# MANOVA

Instructor: Cheng-Ying Chou

# The Multivariate Analysis of Variance

- Introduction
- Review of Univariate One-way Analysis of Variance (1-way ANOVA)
- Multivariate One-way Analysis of Variance (1-way MANOVA)
- Homogeneity of Covariance Matrices
- Summary

# Introduction

- Example
  - Evaluation of four different teaching methods
  - 40 pupils in a primary school were randomly assigned to one of four methods.
  - After a 12-week period, the scores of the tests on reading comprehension (X1) and analytical ability (X2) were obtained (Shen, 1998)

# Introduction

| | Methods | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | I | | II | | III | | IV | |
| | X1 | X2 | X1 | X2 | X1 | X2 | X1 | X2 |
| | 7 | 9 | 43 | 21 | 24 | 14 | 27 | 17 |
| | 37 | 16 | 39 | 17 | 30 | 15 | 20 | 13 |
| | 44 | 16 | 21 | 9 | 17 | 8 | 25 | 15 |
| | 28 | 14 | 32 | 16 | 34 | 15 | 13 | 11 |
| | 42 | 16 | 21 | 9 | 32 | 16 | 32 | 17 |
| | 23 | 11 | 14 | 6 | 35 | 17 | 4 | 10 |
| | 45 | 25 | 24 | 11 | 19 | 11 | 3 | 7 |
| | 32 | 15 | 24 | 12 | 6 | 6 | 36 | 19 |
| | 37 | 16 | 26 | 12 | 30 | 14 | 28 | 16 |
| | 45 | 19 | 26 | 12 | 19 | 10 | 21 | 15 |
| mean | 34.0 | 15.7 | 27.0 | 12.5 | 24.6 | 12.6 | 20.9 | 14.0 |

# Introduction

- Univariate Questions
  - Does the score of reading comprehension (or analytical ability) differ among four teaching methods.
- Multivariate Questions
  - Is there any difference in the effectiveness as measured by reading comprehension and analytical ability among the four teaching methods?

# Introduction

- Dataset: Measurements of Egyptian Skull
  - Univariate Questions:
    - Is there any difference or change in each of the four measurements across time?
  - Multivariate Questions:
    - Is there any difference or change in the four measurements across time?

# Review of Univariate 1-way ANOVA

- Consider one treatment factor: such as teaching methods or time period

- $X_{ij}$: the measurement of object i in treatment j, $i = 1,\ldots,n_j$; $j = 1,\ldots,m$
- The model for one-way ANOVA is

  $X_{ij} = \mu + \tau_j + \varepsilon_{ij}$, where

  $\mu$: overall mean

  $\tau_j$: the jth treatment effect

  $\varepsilon_{ij}$: random error $\sim N(0, \sigma^2)$

# One way ANOVA

- $H_0$: $\mu_1 = \mu_2 = \cdots = \mu_m = \mu$
- $H_1$: $\mu_1 \neq \mu_2 \neq \cdots \neq \mu_m$

$$X_{11}, X_{21}, \cdots X_{n_1 1} \sim N(\mu_1, \Sigma)$$

$$X_{12}, X_{22}, \cdots X_{n_2 2} \sim N(\mu_2, \Sigma)$$

$$\vdots \qquad\qquad \vdots$$

$$X_{1m}, X_{2m}, \cdots X_{n_m m} \sim N(\mu_m, \Sigma)$$

# Review of Univariate 1-way ANOVA

Sums of squares for one-way ANOVA

Treatment means: $\overline{X}_{j.} = \sum_{i=1}^{n_j} X_{ij} / n_j$

Overall mean: $\overline{X} = \sum_{j=1}^{m} \overline{X}_{j.} / m$

Partition of sums of squares

$$\sum_{j=1}^{m}\sum_{i=1}^{n_j}(X_{ij} - \overline{X})^2 = \sum_{j=1}^{m} n_j(\overline{X}_{j.} - \overline{X})^2 + \sum_{j=1}^{m}\sum_{i=1}^{n_j}(X_{ij} - \overline{X}_{j.})^2$$

