Inferences about a Mean Vector

- In the following lectures, we test hypotheses about a \( p \times 1 \) population mean vector \( \mu = (\mu_1, \mu_2, \ldots, \mu_p)' \).
- We could test \( p \) disjoint hypothesis (one for each \( \mu_j \) in \( \mu \)) but that would not take advantage of the correlations between the measured traits (\( X_1, X_2, \ldots, X_p \)).
- We first review hypothesis testing in the univariate case, and then develop the multivariate Hotelling's \( T^2 \) statistic and the likelihood ratio statistic for multivariate hypothesis testing.
- We consider applications to repeated measures (longitudinal) studies.
- We also consider situations when data are incomplete (data are missing at random).

Approaches to Multivariate Inference

- Define a reasonable distance measure. An estimated mean vector that is too "far away" from the hypothesized mean vector \( \mu_0 \) gives evidence against the null hypothesis.
- Construct a likelihood ratio test based on the multivariate normal distribution.
- "Union-Intersection" approach: Consider a univariate test of \( H_0 : a'\mu = a'\mu_0 \) versus \( H_a : a'\mu \neq a'\mu_0 \) for some linear combination of the traits \( a'X \). Optimize over possible values of \( a \).

Review of Univariate Hypothesis Testing

- Is \( \mu_0 \) a plausible value for the population mean \( \mu \)?
- We formulate the problem as a hypothesis testing problem. The competing hypotheses are
  \[ H_0 : \mu = \mu_0 \text{ and } H_a : \mu \neq \mu_0. \]
- Given a sample \( X_1, \ldots, X_n \) from a normal population, we compute the test statistic
  \[ t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}. \]
- If \( t \) is 'small', then \( \bar{X} \) is close enough to \( \mu_0 \) to suggest that \( \mu_0 \) is plausible and we fail to reject \( H_0 \).

Univariate Hypothesis Testing (cont'd)

- When \( H_0 \) is true, the statistic \( t \) has a student \( t \) distribution with \( n - 1 \) degrees of freedom. We reject the null hypothesis at level \( \alpha \) when \( |t| > t_{(n-1)}(\alpha/2) \).
- Notice that rejecting \( H_0 \) when \( t \) is large is equivalent to rejecting \( H_0 \) when the squared standardized distance
  \[ t^2 = \frac{(\bar{X} - \mu_0)^2}{s^2/\bar{n}} = n(\bar{X} - \mu_0)(s^2)^{-1}(\bar{X} - \mu_0) \]
  is large.
- We reject \( H_0 \) when
  \[ n(\bar{X} - \mu_0)(s^2)^{-1}(\bar{X} - \mu_0) > t_{(n-1)}^2(\alpha/2) \]
  i.e., the squared standardized distance exceeds the upper \( \alpha \) percentile of a central F-distribution with \( (1, n - 1) \) df.
Univariate Hypothesis Testing (cont’d)

• If we fail to reject $H_0$, we conclude that $\mu_0$ is close (in units of standard deviations of $\bar{X}$) to $\bar{X}$, and thus is a plausible value for $\mu$.

• The set of plausible values for $\mu$ is the set of all values that lie in the $100(1-\alpha)\%$ confidence interval for $\mu$:
  \[ \bar{x} - t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}} \leq \mu_0 \leq \bar{x} + t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}}. \]

• The confidence interval consists of all the $\mu_0$ values that would not be rejected by the $\alpha$ level test of $H_0 : \mu = \mu_0$.

• Before collecting the data, the potential interval is random and has probability $1 - \alpha$ of containing the true population mean $\mu$.

Hotelling’s $T^2$ Statistic

• Consider now the problem of testing whether the $p \times 1$ vector $\mu_0$ is plausible for the population mean vector $\mu$.

• The squared distance
  \[ T^2 = (\bar{X} - \mu_0)'\left(\frac{1}{n}S\right)^{-1}(\bar{X} - \mu_0) = n(\bar{X} - \mu_0)'S^{-1}(\bar{X} - \mu_0) \]

  is called the Hotelling $T^2$ statistic.

• In the expression above,
  \[ \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i, \quad S = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(X_i - \bar{X})' \]

Hotelling’s $T^2$ Statistic (cont’d)

• As we noted earlier,
  \[ T^2 = (\bar{X} - \mu_0)'\left(\frac{1}{n}S\right)^{-1}(\bar{X} - \mu_0) = n(\bar{X} - \mu_0)'S^{-1}(\bar{X} - \mu_0) \]

  has an approximate central chi-square distribution with $p$ df when $\mu_0$ is correct.

• The exact F-distribution relies on the normality assumption.

• We reject the null hypothesis $H_0 : \mu = \mu_0$ for the $p$-dimensional vector $\mu$ at level $\alpha$ when
  \[ T^2 > \frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha), \]

  where $F_{p,n-p}(\alpha)$ is the upper $\alpha$ percentile of the central $F$ distribution with $p$ and $n - p$ degrees of freedom.

• Note that
  \[ \frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha) > \chi^2_p(\alpha) \]

  but these quantities are nearly equal for large values of $n - p$.  

275

276

277

278
Example 5.2: Female Sweat Data

- Perspiration from a sample of 20 healthy females was analyzed. Three variables were measured for each woman:
  - $X_1 =$ sweat rate
  - $X_2 =$ sodium content
  - $X_3 =$ potassium content

- The question is whether $\mu_0 = [4, 50, 10]'$ is plausible for the population mean vector.

Example 5.2: Sweat Data (cont’d)

- At level $\alpha = 0.1$, we reject the null hypothesis if
  $$T^2 = 20(\bar{x} - \mu_0)'S^{-1}(\bar{x} - \mu_0) \overset{\cdot\cdot}{>} \frac{(n-1)p}{(n-p)}F_{p,n-p}(0.1) = \frac{19(3)}{17}F_{3,17}(0.1) = 8.18.$$ 

- From the data displayed in Table 5.1:
  $$\bar{x} = \begin{bmatrix} 4.64 \\ 45.4 \\ 9.96 \end{bmatrix} \quad \text{and} \quad \bar{x} - \mu_0 = \begin{bmatrix} 4.64 - 4 \\ 45.4 - 50 \\ 9.96 - 10 \end{bmatrix} = \begin{bmatrix} 0.64 \\ -4.6 \\ -0.04 \end{bmatrix}.$$ 

/* This program computes a one-sample Hotelling T-squared test for the female sweat data in Example 5.2. It is posted as sweat.sas in the SAS code folder of the course web page. */

```sas
data set1;
  input subject x1-x3;
  label x1 = "sweat rate";
  label x2 = "sodium";
  label x3 = "potassium";
  datalines;
  1 3.7 48.5 9.3  
  2 5.7 65.1 8.0  
  3 3.8 47.2 10.9  
  4 3.2 53.2 12.0 
run;
```
ods rtf file="sweat.rtf";
ods graphics on;
/* Establish the null hypothesis */
data set1; set set1;
d1=x1-4;
d2=x2-50;
d3=x3-10;
z=1;
run;
/* Print the data file */
proc print data=set1;
title "Sweat Rate Data";
run;
/* Check Univariate Normality */
proc univariate data=set1 normal;
var d1-d3; qqplot;
run;

<table>
<thead>
<tr>
<th>Obs</th>
<th>subject</th>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>d1</th>
<th>d2</th>
<th>d3</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3.7</td>
<td>48.5</td>
<td>9.3</td>
<td>-0.3</td>
<td>-1.5</td>
<td>-0.7</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>5.7</td>
<td>65.1</td>
<td>8.0</td>
<td>1.7</td>
<td>15.1</td>
<td>-2.0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3.8</td>
<td>47.2</td>
<td>10.9</td>
<td>-0.2</td>
<td>-2.8</td>
<td>-0.9</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>19</td>
<td>4.1</td>
<td>44.1</td>
<td>11.2</td>
<td>0.1</td>
<td>-5.9</td>
<td>1.2</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>5.5</td>
<td>40.9</td>
<td>9.4</td>
<td>1.5</td>
<td>-9.1</td>
<td>-0.6</td>
<td>1</td>
</tr>
</tbody>
</table>
Sweat Rate Data

The UNIVARIATE Procedure

Variable: d1

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student's t</td>
<td>t</td>
<td>1.68733</td>
<td>Pr &gt;</td>
</tr>
<tr>
<td>Sign</td>
<td>M</td>
<td>2</td>
<td>Pr &gt;=</td>
</tr>
<tr>
<td>Signed Rank</td>
<td>S</td>
<td>38.5</td>
<td>Pr &gt;=</td>
</tr>
</tbody>
</table>

Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>W</td>
<td>0.975781</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D</td>
<td>0.159403</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq</td>
<td>0.051521</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq</td>
<td>0.275722</td>
</tr>
</tbody>
</table>

Sweat Rate Data

The UNIVARIATE Procedure

Variable: d2

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student's t</td>
<td>t</td>
<td>-1.45542</td>
<td>Pr &gt;</td>
</tr>
<tr>
<td>Sign</td>
<td>M</td>
<td>-2</td>
<td>Pr &gt;=</td>
</tr>
<tr>
<td>Signed Rank</td>
<td>S</td>
<td>-32.5</td>
<td>Pr &gt;=</td>
</tr>
</tbody>
</table>

Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>W</td>
<td>0.985837</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D</td>
<td>0.109667</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq</td>
<td>0.026283</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq</td>
<td>0.165654</td>
</tr>
</tbody>
</table>
Sweat Rate Data

The UNIVARIATE Procedure

Variable: d3

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student's t</td>
<td>-0.08218</td>
<td>0.9354</td>
</tr>
<tr>
<td>Sign</td>
<td>-1</td>
<td>0.8238</td>
</tr>
<tr>
<td>Signed Rank</td>
<td>-6.5</td>
<td>0.8194</td>
</tr>
</tbody>
</table>

Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>0.963849</td>
<td>0.6233</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>0.129105</td>
<td>0.1500</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>0.0402</td>
<td>0.2500</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>0.267019</td>
<td>0.2500</td>
</tr>
</tbody>
</table>

