $B = \sum_{l}^{g} n_l(\bar{x}_l - \bar{x})(\bar{x}_l - \bar{x})'$

$$= 5\begin{bmatrix} 2 \\ 3 \end{bmatrix}[2\ 3] + 3\begin{bmatrix} -2 \\ -1 \end{bmatrix}[-2\ -1] + 4\begin{bmatrix} -1 \\ -3 \end{bmatrix}[-1\ -3]$$

$$= 5\begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix} + 3\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} + 4\begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}$$

$$= \begin{bmatrix} 20 & 30 \\ 30 & 45 \end{bmatrix} + \begin{bmatrix} 12 & 6 \\ 6 & 3 \end{bmatrix} + \begin{bmatrix} 4 & 12 \\ 12 & 36 \end{bmatrix} = \begin{bmatrix} 36 & 48 \\ 48 & 84 \end{bmatrix}$$

$$W = \sum^{g}\sum^{n_l}(x_{lj} - \bar{x}_l)(x_{lj} - \bar{x}_l)' = \sum_{l=1}^{g}(n_l - 1)S_l$$

$$= 4\begin{bmatrix} 2.5 & -1.5 \\ -1.5 & 2 \end{bmatrix} + 2\begin{bmatrix} 1 & -1.5 \\ -1.5 & 3 \end{bmatrix} + 3\begin{bmatrix} +2 & -1.33 \\ -1.33 & 1.33 \end{bmatrix} = \begin{bmatrix} 18 & -12.99 \\ -12.99 & 17.99 \end{bmatrix}$$

$$B+W = \begin{bmatrix} 54 & 35.01 \\ 35.01 & 101.99 \end{bmatrix} \quad \Lambda^* = \frac{|W|}{|B+W|} = \frac{155.0799}{42481.7599} = 0.036$$

→ Reject $H_0$ since it's close to 0.

b) Simultaneous CR for trt effect

$$\tau_{ki}-\tau_{\ell i} = (\bar{x}_{ki}-\bar{x}_{\ell i}) \pm t_{n-g}\left(\frac{\alpha}{2pg\frac{(g-1)}{2}}\right)\sqrt{\left(\frac{1}{n_k}+\frac{1}{n_\ell}\right)\frac{w_{ii}}{n-g}}$$

$$\tau_{11}-\tau_{21} = (\bar{x}_{11}-\bar{x}_{21}) \pm t_{12-3}\left(\frac{0.05}{2.3.\frac{2\,2}{2}}\right)\sqrt{\left(\frac{1}{5}+\frac{1}{3}\right)\frac{18}{12-3}}$$

$$= 4 \pm (3.25)(1.033) \Rightarrow (0.643,\ 7.35) \rightarrow \text{not contain zero}$$

(this makes the difference

$$\tau_{22}-\tau_{32} = 4-2) \pm 3.25\sqrt{\left(\frac{1}{3}+\frac{1}{4}\right)\left(\frac{17.99}{12-3}\right)}$$

$$2 \pm 3.25\,(1.08)$$

$$(-1.51,\ 5.51)\ \times$$

2) a) $n = \#$ of obs, $m = \#$ of resp, $r = \#$ of pred

$$\underset{m\times n}{y} = \underset{(n\times(r+1))}{z}\ \underset{((r+1)\times m)}{\beta} + \underset{n\times m}{\varepsilon}$$

$$y=\begin{bmatrix}5 & -3\\ 3 & -1\\ 4 & -1\\ 2 & 2\\ 1 & 3\end{bmatrix} \quad z=\begin{bmatrix}1 & -2\\ 1 & -1\\ 1 & 0\\ 1 & 1\\ 1 & 2\end{bmatrix} \quad \varepsilon=\begin{bmatrix}\varepsilon_{11} & \varepsilon_{12}\\ \cdot & \cdot\\ \cdot & \cdot\\ \varepsilon_{51} & \varepsilon_{52}\end{bmatrix} \quad \beta=\begin{bmatrix}\beta_{01} & \beta_{02}\\ \beta_{11} & \beta_{12}\end{bmatrix}$$

$$\hat{\beta} = (z'z)^{-1}z'y = \begin{bmatrix}3 & 0\\ -0.9 & 1.5\end{bmatrix}$$

$$(z'z) = \begin{bmatrix}1 & 1 & 1 & 1 & 1\\ -2 & -1 & 0 & 1 & 2\end{bmatrix}\begin{bmatrix}1 & -2\\ 1 & -1\\ 1 & 0\\ 1 & 1\\ 1 & 2\end{bmatrix} = \begin{bmatrix}5 & 0\\ 0 & 10\end{bmatrix} \quad (z'z)^{-1} = \frac{1}{50}\begin{bmatrix}10 & 0\\ 0 & 5\end{bmatrix} = \begin{bmatrix}0.2 & 0\\ 0 & 0.1\end{bmatrix}$$

$$z'y = \begin{bmatrix}1 & 1 & 1 & 1 & 1\\ -2 & -1 & 0 & 1 & 2\end{bmatrix}\begin{bmatrix}5 & -3\\ 3 & -1\\ 4 & -1\\ 2 & 2\\ 1 & 3\end{bmatrix} = \begin{bmatrix}15 & 0\\ -9 & 15\end{bmatrix}$$

b) Hence, the predicted values

$$\hat{y} = Z\hat{\beta}$$

$$\hat{y} = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ -0.9 & 1.5 \end{bmatrix} = \begin{bmatrix} 4.8 & -3.0 \\ 3.9 & -1.5 \\ 3 & 0 \\ 2.1 & 1.5 \\ 1.2 & 3.0 \end{bmatrix}$$

The residual matrix

$$\hat{\varepsilon} = y - \hat{y} = \begin{bmatrix} 0.2 & 0 \\ -0.9 & 0.5 \\ 1 & -1 \\ 0.1 & 0.5 \\ -0.2 & 0 \end{bmatrix}$$

Bunu sor.

$$y'y = \hat{y}'\hat{y} + \hat{\varepsilon}'\hat{\varepsilon}$$

$$\begin{bmatrix} 55 & -15 \\ -15 & 24 \end{bmatrix} = \begin{bmatrix} 53.1 & -13.5 \\ -13.5 & 22.5 \end{bmatrix} + \begin{bmatrix} 1.9 & -1.5 \\ -1.5 & 1.5 \end{bmatrix}$$