SST (Total) = SSB (Between) + SSW(Within)

n-1 = (m-1) + (n-m)

where $n = \sum_{j=1}^{m} n_j$

# Review of Univariate 1-way ANOVA

$H_0$: $\tau_1 = \ldots = \tau_m$ vs. $H_a$: $\tau_j \neq \tau_{j'}$ for at least one pair $1 = j \neq j' \leq m$

<u>One-way ANOVA Table</u>

| SOV | df | SS | MS |
|---|---|---|---|
| Between samples | m-1 | SSB | MSB=SSB/(m-1) |
| Within samples | n-m | SSW | MSW=SSW/(n-m) |
| Total | n-1 | SST | |

Reject the null hypothesis if $F = MSB/MSW > F_{\alpha, m-1, n-m}$

# Review of Univariate 1-way ANOVA

- Example: Reading Comprehension
  - $n=40$, $n1=n2=n3=n4=10$, grand total is 1065
  - $SST = 7^2 + \ldots + 21^2 - (1065)^2/40 = 4815.375$ ($dfT=39$)
  - $SSB = (340^2 + \ldots + 209^2)/10 - (1065)^2/40$
    $= 914.075$ ($dfB=3$)
  - $SSW = 4815.375 - 914.075 = 3901.300$ ($dfW=36$)