/* Use PROC IML to compute a chi-square Q-Q plot for assessing multivariate normality, and a related goodness-of-fit test */

data set3; set set1;
  keep d1-d3;
  run;

proc iml;
  start mvn;
  use set3; /* Enter the data */
  read all into X;
  N=NCOL(X); /* Number of observations is N */
  P=NCOL(X); /* Number of traits is P */
  SUM=X[+,]; /* Total for each trait */
  A=X'*X-SUM'/N; /* Corrected crossproducts matrix */
  S=A/(N-1); /* Sample covariance matrix */
  XBAR=SUM/N; /* Sample mean vector */
  SCALE=INV(SQRT(DIAG(A)));
  R=SCALE*A*SCALE; /* Sample correlation matrix */
  /* Compute plotting positions for a chi-square probability plot */
  SUM=W[+,]; /* Total for each variable */
  A=W'*W-SUM'/N; /* Corrected crossproducts matrix */
  S=A/(N-1); /* Sample covariance matrix */
  XBAR=SUM/N; /* Sample mean vector */
  SCALE=INV(SQRT(DIAG(A)));
  R=SCALE*A*SCALE; /* Sample correlation matrix */
\[ E = X - (J(N,1) \times X\text{BAR}); \]
\[ D = \text{VECDIAG}(E \times \text{INV}(S) \times E'); \quad /* \text{Squared Mah. distances} */ \]
\[ \text{RD} = \text{RANK}(D); \quad /* \text{Compute ranks} */ \]
\[ \text{RD} = (\text{RD} - .5)/N; \]
\[ \text{PD2} = P/2; \]
\[ Q = 2 \times \text{GAMINV}(\text{RD}, \text{PD2}); \quad /* \text{Plotting positions} */ \]
\[ \text{DQ} = D || Q; \]
\[ \text{CREATE CHISQ FROM DQ}; \quad /* \text{Open a file to store results} */ \]
\[ \text{APPEND FROM DQ}; \]
\[ /* \text{Compute test statistic} */ \]
\[ \text{rpn} = t(D) \times Q - (\text{sum}(D) \times \text{sum}(Q))/N; \]
\[ \text{rpn} = \text{rpn}/\sqrt{((\text{ssq}(D) - (\text{sum}(D)^2)/N) \times (\text{ssq}(Q) - (\text{sum}(Q)^2)/N))}; \]

/* Simulate a p-value for the correlation test */

\[ \text{ns} = 10000; \]
\[ \text{pvalue} = 0; \]
\[ \text{do } i = 1 \text{ to } \text{ns}; \]
\[ \text{do } j1 = 1 \text{ to } n; \]
\[ \text{do } j2 = 1 \text{ to } p; \]
\[ x[j1,j2] = \text{rannor}(-100); \]
\[ \text{end}; \]
\[ \text{SUMX} = X[+, \ ]; \]
\[ A = t(X) \times X - t(\text{SUMX}) \times \text{SUMX}/N; \]
\[ S = A/(N-1); \]
\[ X\text{BAR} = \text{SUMX}/N; \]

\[ \text{run mvn; } \]

/* Display the chi-square probability plot */

\[ \text{title h=2 "Chi-Square Plot"}; \]
\[ \text{proc sgplot data=chisq; } \]
\[ \text{scatter x=col2 y=col1 / } \]
\[ \text{markerattrs=(size=12 symbol=CircleFilled color=black); } \]
\[ \text{yaxis label="Ordered Distances" } \]
\[ \text{labelattrs=(size=17) valueattrs=(size=15); } \]
\[ \text{xaxis label="Chi-square Quantiles" } \]
\[ \text{labelattrs=(size=17) valueattrs=(size=15); } \]

\[ \text{print,,,,,"Correlation Test of Normality"}; \]
\[ \text{print } N \text{ } P \text{ } \text{rpn} \text{ } \text{pvalue}; \]
\[ \text{finish}; \]
The GLM Procedure

Dependent Variable: d1

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1</td>
<td>8.19200000</td>
<td>8.19200000</td>
<td>2.85</td>
<td>0.1080</td>
</tr>
<tr>
<td>Error</td>
<td>19</td>
<td>54.70800000</td>
<td>2.87936842</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>20</td>
<td>62.90000000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Parameter Estimate Standard Error t Value Pr > |t|
| z             | 0.6400000000 | 0.37943171 | 1.69 | 0.1080 |

The GLM Procedure

Dependent Variable: d2

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1</td>
<td>423.200000</td>
<td>423.200000</td>
<td>2.12</td>
<td>0.1619</td>
</tr>
<tr>
<td>Error</td>
<td>19</td>
<td>3795.980000</td>
<td>199.788421</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>20</td>
<td>4219.180000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Parameter Estimate Standard Error t Value Pr > |t|
| z             | -4.600000000 | 3.16069454 | -1.46 | 0.1619 |

/* Compute the Hotelling T-squared test */

title "One Sample Hotelling T-squared Test";
proc glm data=set1;
  model d1 d2 d3 = z / noint solution;
  manova h=z / printh printe;
  run;
ods graphics off;
ods rtf close;
### The GLM Procedure

#### Dependent Variable: d3

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1</td>
<td>0.02450000</td>
<td>0.02450000</td>
<td>0.01</td>
<td>0.9354</td>
</tr>
<tr>
<td>Error</td>
<td>19</td>
<td>68.92550000</td>
<td>3.62765789</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>20</td>
<td>68.95000000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Parameter Estimates

| Parameter | Estimate | Standard Error | t Value | Pr > |t| |
|-----------|----------|----------------|---------|------|----|
| z         | -0.03500000 | 0.42589071 | -0.08   | 0.9354  |

### The GLM Procedure

#### Multivariate Analysis of Variance

**E = Error SSCP Matrix**

<table>
<thead>
<tr>
<th></th>
<th>d1</th>
<th>d2</th>
<th>d3</th>
</tr>
</thead>
<tbody>
<tr>
<td>d1</td>
<td>54.708</td>
<td>190.19</td>
<td>-34.372</td>
</tr>
<tr>
<td>d2</td>
<td>190.19</td>
<td>3795.98</td>
<td>-107.16</td>
</tr>
<tr>
<td>d3</td>
<td>-34.372</td>
<td>-107.16</td>
<td>68.9255</td>
</tr>
</tbody>
</table>

**H = Type III SSCP Matrix for z**

<table>
<thead>
<tr>
<th></th>
<th>d1</th>
<th>d2</th>
<th>d3</th>
</tr>
</thead>
<tbody>
<tr>
<td>d1</td>
<td>8.192</td>
<td>-58.88</td>
<td>-0.448</td>
</tr>
<tr>
<td>d2</td>
<td>-58.88</td>
<td>423.2</td>
<td>3.22</td>
</tr>
<tr>
<td>d3</td>
<td>-0.448</td>
<td>3.22</td>
<td>0.0245</td>
</tr>
</tbody>
</table>

### MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall z Effect

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks' Lambda</td>
<td>0.66112774</td>
<td>2.90</td>
<td>3</td>
<td>17</td>
<td>0.0649</td>
</tr>
<tr>
<td>Pillai's Trace</td>
<td>0.13887226</td>
<td>2.90</td>
<td>3</td>
<td>17</td>
<td>0.0649</td>
</tr>
<tr>
<td>Hotelling-Lawley Trace</td>
<td>0.51256699</td>
<td>2.90</td>
<td>3</td>
<td>17</td>
<td>0.0649</td>
</tr>
<tr>
<td>Roy's Greatest Root</td>
<td>0.51256699</td>
<td>2.90</td>
<td>3</td>
<td>17</td>
<td>0.0649</td>
</tr>
</tbody>
</table>

---

# R code for computing a one-sample Hotelling T-squared test.

# This code is posted in the R code folder on the course web page as 

# sweat.R

# Data are read into a data frame from the file

# sweat.dat

# that is posted in the data folder on the course web page: 

sweat <- read.table("c:/stat501/data/sweat.dat", header=F, col.names=c("subject", "x1", "x2", "x3"))

sweat
# Compute sample mean vector and sample covariance matrix

```r
xbar <- sapply(sweat[, 2:4], mean)
xbar
```

```
x1  x2  x3
4.640 45.400 9.965
```

```r
xvar <- var(sweat[, 2:4])
xvar
```

```
x1  x2  x3
x1  2.879368 10.01  -1.809
x2  10.010000 199.7884 -5.640
x3  -1.809053 -5.640000 3.627658
```

# Use the sapply function to produce the same qqplots with a single set of code

```r
layout(matrix(1:4, nc=2))
sapply(colnames(sweat[, 2:4]), function(x) {
  qqnorm(sweat[[x]], main=x)
  qqline(sweat[[x]])
})
```

# Compute values of the Shapiro-Wilk statistic

```r
sapply(colnames(sweat[, 2:4]), function(x) {
  shapiro.test(sweat[[x]])
})
```

```
x1
statistic 0.9757811
p.value 0.8689236

x2
statistic 0.9858372
p.value 0.9861875

x3
statistic 0.9638493
p.value 0.6232605
```
# Produce a chi-square probability plot  
# to assess joint normality

layout(matrix(1:1, nc=1))  
x <- sweat[,2:4]  
d <- apply(x, 1, function(x) t(x-xbar)%*%solve(xvar)%*%(x-xbar))  
plot(qc<-qchisq((1:nrow(x)-0.5)/nrow(x), df=ncol(x)),  
orderd <- sort(d),  
xlab = "Chi-square Quantiles",  
ylab = "Ordered Distances",  
xlim = range(qc)*c(1, 1.1))  
abline(a=0, b=1)

# Compute the correlation test

r<-cor(orderd,qc)  
r  
[1] 0.975713

# Simulate a p-value for the correlation test

ns<-10000  
pvalue<-0  
nt<-nrow(x)  
nc<-ncol(x)  
nn<-nt*nc

for(j in 1:ns) {  
x <- as.matrix(rnorm(nt,0,1))  
for(i in 2:nc){x<-cbind(x,rnorm(nt,0,1))}  
d<-NULL  
for(i in 1:nt){  
d<-cbind(d,t(x[i,]-apply(x,2,mean))%*%solve(var(x))%*%  
(x[i,]-apply(x,2,mean)))}  
d<-sort(d)  
q<-NULL  
for(i in 1:nt){  
q<-cbind(q,qchisq(i/(nt+1),nc))}
Invariance property of Hotelling's $T^2$

- The $T^2$ statistic is invariant to changes in units of measurements of the form
  \[ Y_{p \times 1} = C_{p \times p} X_{p \times 1} + d_{p \times 1}, \]
  with $C$ non-singular. An example of such a transformation is the conversion of temperature measurements from Fahrenheit to Celsius.

- Note that given observations $x_1, \ldots, x_n$, we find that
  \[ \bar{y} = C \bar{x} + d, \]
  and $S_y = CSC'$. 

- Similarly, $E(Y) = C \mu + d$ and the hypothesized value is $\mu_{Y,0} = C \mu_0 + d$. 