# R code

```
setwd("/Users/cychou/Documents/Work/Teaching/Exploratory Multivariate Data Analysis/Lecture Slides/R_code");
methods<-read.csv("methods.csv",header=T,sep=",");
mod1<-lm(X1~method,data=methods)
anova(mod1)
```

```
Response: X1
          Df Sum Sq Mean Sq F value Pr(>F)
method     3  914.1  304.69  2.8116 0.0531 .
Residuals 36 3901.3  108.37
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# Review of Univariate 1-way ANOVA

- ANOVA Table for Reading Comprehension

| SOV | df | SS | MS | F |
|---|---|---|---|---|
| Between | 3 | 914.075 | 304.692 | 2.81 |
| Within | 36 | 3901.300 | 108.369 | |
| Total | 39 | 4815.375 | | |

- Since $F=2.81 < F_{0.05,3,36} = 2.90$, fail to reject the null hypothesis of equal teaching effectiveness at the 5% significance level

# Multivariate 1-way ANOVA

- Consider one treatment factor: such as teaching methods or time period
- $X_{ijk}$: the measurement of variable k for object i in treatment j, $i = 1,\ldots,n_j$; $j = 1,\ldots,m$; $k = 1,\ldots,p$
- The model for one-way MANOVA is

$$X_{ijk} = \mu + \tau_{jk} + \epsilon_{ijk}$$

where $\mu$: overall mean

$\tau_{jk}$: the jth treatment effect

# One way MANOVA

- $H_0$: 
$$\begin{aligned} \mu_{11} &= \mu_{21} = \cdots = \mu_{m1} \\ \mu_{12} &= \mu_{22} = \cdots = \mu_{m2} \\ &\vdots \qquad\quad \vdots \\ \mu_{1p} &= \mu_{2p} = \cdots = \mu_{mp} \end{aligned}$$

$$\begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{1p} \end{pmatrix} = \begin{pmatrix} \mu_{21} \\ \mu_{22} \\ \vdots \\ \mu_{2p} \end{pmatrix} = \cdots \begin{pmatrix} \mu_{m1} \\ \mu_{m2} \\ \vdots \\ \mu_{mp} \end{pmatrix}$$

Measured variable k=1 .........p

X1 X2

Treatment j
1
2
.
.
m

Teaching methods
I, II, III, IV

$$Y_{11} = \begin{pmatrix} Y_{111} \\ Y_{211} \\ \vdots \\ Y_{n_1 11} \end{pmatrix} \cdots Y_{1p} = \begin{pmatrix} Y_{11p} \\ Y_{21p} \\ \vdots \\ Y_{n_m 1p} \end{pmatrix}$$

$$\vdots \qquad \ddots \qquad \vdots$$

$$Y_{m1} = \begin{pmatrix} Y_{11p} \\ Y_{21p} \\ \vdots \\ Y_{n_1 1p} \end{pmatrix} \cdots Y_{mp} = \begin{pmatrix} Y_{1mp} \\ Y_{2mp} \\ \vdots \\ Y_{n_m mp} \end{pmatrix}$$

# Multivariate 1-way ANOVA

The overall mean for variable k

$$\overline{X}_k = \sum_{j=1}^{m}\sum_{i=1}^{n_j} X_{ijk} / n$$

The treatment mean for treatment j of variable k

$$\overline{X}_{jk} = \sum_{i=1}^{n_j} X_{ijk} / n_j$$

Total sums of squares and cross-products

$$t_{kk} = \sum_{j=1}^{m}\sum_{i=1}^{n_j} (X_{ijk} - \overline{X}_k)^2,\ \mathrm{dfT{=}n\text{-}1}$$

$$t_{kk'} = \sum_{j=1}^{m}\sum_{i=1}^{n_j} (X_{ijk} - \overline{X}_k)(X_{ijk'} - \overline{X}_{k'}), \mathrm{dfT} = \mathrm{n\text{-}1}$$

# Multivariate 1-way ANOVA

Within-sample sums of squares and cross-products

$$w_{kk} = \sum_{j=1}^{m} \sum_{i=1}^{n_j} (X_{ijk} - \overline{X}_{jk})^2, dfW = n\text{-}m$$

$$w_{kk'} = \sum_{j=1}^{m} \sum_{i=1}^{n_j} (X_{ijk} - \overline{X}_{jk})(X_{ijk'} - \overline{X}_{jk'}), dfW = n\text{-}m$$

Between-sample sums of squares and cross-products

$b_{kk} = t_{kk} - w_{kk}$, $dfB = m\text{-}1$, and

$b_{kk'} = t_{kk'} - w_{kk'}$, $dfB = m\text{-}1$

# Multivariate 1-way ANOVA

Total sums of squares and cross-products matrix

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1p} \\ t_{12} & t_{22} & \cdots & t_{2p} \\ .. & .. & \cdots & .. \\ t_{1p} & t_{2p} & \cdots & t_{pp} \end{pmatrix}, \text{dfT=n-1}$$

$$t_{kk'} = \sum_{j=1}^{m} \sum_{i=1}^{n_j} (X_{ijk} - \overline{X}_k)(X_{ijk'} - \overline{X}_{k'})$$

Within-sample sums of squares and cross-products matrix

$$\mathbf{W} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1p} \\ w_{12} & w_{22} & \cdots & w_{2p} \\ .. & .. & \cdots & .. \\ w_{1p} & w_{2p} & \cdots & w_{pp} \end{pmatrix}, \text{dfW=n-m}$$

$$w_{kk'} = \sum_{j=1}^{m} \sum_{i=1}^{n_j} (X_{ijk} - \overline{X}_{jk})(X_{ijk'} - \overline{X}_{jk'})$$

Between-sample sums of squares and cross-products matrix

$$\mathbf{B} = \mathbf{T} - \mathbf{W}, \text{dfB} = m - 1$$

# Multivariate 1-way ANOVA

- The Wilks' lambda statistic

$$\Lambda = |\mathbf{W}| / |\mathbf{T}| = \frac{|\mathbf{W}|}{|\mathbf{W}+\mathbf{B}|},$$

where

$|\mathbf{W}|$ is the determinant of $\mathbf{W}$ and $|\mathbf{T}|$ is the determinant of $\mathbf{T}$.

$0 \le \Lambda \le 1$

Small sample distribution