Invariance property of Hotelling's $T^2$ (cont'd)

- We now show that the $T^2_y = T^2_x$.

\[
T^2_y = n(\bar{y} - \mu_{Y,0})'(S_y^{-1}(\bar{y} - \mu_{Y,0}))
= n(C(\bar{x} - \mu_0)'(CSC')^{-1}(C(\bar{x} - \mu_0)))
= n(\bar{x} - \mu_0)'C'(C')^{-1}S^{-1}C^{-1}C(\bar{x} - \mu_0)
= n(\bar{x} - \mu_0)'S^{-1}(\bar{x} - \mu_0).
\]

- The Hotelling $T^2$ test is the most powerful test in the class of tests that are invariate to full rank linear transformations.
Likelihood Ratio Test and Hotelling’s $T^2$

- Compare the maximum value of the multivariate normal likelihood function under no restrictions against the restricted maximized value with the mean vector held at $\mu_0$. The hypothesized value $\mu_0$ will be plausible if it produces a likelihood value almost as large as the unrestricted maximum.

- To test $H_0 : \mu = \mu_0$ against $H_a : \mu \neq \mu_0$ we construct the ratio:

$$\text{Likelihood ratio} = \Lambda = \frac{\max_{\{\Sigma\}} L(\mu_0, \Sigma)}{\max_{\{\mu, \Sigma\}} L(\mu, \Sigma)} = \left( \frac{\mid \Sigma \mid}{\mid \Sigma_0 \mid} \right)^{n/2},$$

where the numerator in the ratio is the likelihood at the MLE of $\Sigma$ given that $\mu = \mu_0$ and the denominator is the likelihood for the unrestricted MLEs for $\mu$, and $\Sigma$.

Derivation of Likelihood Ratio Test

$$|\Sigma_0|^{-n/2} \exp \left\{ -\frac{1}{2} \sum_i (x_i - \mu_0)' \Sigma_0^{-1} (x_i - \mu_0) \right\}$$

$$= |\Sigma_0|^{-n/2} \exp \left\{ -\frac{1}{2} \sum_i \text{tr} \left( (x_i - \mu_0)' \Sigma_0^{-1} (x_i - \mu_0) \right) \right\}$$

$$= |\Sigma_0|^{-n/2} \exp \left\{ -\frac{1}{2} \text{tr} \left( \Sigma_0^{-1} \sum_i (x_i - \mu_0)(x_i - \mu_0)' \right) \right\}$$

$$= |\Sigma_0|^{-n/2} \exp \left\{ -\frac{1}{2} \text{tr} \left( \Sigma_0^{-1} \Sigma_0 \right) \right\}$$

$$= |\Sigma_0|^{-n/2} \exp \left\{ -\frac{n p}{2} \right\}.$$
Derivation of Likelihood Ratio Test

\[ \Lambda = \frac{|\hat{\Sigma}_0|^{-n/2} \exp\{-\frac{mp}{2}\}}{|\hat{\Sigma}|^{-n/2} \exp\{-\frac{mp}{2}\}} \]

\[ = \frac{|\hat{\Sigma}_0|^{-n/2}}{|\hat{\Sigma}|^{-n/2}} \]

\[ = \frac{|\hat{\Sigma}|^{n/2}}{|\hat{\Sigma}_0|^{n/2}} \]

\[ = \left( \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} \right)^{n/2} . \]

- \( \mu_0 \) is a plausible value for \( \mu \) if \( \Lambda \) is close to one.

Relationship between \( \Lambda \) and \( T^2 \)

- It is just a matter of algebra to show that
  \[ \Lambda^{2/n} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} \]
  \[ = \left[ 1 + \frac{T^2}{n-1} \right]^{-1} , \text{ or} \]
  \[ \frac{|\hat{\Sigma}_0|}{|\hat{\Sigma}|} = 1 + \frac{T^2}{n-1} . \]

- For large \( T^2 \), the likelihood ratio is small and both lead to rejection of \( H_0 \).

Union-Intersection Derivation of \( T^2 \)

- Consider a reduction from \( p \)-dimensional observation vectors to univariate observations
  \[ Y_j = a'X_j = a_1X_1 + a_2X_2 + \cdots + a_pX_p \sim NID(a'\mu_0, a'\Sigma a) \]
  where \( a' = (a_1, a_2, \ldots, a_p) \)

- The null hypothesis \( H_0 : \mu = \mu_0 \) is true if and only if all null hypotheses of the form \( H_{(0,a)} : a'\mu = a'\mu_0 \) are true.

- Test \( H_{(0,a)} : a'\mu = a'\mu_0 \) versus \( H_{(A,a)} : a'\mu \neq a'\mu_0 \) with
  \[ t^2(a) = \left[ \frac{\bar{Y} - a'\mu_0}{s\sqrt{\frac{n}{n-p}}} \right]^2 = \left[ \frac{a'\bar{X} - a'\mu_0}{\sqrt{n}a'Sa} \right]^2 \]

323

324

325

326
Union-Intersection Derivation of $T^2$

- If you cannot reject the null hypothesis for the $a$ that maximizes $t^2_{(a)}$, you cannot reject any of the the univariate null hypotheses and you cannot reject the multivariate null hypothesis $H_0: \mu = \mu_0$.

- From a problem you solved on an assignment, a vector that maximizes $t^2_{(a)}$ is $a = S^{-1}(\bar{X} - \mu)$

- Consequently, The maximum squared t-test is $T^2 = n(\bar{X} - \mu_0)'S^{-1}(\bar{X} - \mu_0)$

Confidence Regions

- Confidence regions are multivariate extensions of univariate confidence intervals.

- Recall the definition of a $100(1-\alpha)\%$ CI for a parameter $\theta$: for $X \sim f(x|\theta), \theta \in \Theta$, the interval $(t_1(X), t_2(X))$ is a $100(1-\alpha)\%$ CI for $\theta$ if
  $$\Pr[t_1(X) \leq \theta \leq t_2(X)] = 1 - \alpha.$$ 

- If $\theta$ represents a univariate mean $\mu$, a $100(1-\alpha)\%$ CI for $\mu$ is given by $[\bar{X} - t_{n-1,\alpha/2} \sqrt{s^2/n}, \bar{X} + t_{n-1,\alpha/2} \sqrt{s^2/n}]$.

- Similarly, for $\theta_{p \times 1}$, the region $R(X)$ is a $100(1-\alpha)\%$ confidence region for $\theta$ if $\Pr[R(X) \text{ will cover the true } \theta] = 1 - \alpha.$

- For the mean vector $\mu_{p \times 1}$, we know that before the sample is selected,
  $$\Pr \left[ n(\bar{X} - \mu)'S^{-1}(\bar{X} - \mu) \leq \frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha) \right] = 1 - \alpha,$$
  meaning that $\bar{X}$ is within statistical distance $[(n-1)pF_{p,n-p}(\alpha)/(n-p)]^{1/2}$ from $\mu$ with probability $1 - \alpha$.

- Once a sample is obtained and $\bar{x}, S$ are computed, the set of values
  $$n(\bar{x} - \mu)'S^{-1}(\bar{x} - \mu) \leq \frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha)$$
  defines an ellipsoidal region $R(X)$ that is likely to cover $\mu$.

Confidence Regions (cont’d)

- To decide whether a hypothesized value $\mu_0$ is contained in the confidence region, we evaluate
  $$n(\bar{x} - \mu_0)'S^{-1}(\bar{x} - \mu_0)$$
  and compare it to the scaled F value above. If the squared distance from $\bar{x}$ to $\mu_0$ is larger than $[(n-1)pF_{p,n-p}(\alpha)/(n-p)]$, $\mu_0$ is not in the confidence region.

- This is exactly equivalent to testing $H_0: \mu = \mu_0$ versus $H_1: \mu \neq \mu_0$ using Hotelling’s $T^2$ statistic.

- Thus, the $100(1-\alpha)\%$ confidence region is composed of all values $\mu_0$ for which the $T^2$ test would NOT reject $H_0: \mu = \mu_0$ versus $H_1: \mu \neq \mu_0$ at level $\alpha$. 
Confidence Regions (cont’d)

- What can we say about the shape of the confidence region? It is a p-dimensional ellipsoid centered at the sample mean vector \( \bar{X} \).

- Recall that if \((\lambda_i, e_i)\) are an eigenvalue-eigenvector pair of \( S \), then letting
  \[
  (n - 1)pF_{p,n-p}(\alpha)/(n - p) = c^2,
  \]
  the \(i\)th axis of the confidence ellipse has half length
  \[
  c \sqrt{\frac{\lambda_i}{n}} = \sqrt{(n - 1)pF_{p,n-p}(\alpha)/(n - p)} \sqrt{\frac{\lambda_i}{n}}
  \]
  along the \(e_i\) direction.

Example: Microwave Ovens

- Recall the microwave oven radiation data in Tables 4.1 and 4.5, where two radiation measurements, \(x_1\) and \(x_2\), were obtained from \(n = 42\) ovens. Here, the \(x_j\) denotes the transformed (by Box-Cox) radiation measurements, using a power \(\lambda = 0.25\).

  - Sample statistics for those data are:
    \[
    \bar{x} = \begin{bmatrix} 0.564 \\ 0.603 \end{bmatrix}, \quad S = \begin{bmatrix} 0.014 & 0.012 \\ 0.012 & 0.015 \end{bmatrix}, \quad S^{-1} = \begin{bmatrix} 203.02 & -163.39 \\ -163.39 & 200.23 \end{bmatrix}.
    \]

  - Eigenvalue and eigenvector pairs for \( S \) are
    \[
    \lambda_1 = 0.026, \quad e_1' = [0.704, \ 0.710] \\
    \lambda_2 = 0.002, \quad e_2' = [-0.71, \ 0.704]
    \]

Thus, beginning from the center of the ellipse at \( \bar{X} \), the axes of the confidence ellipse are

\[
\pm \sqrt{\frac{(n - 1)pF_{p,n-p}(\alpha)}{n}} \sqrt{\frac{\lambda_i}{n}} \text{ along the } e_i \text{ direction.}
\]

- Larger differences in the sample variances across the \(p\) measurements (due to ‘real’ causes or to differences in the scale of the measurements), will create larger ratios of eigenvalues (correlations are also involved).

Example: Microwave Ovens (cont’d)

- The 95% CR for \(\mu\) is given by all values \(\mu_1, \mu_2\) that satisfy:
  \[
  42[0.564 - \mu_1 \quad 0.603 - \mu_2]' \begin{bmatrix} 203.02 & -163.39 \ -163.39 & 200.23 \end{bmatrix} \begin{bmatrix} 0.564 - \mu_1 \\ 0.603 - \mu_2 \end{bmatrix} \leq 6.62,
  \]

  where
  \[
  \frac{2(41)}{40} F_{2,40}(0.05) = \frac{2(41)}{40} 3.23 = 6.62.
  \]

- Is \(\mu_0 = [0.562 \quad 0.589]'\) a plausible value for \(\mu\)? To check, plug \(\mu_0\) into the expression above and see if it satisfies the inequality. In this case, we get 1.30 which is less than 6.62, and conclude that \(\mu_0\) is plausible at the 95% level.
Example: Microwave Ovens (cont’d)

The joint confidence ellipsoid is centered at \( \bar{x} = [0.564\ 0.603]' \) and the half lengths of the two axes are \( \sqrt{0.026} \sqrt{\frac{2(41)}{42(40)}} = 0.018, \quad \sqrt{0.002} \sqrt{\frac{2(41)}{42(40)}} = 0.064. \)

The axis are in the direction of the two eigenvectors when \( \bar{x} \) is taken as the origin.

The ratio \( \frac{\sqrt{0.026}}{\sqrt{0.002}} = 3.6 \) indicates that the major axis is 3.6 times longer than the minor axis.

Simultaneous Confidence Statements

Often we are interested in drawing inference about each \( \mu_j \).

One possibility is to construct ordinary confidence intervals

\[ \bar{x}_j \pm t_{n-1}(\frac{\alpha}{2}) \sqrt{\frac{s_{jj}}{n}}, \]

for each \( \mu_j \). One problem is that the combined set of individual intervals result in a simultaneous confidence level that is less than the nominal \( 1 - \alpha \).

There are various ways of constructing a collection of individual confidence intervals so that the joint confidence level for the family of parameters remains at \( 1 - \alpha \).

Intuitively, CI’s that protect against erosion of the confidence level will be wider than the individual \((1 - \alpha) \times 100\%\) CI’s.

Simultaneous Confidence Statements (cont’d)

Suppose that we have \( p \) variables. The population mean of the first variable \( \mu_1 \) can be written as

\[ a'_1 \mu = [1 \ 0 \ ... \ 0] \mu, \]

and in general, \( \mu_j = a'_j \mu \) where \( a'_j \) is the \( p \times 1 \) row vector with a one in the \( j \)th position and zeros in all other positions.

Given a sample \( x_1, x_2, ..., x_n \) of \( p \)-dimensional vectors, an estimator of \( \mu_j \) is \( a'_j \bar{x} \), with an estimated variance of \( a'_j Sa_j/n. \)

Then, an ordinary \((1 - \alpha) \times 100\%\) CI for \( \mu_j \) can be written as

\[ a'_j \bar{x} \pm t_{n-1}(\alpha/2) \sqrt{\frac{a'_j S a_j}{n}}. \]
Simultaneous Confidence Statements (cont'd)

- An alternative way to interpret the ordinary $(1 - \alpha) \times 100\%$ confidence interval is as follows: the CI is the set of values of $a'\mu$ for which

$$|t| = \left| \frac{\sqrt{n}(a_j'\bar{x} - a_j'\mu)}{\sqrt{a_j'Sa_j}} \right| \leq t_{n-1}(\alpha/2),$$

or, equivalently

$$t^2 = \frac{n(a_j'\bar{x} - a_j'\mu)^2}{a_j'Sa_j} \leq \frac{n(a_1'(\bar{x} - \mu))^2}{a_1'Sa_1} \leq t_{n-1}^2(\alpha/2).$$

Simultaneous Confidence Statements (cont'd)

- Intuitively, if we wish to construct a set of tests for many different vectors $a$ and have confidence level $1 - \alpha$ that all intervals will cover the true $a'\mu$, we will need a larger critical value on the right-hand side of the inequality.

- What is the maximum value that the statistic $t^2$ can reach for some vector $a$?

$$\max_a t^2 = \max_a \frac{n(a'(\bar{x} - \mu))^2}{a'Sa} = n(\bar{x} - \mu)'S^{-1}(\bar{x} - \mu) = T^2,$$

using the maximization lemma (2.50) on page 80 of your textbook (you checked this on an assignment).

- The maximum $T^2$ is achieved when $a$ is proportional to $S^{-1}(\bar{x} - \mu)$.

Simultaneous Confidence Statements (cont'd)

- Let $X_1, \ldots, X_n$ be a sample from $N_p(\mu, \Sigma)$. Then simultaneously for all $a$, the intervals given by

$$a'\bar{X} \pm c\sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(\alpha)} \frac{a'Sa}{n}$$

will cover $a'\mu$ with probability of at least $1 - \alpha$.

**Proof:** recall that

$$T^2 = n(\bar{x} - \mu)'S^{-1}(\bar{x} - \mu) \leq c^2 \iff \frac{n(a'(\bar{x} - \mu))^2}{a'Sa} \leq c^2$$

for every $a$.
Simultaneous Confidence Statements (cont’d)

- The intervals we just defined are called $T^2$ because their length is determined by the sampling distribution of $T^2$.

- For $a$ the vector with zeros everywhere and 1 in the $j$th position, the $T^2$ interval is
  \[
  \bar{x}_j - \sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(\alpha)} \frac{s_{jj}}{n} \leq \mu_j \leq \bar{x}_j + \sqrt{\frac{p(n-1)}{(n-p)} F_{p,n-p}(\alpha)} \frac{s_{jj}}{n}.
  \]

- Note that for $a$ the vector with zeros everywhere except 1 in the $j$th position and -1 in the $k$th position, the interval would correspond to $\mu_j - \mu_k$. In this case,
  \[
  a' \bar{x} = \bar{x}_j - \bar{x}_k, \quad \text{and} \quad a'Sa = s_{jj} - 2s_{jk} + s_{kk}.
  \]

Example: Microwave Ovens (cont’d)

- Before we had obtained a simultaneous 95% confidence ellipsoid for $\mu_1$ and $\mu_2$, the means of the fourth root of radiation with door closed and door open.

- We now compute 95% $T^2$ intervals for the two means. First note that
  \[
  \sqrt{\frac{p(n-1)}{n(n-p)} F_{p,n-p}(0.05)} = \sqrt{\frac{2(41)}{42(40)}} \cdot 3.23 = 0.397.
  \]
  is common to both intervals.

Example: Microwave Ovens (cont’d)

- For $\mu_1, \mu_2$:
  \[
  \bar{x}_1 \pm 0.397 \sqrt{s_{11}} \Rightarrow 0.564 \pm (0.397 \times 0.12) \Rightarrow 0.564 \pm 0.0476
  \]
  \[
  \bar{x}_2 \pm 0.397 \sqrt{s_{22}} \Rightarrow 0.603 \pm (0.397 \times 0.121) \Rightarrow 0.603 \pm 0.048.
  \]

- For the difference between doors closed and open:
  \[
  \bar{x}_1 - \bar{x}_2 \pm 0.397 \sqrt{s_{11} - 2s_{12} + s_{22}} \Rightarrow -0.039 \pm (0.397 \times 0.0748) \Rightarrow [-0.069, -0.009],
  \]
  suggesting that closing the door significantly reduces the (fourth root) radiation emitted by the ovens.

- The $T^2$ intervals are shadows or projections of the confidence ellipse onto the component axes.

Example: Microwave Ovens (cont’d)

The $T^2$ intervals are shadows or projections of the confidence ellipse onto the component axes.
Comparison of simultaneous and ordinary $t$ intervals

- The ordinary one-at-a-time $t$ intervals each have coverage probability $1 - \alpha$, but the joint coverage probability of $p$ intervals is not known.

- In the special case where the covariance matrix $\Sigma$ is diagonal, the joint coverage probability of $p$ ordinary $t$ intervals is $(1 - \alpha)^p$.

- Clearly, to guarantee $1 - \alpha$ joint coverage probability, the $t$ intervals need to be made wider.

- How much wider depends on $p$, $n$ and $\alpha$.

Comparison of confidence intervals (cont’d)

- The multipliers of $(s_{jj}/n)^{1/2}$ in the simultaneous intervals and in the $t$ intervals are, respectively:
  \[
  \left(\frac{p(n - 1)}{(n - p)}\right)^{1/2} F_{p,n-p}(\alpha), \quad \text{and} \quad t_{n-1}(\alpha/2).
  \]

- For example, for $\alpha = 0.05$, $n = 15$ and $p = 4$, the simultaneous intervals are
  \[
  \frac{(4.14 - 2.145)}{2.145} \times 100\% = 93\% \text{ wider.}
  \]

The Bonferroni method for multiple comparisons

- The Bonferroni method is useful when we wish to make a small number $m$ of comparisons for linear combinations $a'_1 \mu, ..., a'_m \mu$.

- Let $C_i$ denote a confidence statement about $a'_i \mu$ such that $\Pr(C_i \text{ true}) = 1 - \alpha_i$. Then
  \[
  \Pr(\text{all } C_i \text{ true}) = 1 - \Pr(\text{ at least one } C_i \text{ false}) \\
  \geq 1 - \sum_i \Pr(C_i \text{ false}) \\
  = 1 - \sum_i (1 - \Pr(C_i \text{ true}) \\
  = 1 - (\alpha_1 + \alpha_2 + ... + \alpha_m).
  \]
The Bonferroni method for multiple comparisons

• Consider, for example, $m$ individual $t$ intervals for $\mu_1, \ldots, \mu_m$, with $\alpha_i = \alpha/m$. From the Bonferroni inequality, we have that:

$$\Pr \left[ \bar{X}_i \pm t_{n-1} \left( \frac{\alpha}{2m} \right) \sqrt{\frac{\overline{S^2}}{n}} \text{ contains } \mu_i, \text{ for all } i \right] \geq 1 - \sum_{i=1}^{m} \frac{\alpha}{m} = 1 - \alpha.$$  

• In general, to make confidence statements about $p$ means, we divide the significance level $\alpha$ by the number of intervals we want to construct $p$.

• Microwave ovens: see $T^2$ and Bonferroni intervals in next figure.

Paired Comparisons and Repeated Measures

• Hotelling’s $T$-squared tests for a single mean vector have useful applications for studies with paired comparisons or repeated measurements.

  1. Paired comparison designs in which two treatments are applied to each sample unit and $p$ variables are measured on each unit under each treatment.

  2. Repeated measures designs in which the same variable is measured at the same set of $p$ time points on each sample unit.