$$F = \frac{1-\Lambda^{1/t}}{\Lambda^{1/t}} \frac{df_2}{df_1} \sim F_{df_1, df_2}, \text{ where } df_1 = p(m\text{-}1),\ df_2 = wt\text{-}(df_1/2) + 1,$$

$w = n\text{-}1\text{-}[(p\text{-}m+2)/2]$, and

$$t=\sqrt{(df_1^2-4)/[p^2+(m-1)^2-5]},$$

when $df_1 = 2$, set t=1.

# Multivariate 1-way ANOVA

- The Wilk's lambda statistic

Special Cases:

(1) When dfB=2 and p is arbitrary

$$F = \frac{\sqrt{\Lambda}}{1-\sqrt{\Lambda}} \frac{\text{dfE-p+1}}{\text{p}} \sim F_{2p,2(\text{dfE-p+1})}$$

(2) When p=2 and dfB is arbitrary

$$F = \frac{\sqrt{\Lambda}}{1-\sqrt{\Lambda}} \frac{\text{dfE-1}}{\text{dfB}} \sim F_{2\text{dfB},2(\text{dfE-1})}$$

Large sample approximation

$$\chi^2 = -[\text{dfE}-(\text{p}-\text{dfB}+1)/2]\ln\Lambda \sim \chi^2_{\text{pdfB}}$$

# Transformation of Wilks' to Exact Upper Tail F tests

| Parameters P, dfT | Statistic having F-distribution | Degrees of freedom |
|---|---|---|
| Any p, dfB=1 | $\frac{1-\Lambda}{\Lambda}\frac{dfW-p+1}{p}$ | P, dfW-p+1 |
| Any p, dfB=2 | $\frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\frac{dfW-p+1}{p}$ | 2p, 2(dfW-p+1) |
| p=1, any dfB | $\frac{1-\Lambda}{\Lambda}\frac{dfW}{dfB}$ | dfB,dfW |
| p=2, any dfB | $\frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\frac{dfW-1}{dfB}$ | 2dfB, 2(dfW-1) |

# Multivariate 1-way ANOVA

- The Roy's largest eigenvalue

Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$ of eignvalues of $\mathbf{W}^{-1}\mathbf{B}$,

the Wilk's lambda can be expressed as

$$\Lambda = \prod_{i=1}^{p} \frac{1}{(1+\lambda_i)}$$

The Roy's largest eigenvalue is $\lambda_1$

Roy's largest root test $\quad \theta = \dfrac{\lambda_1}{1+\lambda_1}$

The distribution of $\lambda_1$

$(df_2 / df_1)\lambda_1 \sim F_{df_1, df_2}$, where

$df_1 = d,\ df_2 = n - m - d - 1,\ d=\max(p, m-1).$

d=max(p,dfB)

# Multivariate 1-way ANOVA

- The Pillai's trace statistic

The Pillai's statistic V is given as

$$V = \sum_{i=1}^{P} \frac{\lambda_i}{(1+\lambda_i)}$$

The distribution of V

$$(n-m-p+s)V/[d(s-V)] \sim F_{df_1, df_2}, \text{where}$$

$$df_1 = sd,\ df_2 = s(n-m-p-s),$$

s=min(p, m-1), and d=max(p,m-1).

# Multivariate 1-way ANOVA

- The Lawes-Hotelling statistic

The Lawley-Hotelling statistic U is given as

$$U=\sum_{i=1}^{p}\lambda_i$$

The distribution of U

$$df_2 U / s df_1 \sim F_{df_1, df_2}, \text{where}$$

$$df_1 = s(2A+s+1),\ df_2 = 2(sB+1),$$

$$A=(|m\text{-}p\text{-}1|\text{-}1)/2, \text{ and } B=(n\text{-}m\text{-}p\text{-}1)/2.$$

# Introduction

Methods

| | I | | II | | III | | IV | | |
|---|---|---|---|---|---|---|---|---|---|
| | X1 | X2 | X1 | X2 | X1 | X2 | X1 | X2 | |
| | 7 | 9 | 43 | 21 | 24 | 14 | 27 | 17 | |
| | 37 | 16 | 39 | 17 | 30 | 15 | 20 | 13 | |
| | 44 | 16 | 21 | 9 | 17 | 8 | 25 | 15 | |
| | 28 | 14 | 32 | 16 | 34 | 15 | 13 | 11 | |
| | 42 | 16 | 21 | 9 | 32 | 16 | 32 | 17 | |
| | 23 | 11 | 14 | 6 | 35 | 17 | 4 | 10 | |
| | 45 | 25 | 24 | 11 | 19 | 11 | 3 | 7 | |
| | 32 | 15 | 24 | 12 | 6 | 6 | 36 | 19 | |
| | 37 | 16 | 26 | 12 | 30 | 14 | 28 | 16 | |
| | 45 | 19 | 26 | 12 | 19 | 10 | 21 | 15 | |
| Sum | 340 | 157 | 270 | 125 | 246 | 126 | 209 | 140 | 1065 548 |
| mean | 34.0 | 15.7 | 27.0 | 12.5 | 24.6 | 12.6 | 20.9 | 14.0 | |

$$\overline{X}_1 = \begin{pmatrix} 34.0 \\ 15.7 \end{pmatrix}, \overline{X}_2 = \begin{pmatrix} 27.0 \\ 12.5 \end{pmatrix}, \overline{X}_3 = \begin{pmatrix} 24.6 \\ 12.6 \end{pmatrix}, \overline{X}_4 = \begin{pmatrix} 20.9 \\ 14.0 \end{pmatrix}$$

# Multivariate 1-way ANOVA

- Example: Teaching Methods

Total sum of squares and cross-products

$$t_{11} = 7^2 + 37^2 + \ldots + 21^2 - (1065)^2 / 40 = 4815.375$$

$$t_{12} = 7x9+37x16+\ldots+21x15 -(1065)(548)/40=1536.500$$

$$t_{22} = 9^2 + 16^2 + \ldots + 15^2 - (548)^2 / 40 = 654.400$$

$$\mathbf{T} = \begin{pmatrix} 4815.375 & 1536.500 \\ & 654.400 \end{pmatrix}, \text{ dfT=39}$$

# Multivariate 1-way ANOVA

- Example: Teaching Methods

Between-sample sum of squares and cross-products

$b_{11} = (340^2 + 270^2 + ... + 209^2)/10 - (1065)^2/40 = 914.075$

$b_{12} =$ (340x157+270x125+...+209x140)/10 -(1065)(548)/40=148.100

$b_{22} = (157^2 + 125^2 + ... + 140^2)/10 - (548)^2/40 = 67.40$

$$\mathbf{B} = \begin{pmatrix} 914.075 & 148.10 \\ & 67.40 \end{pmatrix}, \text{ dfB=3.}$$

$$\mathbf{W} = \mathbf{T} - \mathbf{B} = \begin{pmatrix} 3901.300 & 1388.400 \\ & 587.000 \end{pmatrix}, \text{ dfW=36}$$

# Teaching methods