$T^2$ and Bonferroni confidence intervals

Paired Comparison Designs

• What is a paired comparisons design? A design where two different treatments (or a treatment and an ‘absence of treatment’ or control) are assigned to each sample unit.

  • Examples:

    – We measure volumes of sales of a certain product in a certain market before and after an advertising campaign

    – We measure blood pressure on a sample of individuals before and after receiving some drug

    – We count traffic accidents at a collection of intersections when stop sign controls or light controls were used for signalization.
Paired Comparison Designs

- One advantage of this type of design is that the differences in the outcome measures under one treatment or the other reflect only the effect of treatment since everything else in the units receiving the treatments is absolutely identical.

- It is not quite so simple: in drug trials there may be carry-over effects to worry about and in 'before/after' experiments other conditions may also change.

Paired Comparisons when \( p = 1 \)

- In the univariate case, let \( X_{1j} \) and \( X_{2j} \) denote the responses of unit \( j \) under treatments 1 and 2, respectively.

- Differences \( D_j = X_{1j} - X_{2j}, \ j = 1, \ldots, n \) reflect the effect of treatment. If \( D \sim \mathcal{N}(\delta, \sigma_d^2) \),

\[
t = \frac{\bar{D} - \delta}{s_d/\sqrt{n}} \sim t_{n-1},
\]

\[
\bar{D} = \frac{1}{n} \sum_j D_j, \quad s_d^2 = \frac{1}{n-1} \sum_j (D_j - \bar{D})^2.
\]

- A 100(1 - \( \alpha \))% confidence interval for \( \delta = E(X_{1j} - X_{2j}) \) is given by

\[
d - t_{n-1}(\alpha/2)s_d/\sqrt{n} \leq \delta \leq d + t_{n-1}(\alpha/2)s_d/\sqrt{n}.
\]

Paired Comparisons when \( p > 1 \)

- When \( p \) variables are measured on each sample unit, we compute vectors of differences ([post treatment]-[pre treatment]):

\[
D_j = \begin{bmatrix} D_{j1} \\ D_{j2} \\ \vdots \\ D_{jp} \end{bmatrix} = \begin{bmatrix} X_{1j1} \\ X_{1j2} \\ \vdots \\ X_{1jp} \end{bmatrix} - \begin{bmatrix} X_{2j1} \\ X_{2j2} \\ \vdots \\ X_{2jp} \end{bmatrix}
\]

for \( j = 1, \ldots, n \) sample units.

- For \( D_j \sim \mathcal{N}(\bar{D}, \Sigma_d) \) we have

\[
T^2 = n(\bar{D} - \delta)'S_D^{-1}(\bar{D} - \delta) \sim \frac{p(n-1)}{n-p} F_{p,n-p}.
\]

Paired Comparisons when \( p > 1 \)

- Where

\[
\bar{D} = \frac{1}{n} \sum_j D_j, \quad S_D = \frac{1}{n-1} \sum_j (D_j - \bar{D})(D_j - \bar{D})'.
\]

- In particular, a test of \( H_0 : \delta = 0 \) versus \( H_1 : \delta \neq 0 \) rejects \( H_0 \) at level \( \alpha \) if

\[
T^2 = n\bar{D}'S_D^{-1}\bar{D} > \frac{p(n-1)}{n-p} F_{p,n-p}(\alpha).
\]

If we fail to reject \( H_0 \), we conclude that there are no substantial treatment effects on any of the \( p \) variables.
Confidence region for $\delta$

- As before, a $100(1 - \alpha)\%$ CR for $\delta$ consists of all $\delta$ such that
  \[
  (\bar{D} - \delta)^{\prime} S_D^{-1} (\bar{D} - \delta) \leq \frac{p(n - 1)}{n(n - p)} F_{p, n - p}(\alpha).
  \]

- Simultaneous $T^2$ intervals for the $p$ individual mean differences are given by
  \[
  \bar{D}_i \pm \sqrt{\frac{p(n - 1)}{(n - p)} F_{p, n - p}(\alpha)} \sqrt{\frac{s_d^2}{n}}.
  \]

- When $n - p$ is large,
  \[
  \frac{p(n - 1)}{(n - p)} F_{p, n - p}(\alpha) \approx \chi^2_p(\alpha).
  \]

Example 6.1: Effluent Study

- Wastewater treatment plants are required to monitor the quality of the treated water that goes into rivers and streams.

- To monitor the reliability of a self-monitoring program, 11 samples of water were divided in half. One half was sent to a state lab and another one to a private lab for analysis.

- Two variables were measured: biochemical oxygen demand (BOD) and suspended solids (SS).

- Question of interest: Is there a difference between labs in mean values for either BOD or SS?

Example 6.1: Effluent Study (cont’d)

- Sample statistics are
  \[
  \bar{d} = \begin{bmatrix} -9.36 \\ 13.27 \end{bmatrix}, \quad S_d = \begin{bmatrix} 199.26 & 88.38 \\ 88.38 & 418.61 \end{bmatrix}.
  \]

- To test $H_0: \delta = 0$ versus $H_1: \delta \neq 0$, we compute $T^2$:
  \[
  T^2 = 11[-9.36 \quad 13.27] \begin{bmatrix} 0.0055 & -0.0012 \\ -0.0012 & 0.0026 \end{bmatrix} \begin{bmatrix} 9.36 \\ 13.27 \end{bmatrix} = 13.6.
  \]

- For $\alpha = 0.05$, $[p(n - 1)/(n - p)] F_{2,9}(0.05) = 9.47$. Since $T^2 > 9.47$, we reject $H_0$. 

<table>
<thead>
<tr>
<th>Sample</th>
<th>Commercial $x_{1j}$ (BOD)</th>
<th>Commercial $x_{1j}$ (SS)</th>
<th>State $x_{2j}$ (BOD)</th>
<th>State $x_{2j}$ (SS)</th>
<th>$d_{j1}$</th>
<th>$d_{j2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>27</td>
<td>25</td>
<td>15</td>
<td>-19</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>23</td>
<td>28</td>
<td>13</td>
<td>-22</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>64</td>
<td>36</td>
<td>22</td>
<td>-18</td>
<td>42</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>44</td>
<td>35</td>
<td>29</td>
<td>-27</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>30</td>
<td>15</td>
<td>31</td>
<td>-4</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>34</td>
<td>75</td>
<td>44</td>
<td>64</td>
<td>-10</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>26</td>
<td>42</td>
<td>30</td>
<td>-14</td>
<td>-4</td>
</tr>
<tr>
<td>8</td>
<td>71</td>
<td>124</td>
<td>54</td>
<td>64</td>
<td>17</td>
<td>60</td>
</tr>
<tr>
<td>9</td>
<td>43</td>
<td>54</td>
<td>34</td>
<td>56</td>
<td>9</td>
<td>-2</td>
</tr>
<tr>
<td>10</td>
<td>33</td>
<td>30</td>
<td>29</td>
<td>20</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>20</td>
<td>14</td>
<td>39</td>
<td>21</td>
<td>-19</td>
<td>-7</td>
</tr>
</tbody>
</table>
Example 6.1: Effluent Study (cont’d)

- The two labs appear to produce significantly different mean measurements for either BOD or SS or both. Which variable has different means?

- The private lab seems to report lower BOD and higher SS than the state lab. We compute the two simultaneous $T^2$ intervals to obtain:

$$\delta_1 : (-22.46, 3.74), \quad \delta_2 : (-5.71, 32.25).$$

- Both intervals include zero, yet we rejected the null hypothesis of equal mean vectors.

---

Example 6.1: Effluent Study (cont’d)

- The $T^2$ intervals are conservative (wider than needed for simultaneous 95% coverage) when only two comparisons are being made.

- The Bonferroni intervals are less conservative, but also cover zero.

- One complication is that normality of the $D_i$’s is suspect; water sample 8 appears to be an outlier. With such a small sample, its effect is substantial.

- Consider the correlation between the measurements of BOD and SS

---

/* This program is posted as effluent.sas in the SAS code folder on the course web page. It computes a one-sample Hotelling T-squared test for the data in example 6.1 */

DATA set1;
  INPUT sample BOD1 SS1 BOD2 SS2;
  dBOD=BOD1-BOD2;
  dSS = SS1-SS2;
  Z=1;
  /* LABEL BOD1 = BOD (private lab)  
     SS1 = SS (private lab)  
     BOD2 = BOD (state lab)  
     SS2 = SS (state lab)  
     dBOD = BOD Difference (Private vs. State)  
     dSS = SS Difference (Private vs State); */
  datalines;
  1 6 27 25 15 2 6 23 28 13 3 18 64 36 22 4 8 44 35 29 5 11 30 15 31 6 34 75 44 64 7 28 26 42 30 8 71 124 54 64 9 43 54 34 56 10 33 30 29 20 11 20 14 39 21 run;
Differences in Effluent Measurements

The CORR Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Sum</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>dBOD</td>
<td>11</td>
<td>-9.36</td>
<td>14.12</td>
<td>-103</td>
<td>-27.00</td>
<td>17.00</td>
</tr>
<tr>
<td>dSS</td>
<td>11</td>
<td>13.27</td>
<td>20.46</td>
<td>146.00</td>
<td>-7.00</td>
<td>60000</td>
</tr>
</tbody>
</table>

Pearson Correlation Coefficients, N = 11

<table>
<thead>
<tr>
<th></th>
<th>dBOD</th>
<th>dSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>dBOD</td>
<td>0.31</td>
<td>0.36</td>
</tr>
<tr>
<td>dSS</td>
<td>0.36</td>
<td>0.31</td>
</tr>
</tbody>
</table>

/* Create a scatter plot for pairs of differences */

proc sgplot data=set1;
scatter x=dBOD y=dSS /
    markerattrs=(size=12 symbol=CircleFilled color=black);
yaxis label="SS (private vs state lab)"
    labelattrs=(size=17) valueattrs=(size=15);
xaxis label="BOD (private vs state lab)"
    labelattrs=(size=17) valueattrs=(size=15);
run;

proc univariate data=set1 normal;
var dBOD dSS; qqplot;
run;
### The UNIVARIATE Procedure

**Variable: dBOD**

#### Tests for Location: \( \mu_0 = 0 \)

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>( p ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student's t</td>
<td>-2.20007</td>
<td>0.0524</td>
</tr>
<tr>
<td>Sign</td>
<td>-2.5</td>
<td>0.2266</td>
</tr>
<tr>
<td>Signed Rank</td>
<td>-22.5</td>
<td>0.0449</td>
</tr>
</tbody>
</table>

#### Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>( p ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>0.920348</td>
<td>0.3215</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>0.184218</td>
<td>&gt;0.1500</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>0.072704</td>
<td>0.2382</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>0.417041</td>
<td>&gt;0.2500</td>
</tr>
</tbody>
</table>

### The UNIVARIATE Procedure

**Variable: dSS**

#### Tests for Location: \( \mu_0 = 0 \)

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>( p ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student's t</td>
<td>2.15153</td>
<td>0.0569</td>
</tr>
<tr>
<td>Sign</td>
<td>1.5</td>
<td>0.5488</td>
</tr>
<tr>
<td>Signed Rank</td>
<td>23</td>
<td>0.0400</td>
</tr>
</tbody>
</table>

#### Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>( p ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>0.81878</td>
<td>0.0166</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>0.284543</td>
<td>0.0140</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>0.157788</td>
<td>0.0165</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>0.866596</td>
<td>0.0186</td>
</tr>
</tbody>
</table>
```latex
title h=2 "Hotelling T-squared Test";
proc glm data=set1;
   model dBOD dSS = Z / noint solution;
   manova H=Z / printhe;
run;

ods graphics off;
ods rtf close;
```

---

**The GLM Procedure**

**Dependent Variable: dBOD**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1</td>
<td>964.4545</td>
<td>964.4545</td>
<td>4.84</td>
<td>0.0524</td>
</tr>
<tr>
<td>Error</td>
<td>10</td>
<td>1992.5454</td>
<td>199.2545</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>11</td>
<td>2957.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Parameter | Estimate  | Standard Error | t Value | Pr > |t| |
|-----------|-----------|----------------|---------|------|---|
| Z         | -9.3636   | 4.2561         | -2.20   | 0.0524 |

---

**The GLM Procedure**

**Dependent Variable: dSS**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1</td>
<td>1937.8182</td>
<td>1937.8182</td>
<td>4.63</td>
<td>0.0569</td>
</tr>
<tr>
<td>Error</td>
<td>10</td>
<td>4186.1818</td>
<td>418.6182</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>11</td>
<td>6124.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Parameter | Estimate  | Standard Error | t Value | Pr > |t| |
|-----------|-----------|----------------|---------|------|---|
| Z         | 13.2727   | 6.1689         | 2.15    | 0.0569 |

---

**Hotelling T-squared Test**

**The GLM Procedure**

**Multivariate Analysis of Variance**

<table>
<thead>
<tr>
<th>E = Error SSCP Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>dBOD</td>
</tr>
</tbody>
</table>

| dBOD | 1992.5454 | 883.09090909 |
| dSS  | 883.09090909 | 4186.1818182 |

<table>
<thead>
<tr>
<th>H = Type III SSCP Matrix for Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>dBOD</td>
</tr>
</tbody>
</table>

| dBOD | 964.4545 | -1367.090909 |
| dSS  | -1367.090909 | 1937.8181818 |

---
Hotelling T-squared Test

The GLM Procedure
Multivariate Analysis of Variance

Characteristic Roots and Vectors of: E Inverse * H, where
H = Type III SSCP Matrix for Z
E = Error SSCP Matrix

<table>
<thead>
<tr>
<th>Characteristic Root</th>
<th>Characteristic Vector V'EV=1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.36393121</td>
<td>0.01912416 -0.01303845</td>
</tr>
<tr>
<td>0.00000000</td>
<td>0.01370761 0.00967044</td>
</tr>
</tbody>
</table>

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Z Effect
H = Type III SSCP Matrix for Z
E = Error SSCP Matrix

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks' Lambda</td>
<td>0.42302415</td>
<td>6.14</td>
<td>2</td>
<td>9</td>
<td>0.0208</td>
</tr>
<tr>
<td>Pillai's Trace</td>
<td>0.57697585</td>
<td>6.14</td>
<td>2</td>
<td>9</td>
<td>0.0208</td>
</tr>
<tr>
<td>Hotelling-Lawley Trace</td>
<td>1.36393121</td>
<td>6.14</td>
<td>2</td>
<td>9</td>
<td>0.0208</td>
</tr>
<tr>
<td>Roy's Greatest Root</td>
<td>1.36393121</td>
<td>6.14</td>
<td>2</td>
<td>9</td>
<td>0.0208</td>
</tr>
</tbody>
</table>

# Code for computing one sample Hotelling T-squared test for a paired comparison study
# effluent.R
#
# Data are read into a data frame from the file
#
# effluent.dat

edat<-read.table("c:/stat501/data/effluent.dat", header=F, col.names=c("sample", "bod1", "ss1", "bod2", "ss2"))

edat

sample  bod1  ss1  bod2  ss2
1       1      6    27    25    15
2       2      6    23    28    13
3       3      18   64    36    22
4       4      8    44    35    29
5       5      11   30    15    31
6       6      34   75    44    64
7       7      28   26    42    30
8       8      71  124    54    64
9       9      43   54    34    56
10      10     33   30    29    20
11      11     20   14    39    21

# Compute sample mean vector and sample covariance matrix
xbar<-sapply(edat[,2:5], mean)
xbar
bod1  ss1  bod2  ss2
25.27273 46.45455 34.63636 33.18182

xvar<-var(edat[,2:5])
xvar
bod1  ss1  bod2  ss2
bod1  387.4182  489.3636  148.7091  296.0455
ss1  489.3636 1014.0727  225.3818  479.6091
bod2  148.7091  225.3818 109.2545  120.3727
ss2  296.0455  479.6091  120.3727  363.7636
# Apply the contrasts

\[
C = \begin{bmatrix}
1 & 0 & -1 & 0 \\
0 & 1 & 0 & -1
\end{bmatrix}
\]

\[
Cxbar \leftarrow C \times xbar
\]

\[
Cxvar \leftarrow C \times xvar \times t(C)
\]

\[
p \leftarrow \text{nrow}(Cxvar)
\]

\[
n \leftarrow \text{nrow}(edat)
\]

\[
nullmean \leftarrow \text{rep}(0, p)
\]

\[
d \leftarrow Cxbar - nullmean
\]

\[
t2 = \frac{n \times t(d) \times \text{solve}(Cxvar) \times d}{(n-p) \times (p \times (n-1))}
\]

\[
pval \leftarrow 1 - \text{pf}(t2mod, p, n-p)
\]

\[
t2
\]

\[
[,1]
\]

\[
[1,] 13.63931
\]

\[
pval
\]

\[
[,1]
\]

\[
[1,] 0.02082779
\]

## Repeated Measures Designs

- Suppose now that \( q \) treatments are applied to each of the \( n \) sample units. Then

\[
X_j' = [X_{j1}, X_{j2}, ..., X_{jp}]', \quad j = 1, ..., n.
\]

- Of interest are contrasts of the elements of \( \mu = E(X_j) \) such as

\[
\begin{bmatrix}
\mu_1 - \mu_2 \\
\mu_1 - \mu_3 \\
\vdots \\
\mu_1 - \mu_q
\end{bmatrix} = \begin{bmatrix}
1 & -1 & 0 & \cdots & 0 \\
1 & 0 & -1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 0 & 0 & \cdots & -1
\end{bmatrix} \mu = C\mu.
\]

- The **contrast matrix** \( C \) has \( p - 1 \) independent rows.

- If there is no effect of treatment, then \( C\mu = 0 \).

## Testing Hypotheses in Repeated Measures Designs

- To test \( H_0 : C\mu = 0 \) we again use \( T^2 \):

\[
T^2 = n(C\bar{x})'(CSC')^{-1}(C\bar{x})
\]

- The null hypothesis is rejected if

\[
T^2 \ge \frac{(n - 1)(p - 1)}{(n - p + 1)} F_{p-1, n-p+1}(\alpha).
\]

- A confidence region for contrasts \( C\mu \) is given by all \( C\mu \) such that

\[
n(C\bar{x})'(CSC')^{-1}(C\bar{x}) \le \frac{(n - 1)(p - 1)}{(n - p + 1)} F_{p-1, n-p+1}(\alpha).
\]
Example 6.2: Dog Anesthetics

A sample of 19 dogs were administered four treatments each:
(1) high CO\textsubscript{2} pressure, (2) low CO\textsubscript{2} pressure,
(3) high pressure + halothane, (4) low pressure + halothane.

Outcome variable was milliseconds between heartbeats.

Three contrasts are of interest:
(\mu_3 + \mu_4) - (\mu_1 + \mu_2) : Effect of halothane
(\mu_1 + \mu_3) - (\mu_2 + \mu_4) : Effect of CO\textsubscript{2} pressure
(\mu_1 + \mu_4) - (\mu_2 + \mu_3) : Interaction between halothane and CO\textsubscript{2}.

Example 6.2: Dog Anesthetics (cont’d)

The 3 × 4 contrast matrix \( C \) is:
\[
\begin{bmatrix}
-1 & -1 & 1 & 1 \\
1 & -1 & 1 & -1 \\
1 & -1 & -1 & 1
\end{bmatrix}.
\]

The test of \( H_0 : C\mu = 0 \) rejects the null if
\[
T^2 = n(C\bar{x})'(CSC')^{-1}(C\bar{x}) = 116 \geq \frac{(18)(3)}{(16)} f_{3,16}(\alpha) = 10.94.
\]

We clearly reject the null, so the question now is whether there is a difference between CO\textsubscript{2} pressure, between halothane levels or perhaps there is no main effect of treatment but there is still an interaction.

Example 6.2: Dog Anesthetics (cont’d)

Three simultaneous confidence intervals (one for each row of \( C \)):
\[
d_1\mu : (\bar{x}_3 + \bar{x}_4) - (\bar{x}_1 + \bar{x}_2) \pm \sqrt{\frac{18(3)}{16} F_{3,16}(0.05)} \sqrt{d_1'Sc_1}
= 209.31 \pm 73.70.
\]

For the effect of CO\textsubscript{2} pressure and the interaction, we find respectively:
-60.05 ± 54.70, -12.79 ± 65.97

Use of halothane produces longer times between heartbeats. This effect is similar at both high and low CO\textsubscript{2} pressure levels because the interaction is not significant (third contrast). Higher CO\textsubscript{2} pressure produces shorter times between heartbeats.