```
# ANOVA
setwd("/Users/cychou/Documents/Work/Teaching/Exploratory Multivariate Data Analysis/Lecture Slides/R_code");
methods<-read.csv("methods.csv",header=T,sep=",");
methods$method<-as.factor(methods$method)
mod1<-lm(X1~method,data=methods)
anova(mod1)
```

```
# MANOVA
install.packages("dplyr")
library(dplyr)

combined<-cbind(methods$X1,methods$X2)
mod2<-manova(combined~method,data=methods)
summary(mod2, test="Pillai")
summary(mod2, test="Wilks")
summary(mod2, test="Roy")
summary(mod2, test="Hotelling-Lawley")
```

# Multivariate 1-way ANOVA

- Example: Teaching Methods

The Wilk's lamba statistic:

$$|\mathbf{T}| = 730349.15, \text{ and } |\mathbf{E}| = 362408.54$$

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{T}|} = \frac{362408.54}{730349.15} = 0.458542$$

$$p = 2$$

$$F = \frac{1-\sqrt{0.458542}}{\sqrt{0.458542}} \frac{40-4-1}{4-1} = 5.5622$$

$$> F_{0.01,\,6,70} = 3.07$$

$$df_1 = 2dfB = 2(4-1) = 6$$

$$df_2 = 2(dfW-1) = 2(40-4-1) = 70$$