Example 6.2: Dog Anesthetics (cont’d)

The same null hypothesis can be represented with different contrast matrices

For example, the null hypothesis that all four treatments have the same mean is also tested with \( H_0 : C\mu = 0 \), where
\[
C = \begin{bmatrix}
-1 & 1 & 0 & 0 \\
0 & -1 & 1 & 0 \\
0 & 0 & -1 & 1
\end{bmatrix} \text{ or } C = \begin{bmatrix}
1 & 0 & 0 & -1 \\
0 & 1 & 0 & -1 \\
0 & 0 & 1 & -1
\end{bmatrix}
\]

The value of \( T^2 \) is unchanged
\[
T^2 = n(C\bar{x})'(CSC')^{-1}(C\bar{x}) = 116
\]

The value of \( T^2 \) will be the same for any set of contrasts that implies the same joint null hypothesis.
/ This program performs a one-sample Hotelling T-squared test for equality of means for repeated measures. The SAS code is posted as dogs.sas

This program analyzes the data posted as dogs.dat */
data set1;
infile "c:\stat501\data\dogs.dat";
input dog x1-x4;Z=1;
/* LABEL X1 = High CO2 without Halothane
   X2 = Low CO2 without Halothane
   X3 = High CO2 with Halothane
   X4 = Low CO2 with Halothane; */
run;

/* Create a scatter plot matrix */
proc sgscatter data=set1; matrix x1 x2 x3 x4/ diagonal=(histogram)
markerattrs=(size=10 symbol=CircleFilled color=black); run;

/* Test the null hypothesis that all treatments have the same mean time between heartbeats */
proc glm data=set1;
   model x1-x4 = z /noint nouni;
   manova H=z M=x1-x2+x3-x4, -x1-x2+x3+x4,
   -x1+x2+x3-x4 prefix=c /printe ;
run;
The GLM Procedure
Multivariate Analysis of Variance

M Matrix Describing Transformed Variables

<table>
<thead>
<tr>
<th></th>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>c2</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>c3</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>

E = Error SSCP Matrix

<table>
<thead>
<tr>
<th></th>
<th>c1</th>
<th>c2</th>
<th>c3</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>93524.947368</td>
<td>19780.315789</td>
<td>-16462.21053</td>
</tr>
<tr>
<td>c2</td>
<td>19780.315789</td>
<td>169780.10526</td>
<td>-16696.73684</td>
</tr>
<tr>
<td>c3</td>
<td>-16462.21053</td>
<td>-16696.73684</td>
<td>136033.15789</td>
</tr>
</tbody>
</table>

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Z Effect on the Variables Defined by the M Matrix Transformation

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks’ Lambda</td>
<td>0.13431200</td>
<td>34.38</td>
<td>3</td>
<td>16</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Pillai’s Trace</td>
<td>0.86568800</td>
<td>34.38</td>
<td>3</td>
<td>16</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Hotelling-Lawley Trace</td>
<td>6.44535118</td>
<td>34.38</td>
<td>3</td>
<td>16</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Roy’s Greatest Root</td>
<td>6.44535118</td>
<td>34.38</td>
<td>3</td>
<td>16</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

/* Do a mixed model analysis of variance. Restructure the data file to put each observation on a different line */

data set3; set set1;
array A(I) x1-x4;
do over A;
x = A;
treat = I;
output;
end;
keep dog treat x;
run;

proc print data=set3;
run;

395 396 397 398
proc glm data=set3;
class treat dog;
model x = dog treat / solution;
random dog;
means treat / bon tukey;
run;

The GLM Procedure

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>21</td>
<td>531413.6974</td>
<td>25305.4142</td>
<td>13.69</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>54</td>
<td>99834.5526</td>
<td>1848.7880</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>75</td>
<td>631248.2500</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>dog</td>
<td>18</td>
<td>305394.5000</td>
<td>16966.3611</td>
<td>9.18</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>treat</td>
<td>3</td>
<td>226019.1974</td>
<td>75339.7325</td>
<td>40.75</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Source | Type III Expected Mean Square
--------|---------------------------------|
dog     | Var(Error) + 4 Var(dog)        |
treat   | Var(Error) + Q(treat)          |

Alpha                  | 0.05
Error Degrees of Freedom | 54
Error Mean Square       | 1848.788
Critical Value of Studentized Range | 3.74890
Minimum Significant Difference | 36.98

Means with the same letter are not significantly different.

<table>
<thead>
<tr>
<th>Tukey Grouping</th>
<th>Mean</th>
<th>N</th>
<th>treat</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>502.89</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>479.26</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>404.63</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>368.21</td>
<td>19</td>
<td>1</td>
</tr>
</tbody>
</table>

/* Use the REPEATED option in PROC GLM to do both analyses */

proc glm data=set1;
model x1-x4 = z / noint nouni;
repeated x 4 contrast / printm printe short summary;
run;
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

x_N represents the contrast between the nth level of x and the last

M Matrix Describing Transformed Variables

<table>
<thead>
<tr>
<th></th>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
</tr>
</thead>
<tbody>
<tr>
<td>x_1</td>
<td>1.0000000</td>
<td>0.000000000</td>
<td>0.000000000</td>
<td>-1.000000000</td>
</tr>
<tr>
<td>x_2</td>
<td>0.0000000</td>
<td>1.000000000</td>
<td>0.000000000</td>
<td>-1.000000000</td>
</tr>
<tr>
<td>x_3</td>
<td>0.0000000</td>
<td>0.000000000</td>
<td>1.000000000</td>
<td>-1.000000000</td>
</tr>
</tbody>
</table>

E = Error SSCP Matrix

x_N represents the contrast between the nth level of x and the last

<table>
<thead>
<tr>
<th></th>
<th>x_1</th>
<th>x_2</th>
<th>x_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x_1</td>
<td>55936</td>
<td>37559</td>
<td>18495</td>
</tr>
<tr>
<td>x_2</td>
<td>37559</td>
<td>84802</td>
<td>29122</td>
</tr>
<tr>
<td>x_3</td>
<td>18495</td>
<td>29122</td>
<td>49158</td>
</tr>
</tbody>
</table>

Repeating Measures Analysis of Variance
Tests of Hypotheses for Within Subjects Effects

Contrast Variable: x_1

Contrast Variable: x_2

Contrast Variable: x_3

Sphericity Tests

<table>
<thead>
<tr>
<th>Variables</th>
<th>DF</th>
<th>Mauchly's Criterion</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transformed Variates</td>
<td>5</td>
<td>0.5040967</td>
<td>11.454509</td>
<td>0.0431</td>
</tr>
<tr>
<td>Orthogonal Components</td>
<td>5</td>
<td>0.8672015</td>
<td>2.381047</td>
<td>0.7943</td>
</tr>
</tbody>
</table>

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of no x*z Effect
H = Type III SSCP Matrix for x*z
E = Error SSCP Matrix

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Num DF</th>
<th>Den DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks' Lambda</td>
<td>0.13431200</td>
<td>34.38</td>
<td>3</td>
<td>16</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Pillai's Trace</td>
<td>0.86568800</td>
<td>34.38</td>
<td>3</td>
<td>16</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Hotelling-Lawley Trace</td>
<td>6.44535118</td>
<td>34.38</td>
<td>3</td>
<td>16</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Roy's Greatest Root</td>
<td>6.44535118</td>
<td>34.38</td>
<td>3</td>
<td>16</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Repeated Measures Analysis of Variance
Analysis of Variance of Contrast Variables

Greenhouse-Geisser Epsilon 0.9171
Huynh-Feldt Epsilon 1.0991
proc mixed data=set3;
class treat dog;
model x = treat / solution;
repeated / type=cs subject=dog r=1;
lsmeans treat / adjust=bon cl;
run;
### The Mixed Procedure

#### Differences of Least Squares Means

<table>
<thead>
<tr>
<th>Effect</th>
<th>treat</th>
<th>_treat</th>
<th>Alpha</th>
<th>Lower</th>
<th>Upper</th>
<th>Adj Lower</th>
<th>Adj Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>treat 1</td>
<td>2</td>
<td>2</td>
<td>0.05</td>
<td>-64.39</td>
<td>-8.45</td>
<td>-74.63</td>
<td>1.78</td>
</tr>
<tr>
<td>treat 1</td>
<td>3</td>
<td>3</td>
<td>0.05</td>
<td>-139.02</td>
<td>-83.08</td>
<td>-149.26</td>
<td>-172.89</td>
</tr>
<tr>
<td>treat 1</td>
<td>4</td>
<td>4</td>
<td>0.05</td>
<td>-162.65</td>
<td>-106.72</td>
<td>-172.89</td>
<td>-96.47</td>
</tr>
<tr>
<td>treat 2</td>
<td>3</td>
<td>3</td>
<td>0.05</td>
<td>-102.60</td>
<td>-46.66</td>
<td>-112.84</td>
<td>-36.42</td>
</tr>
<tr>
<td>treat 2</td>
<td>4</td>
<td>4</td>
<td>0.05</td>
<td>-126.23</td>
<td>-70.29</td>
<td>-136.47</td>
<td>-60.05</td>
</tr>
<tr>
<td>treat 3</td>
<td>4</td>
<td>4</td>
<td>0.05</td>
<td>-51.60</td>
<td>4.33</td>
<td>-61.84</td>
<td>14.57</td>
</tr>
</tbody>
</table>

#### Covariance Parameter Estimates

<table>
<thead>
<tr>
<th>Cov Parm</th>
<th>Subject</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN(1,1)</td>
<td>dog</td>
<td>2819.29</td>
</tr>
<tr>
<td>UN(2,1)</td>
<td>dog</td>
<td>3568.42</td>
</tr>
<tr>
<td>UN(2,2)</td>
<td>dog</td>
<td>7963.13</td>
</tr>
<tr>
<td>UN(3,1)</td>
<td>dog</td>
<td>2943.50</td>
</tr>
<tr>
<td>UN(3,2)</td>
<td>dog</td>
<td>5303.99</td>
</tr>
<tr>
<td>UN(3,3)</td>
<td>dog</td>
<td>6851.32</td>
</tr>
<tr>
<td>UN(4,1)</td>
<td>dog</td>
<td>2295.36</td>
</tr>
<tr>
<td>UN(4,2)</td>
<td>dog</td>
<td>4065.46</td>
</tr>
<tr>
<td>UN(4,3)</td>
<td>dog</td>
<td>4499.64</td>
</tr>
<tr>
<td>UN(4,4)</td>
<td>dog</td>
<td>4878.99</td>
</tr>
</tbody>
</table>

#### Code Snippet

```r
proc mixed data=set3;
  class treat dog;
  model x = treat / solution;
  repeated / type=un subject=dog r=1;
  lsmeans treat / adjust=bon cl;
run;

ods graphics off;
ods rtf close;
```
Fit Statistics
-2 Res Log Likelihood: 785.2
AIC (smaller is better): 805.2
AICC (smaller is better): 808.8
BIC (smaller is better): 814.6