```
> summary(mod2, test="Wilks")
          Df   Wilks approx F num Df den Df    Pr(>F)
method     3 0.45854   5.5622      6     70 9.248e-05 ***
Residuals 36
```

# Multivariate 1-way ANOVA

- Example: Teaching Methods

The Roy's largest eigenvalue:

The two eigenvalues of $\mathbf{W}^{-1}\mathbf{B}$ are 0.9569463 and 0.11440143 and d=max(p,m-1)=max(2,4-1)=3

$$F=\frac{(40\text{-}4\text{-}3\text{-}1)0.9569463}{3}$$

$$=10.20742$$

$$> F_{0.05,3,32} = 2.90$$

Reject the null hypothesis of equal effectiveness among four teaching methods at the 5% significance level.

```
> summary(mod2, test="Roy")
          Df    Roy approx F num Df den Df   Pr(>F)
method     3 0.95695   11.483      3     36 1.98e-05 ***
Residuals 36
```

# Multivariate 1-way ANOVA

- Example: Egyptian Skull

Mean Vectors

| Sample | X1 | X2 | X3 | X4 |
|---|---|---|---|---|
| Early predynastic | 131.37 | 133.60 | 99.17 | 50.53 |
| Late predynastic | 132.37 | 132.70 | 99.07 | 50.23 |
| 12th and 13th dynasties | 134.47 | 133.80 | 96.03 | 50.57 |
| Ptolemaic period | 135.50 | 132.30 | 94.53 | 51.97 |
| Roman period | 136.17 | 130.33 | 93.50 | 51.37 |

# Multivariate 1-way ANOVA

- Example: Egyptian Skull

Total sums of squares and cross-products

$$\mathbf{T} = \begin{pmatrix} 3563.89 & -222.81 & -615.16 & 426.73 \\ & 3635.16 & 1046.28 & 346.47 \\ & & 4309.27 & -16.40 \\ & & & 1533.33 \end{pmatrix}$$

dfT=149

Within-sample sums of squares and cross-products

$$\mathbf{W} = \begin{pmatrix} 3061.07 & 5.33 & 11.47 & 291.30 \\ & 3405.27 & 754.00 & 412.53 \\ & & 3505.97 & 164.33 \\ & & & 1472.13 \end{pmatrix}$$

dfW=145

# Egyptian skull MANOVA

```
install.packages("HSAUR")
library(HSAUR) # load the library, must be installed
before
data("skulls", package = "HSAUR") # load into the
workspace
skulls # see the content of the data.frame
help(skulls)
summary(skulls)
#write.csv(skulls, file = "skulls.csv",row.names=FALSE)
Y<-cbind(skulls$mb, skulls$bh, skulls$bl, skulls$nh);
X<-skulls$epoch;
fit<-manova(Y~X, data=skulls)
summary(fit, test="Pillai")
summary(fit, test="Wilks")
summary(fit, test="Roy")
summary(fit, test="Hotelling-Lawley")
```

# Multivariate 1-way ANOVA

- Example: Egyptian Skull

Pillai's trace statistic V=0.3533

d=max(p,m-1)=max(4,5-1)=4

s=min(p,m-1)=min(4,5-1)=4

$df_1$ = sd=4x4=16

$df_2$ = s(n-m-p+s)=4(150-5-4+4)=580

F=(n-m-p+s)V/[d(s-V)]

=(150-5-4+4)0.3533/[4(4-0.3533)]

=3.51 >$F_{0.05,16,580}$ = 1.42

Reject the null hypothesis of equal mean measurements over time.

```
> summary(fit, test="Pillai")
           Df  Pillai approx F num Df den Df    Pr(>F)
X           4 n0.35331    3.512     16    580  4.675e
Residuals 145
```

# Multivariate 1-way ANOVA

- Example: Egyptian Skull

Lawley-Hotelling statistic U=0.4818

$s=\min(p,m-1)=\min(4,5-1)=4$

$A=(|m\text{-}p\text{-}1|-1)/2=(|5\text{-}4\text{-}1|-1)/2=-0.5$

$B=(n\text{-}m\text{-}p\text{-}1)/2=(150\text{-}5\text{-}4\text{-}1)/2=70$

$df_1=s(2A+s+1)=4(-1+4+1)=16$

$df_2=2(sB+1)=2(4\text{x}70+1)=562$

$F=df_2U/df_1$

$=562\text{x}0.4818/16$

$=4.23>F_{0.05,16,562}=1.423$

Reject the null hypothesis of equal mean measurements over time.

```
> summary(fit, test="Hotelling-Lawley")
           Df Hotelling-Lawley approx F num Df den Df    Pr(>F)
X           4          0.48182    4.231     16    562 8.278e-08 ***
Residuals 145
```

In a classical experiment carried out from 1918 to 1934, apple trees of different rootstocks were compared. The data for eight trees from each of six rootstocks are given in rootstock.dat. The variables are

- y1 = trunk girth at 4 years (mm x 100)
  y2 = extension growth at 4 years (m)
  y3 = trunk girth at 15 years (mm x 100)
  y4 = weight of tree above ground at 15 years (lb x 1000)

**Table 6.2** Rootstock Data

| Rootstock | $y_1$ | $y_2$ | $y_3$ | $y_4$ |
|---|---|---|---|---|
| 1 | 1.11 | 2.569 | 3.58 | .760 |
| 1 | 1.19 | 2.928 | 3.75 | .821 |
| 1 | 1.09 | 2.865 | 3.93 | .928 |
| 1 | 1.25 | 3.844 | 3.94 | 1.009 |
| 1 | 1.11 | 3.027 | 3.60 | .766 |
| 1 | 1.08 | 2.336 | 3.51 | .726 |
| 1 | 1.11 | 3.211 | 3.98 | 1.209 |
| 1 | 1.16 | 3.037 | 3.62 | .750 |
| 2 | 1.05 | 2.074 | 4.09 | 1.036 |
| 2 | 1.17 | 2.885 | 4.06 | 1.094 |
| 2 | 1.11 | 3.378 | 4.87 | 1.635 |
| 2 | 1.25 | 3.906 | 4.98 | 1.517 |
| 2 | 1.17 | 2.782 | 4.38 | 1.197 |
| 2 | 1.15 | 3.018 | 4.65 | 1.244 |
| 2 | 1.17 | 3.383 | 4.69 | 1.495 |
| 2 | 1.19 | 3.447 | 4.40 | 1.026 |

The matrices $\mathbf{H}$, $\mathbf{E}$, and $\mathbf{E}+\mathbf{H}$ are given by

$$\mathbf{H} = \begin{pmatrix} .074 & .537 & .332 & .208 \\ .537 & 4.200 & 2.355 & 1.637 \\ .332 & 2.355 & 6.114 & 3.781 \\ .208 & 1.637 & 3.781 & 2.493 \end{pmatrix},$$

$$\mathbf{E} = \begin{pmatrix} .320 & 1.697 & .554 & .217 \\ 1.697 & 12.143 & 4.364 & 2.110 \\ .554 & 4.364 & 4.291 & 2.482 \\ .217 & 2.110 & 2.482 & 1.723 \end{pmatrix},$$

$$\mathbf{E}+\mathbf{H} = \begin{pmatrix} .394 & 2.234 & .886 & .426 \\ 2.234 & 16.342 & 6.719 & 3.747 \\ .886 & 6.719 & 10.405 & 6.263 \\ .426 & 3.747 & 6.263 & 4.216 \end{pmatrix}.$$

# Wilks' statistic

# Pillai’s statistic

# Lawley-Hotelling statistic

# Roy's test statistic