Type 3 Tests of Fixed Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Num DF</th>
<th>Den DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>treat</td>
<td>3</td>
<td>18</td>
<td>38.67</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Least Squares Means

<table>
<thead>
<tr>
<th>Effect</th>
<th>treat</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>Alpha</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>treat</td>
<td>1</td>
<td>368.21</td>
<td>12.1813</td>
<td>18</td>
<td>30.23</td>
<td>&lt;.0001</td>
<td>0.05</td>
<td>342.62</td>
<td>393.80</td>
</tr>
<tr>
<td>treat</td>
<td>2</td>
<td>404.63</td>
<td>20.4722</td>
<td>18</td>
<td>19.76</td>
<td>&lt;.0001</td>
<td>0.05</td>
<td>361.62</td>
<td>447.64</td>
</tr>
<tr>
<td>treat</td>
<td>3</td>
<td>479.26</td>
<td>18.9894</td>
<td>18</td>
<td>25.24</td>
<td>&lt;.0001</td>
<td>0.05</td>
<td>439.37</td>
<td>519.16</td>
</tr>
<tr>
<td>treat</td>
<td>4</td>
<td>502.89</td>
<td>16.0246</td>
<td>18</td>
<td>31.38</td>
<td>&lt;.0001</td>
<td>0.05</td>
<td>469.23</td>
<td>536.56</td>
</tr>
</tbody>
</table>

Differences of Least Squares Means

<table>
<thead>
<tr>
<th>Effect</th>
<th>treat</th>
<th>treat</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>Adjustment</th>
<th>Adj P</th>
</tr>
</thead>
<tbody>
<tr>
<td>treat</td>
<td>1</td>
<td>2</td>
<td>-36.4211</td>
<td>13.8518</td>
<td>18</td>
<td>-2.63</td>
<td>0.0170</td>
<td>Bonferroni</td>
<td>0.1021</td>
</tr>
<tr>
<td>treat</td>
<td>1</td>
<td>3</td>
<td>-111.05</td>
<td>14.1116</td>
<td>18</td>
<td>-7.87</td>
<td>&lt;.0001</td>
<td>Bonferroni</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>treat</td>
<td>1</td>
<td>4</td>
<td>-134.68</td>
<td>12.7889</td>
<td>18</td>
<td>-10.53</td>
<td>&lt;.0001</td>
<td>Bonferroni</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>treat</td>
<td>2</td>
<td>3</td>
<td>-74.6316</td>
<td>14.8793</td>
<td>18</td>
<td>-5.02</td>
<td>&lt;.0001</td>
<td>Bonferroni</td>
<td>0.0005</td>
</tr>
<tr>
<td>treat</td>
<td>2</td>
<td>4</td>
<td>-98.2632</td>
<td>15.7467</td>
<td>18</td>
<td>-6.24</td>
<td>&lt;.0001</td>
<td>Bonferroni</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>treat</td>
<td>3</td>
<td>4</td>
<td>-23.6316</td>
<td>11.9891</td>
<td>18</td>
<td>-1.97</td>
<td>0.0643</td>
<td>Bonferroni</td>
<td>0.3857</td>
</tr>
</tbody>
</table>

# Code for computing one sample Hotelling
# T-squared test using R
#
# dogs.R
#
# Data are read into a data frame from the file
#
# dogs.dat

dogdat<-read.table("c:/stat501/data/dogs.dat", header=F, col.names=c("dog", "x1", "x2", "x3", "x4"))
# Compute sample mean vector and sample covariance matrix

```r
xbar <- sapply(dogdat[, 2:5], mean)
xbar
x1 x2 x3 x4
368.2105 404.6316 479.2632 502.8947

xvar <- var(dogdat[, 2:5])
xvar
x1 x2 x3 x4
x1 2819.287 3568.415 2943.497 2295.357
x2 3568.415 7963.135 5303.991 4065.459
x3 2943.497 5303.991 6851.316 4499.640
x4 2295.357 4065.459 4499.640 4878.988
```

# Compute correlations

```r
xcorr <- round(cor(dogdat[, 2:5]), 4)
xcorr
x1 x2 x3 x4
x1 1.0000 0.7531 0.6697 0.6189
x2 0.7531 1.0000 0.7181 0.6522
x3 0.6697 0.7181 1.0000 0.7783
x4 0.6189 0.6522 0.7783 1.0000
```

# Display a scatterplot matrix

```r
pairs(dogdat[, 2:5], panel=function(x, y, ...) {points(x, y, ...) abline(lm(y ~ x), col = "black"), cex = 1.5)
```

# Apply the contrasts

```r
C <- matrix(c(1, -1, 1, -1, -1, -1, 1, 1, -1, 1, 1, -1),
            3, 4, byrow = T)
C
[1,]  1 -1  1 -1
[2,] -1 -1  1  1
[3,] -1  1  1 -1

Cxbar <- C %*% xbar
Cxvar <- C %*% xvar %*% t(C)
```
# Compute Hotelling statistic

```r
p <- nrow(Cxvar)
n <- nrow(dogdat)
nullmean <- c(0, 0, 0)
d <- Cxbar-nullmean
t2<-n*t(d)%*%solve(Cxvar)%*%d;
t2mod<-(n-p)*t2/(p*(n-1))
pval <- 1- pf(t2mod,p,n-p)
```

\[
t2 \\
[1,] 116.0163
\]

\[
pval \\
[1,] 3.317767e-07
\]

## Comparing Mean Vectors for Two Populations

- Independent random samples, one sample from each of two populations
- Randomized experiment: \( n_1 \) units are randomly allocated to treatment 1 and \( n_2 \) units are randomly allocated to treatment 2. Sample sizes need not be equal.
- Measure \( p \) outcomes (or variables or traits) on each unit

### Comparing two mean vectors (cont’d)

- If \( n_1, n_2 \) are large, the following assumptions are all we need to make inferences about the difference between treatments \( \mu_1 - \mu_2 \):
  1. \( X_{11}, X_{12}, ..., X_{1n_1} \sim p\text{-variate distribution}(\mu_1, \Sigma_1) \).
  2. \( X_{21}, X_{22}, ..., X_{2n_2} \sim p\text{-variate distribution}(\mu_2, \Sigma_2) \).
  3. \( X_{11}, X_{12}, ..., X_{1n_1} \) are independent
  4. \( X_{21}, X_{22}, ..., X_{2n_2} \) are independent.
  5. \( X_{11}, X_{12}, ..., X_{1n_1} \) are independent of \( X_{21}, X_{22}, ..., X_{2n_2} \).
Comparing two mean vectors (cont’d)

- When sample sizes are small, we also need that both distributions are multivariate normal.

- We will first examine the situation where the population covariance matrices are the same: $\Sigma_1 = \Sigma_2$. This is a strong assumption because it says that all $p(p+1)/2$ variances and covariances are the same in both populations.

Pooled estimate of the covariance matrix

- If $\Sigma_1 = \Sigma_2 = \Sigma$, then

$$S_1 = \frac{1}{n_1 - 1} \sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)(x_{1j} - \bar{x}_1)', \quad S_2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)(x_{2j} - \bar{x}_2)'$$

are both unbiased estimates of $\Sigma$. Then, we can pool or average the information from the two samples to obtain an estimate of the common covariance matrix:

$$S_{pool} = \frac{\sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)(x_{1j} - \bar{x}_1)'}{n_1 + n_2 - 2} + \frac{\sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)(x_{2j} - \bar{x}_2)'}{n_1 + n_2 - 2} \cdot S_1 + \frac{\sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)(x_{2j} - \bar{x}_2)'}{n_1 + n_2 - 2} \cdot S_2.$$
Example 6.3: Two Processes for Manufacturing Soap

- Objective was to compare two processes for manufacturing soap. Outcome measures were $X_1 = \text{lather}$ and $X_2 = \text{mildness}$, and $n_1 = n_2 = 50$.

- Sample statistics for sample 1 were
  
  $\bar{x}_1 = \begin{bmatrix} 8.3 \\ 4.1 \end{bmatrix}$, $S_1 = \begin{bmatrix} 2 & 1 \\ 1 & 6 \end{bmatrix}$,

  and for sample 2:
  
  $\bar{x}_2 = \begin{bmatrix} 10.2 \\ 3.9 \end{bmatrix}$, $S_2 = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$.

The pooled estimate of the common covariance matrix and the difference in sample mean vectors are

$S_{pool} = \frac{49}{98}S_1 + \frac{49}{98}S_2 = \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}$, $\bar{x}_1 - \bar{x}_2 = \begin{bmatrix} -1.9 \\ 0.2 \end{bmatrix}$.

We reject $H_0 : \mu_1 - \mu_2 = 0$ at level $\alpha = .05$ because

$T^2 = (\bar{x}_1 - \bar{x}_2 - 0)' \left[ \frac{1}{n_1} + \frac{1}{n_2} \right] S_{pool}^{-1} (\bar{x}_1 - \bar{x}_2 - 0) = 15.66$

is larger than

\[
\frac{(50 + 50 - 2)(2)}{(50 + 50 - 2 - 1)} F_{2.97(.05)} = 6.26.
\]

Example 6.3: Two processes for manufacturing soap (cont’d)

- A 95% confidence ellipse for the difference between the two means is centered at $\bar{x}_1 - \bar{x}_2$.

- Eigenvalues and eigenvectors of the pooled covariance matrix are given by

$\lambda = \begin{bmatrix} 5.303 \\ 1.697 \end{bmatrix}$, $E = [e_1, e_2] = \begin{bmatrix} 0.290 & 0.957 \\ 0.957 & -0.290 \end{bmatrix}$.

Two processes for manufacturing soap (cont’d)

- Since

\[
\left( \frac{1}{n_1} + \frac{1}{n_2} \right)c^2 = \left( \frac{1}{n_1} + \frac{1}{n_2} \right) \frac{(n_1 + n_2 - 2)p}{(n_1 + n_2 - p - 1)} F_{p,n_1+n_2-p-1}(\alpha)
\]

\[
= \left( \frac{1}{50} + \frac{1}{50} \right) \frac{(98)(2)}{(97)} F_{2.97(0.05)} = 0.25,
\]

we know that the ellipse extends $\sqrt{5.303 \times 0.25} = 1.15$ and $\sqrt{1.697 \times 0.25} = 0.65$ units in the $e_1$ and $e_2$ directions, respectively.
Two processes for manufacturing soap (cont’d)

- Since $\mu_1 - \mu_2 = 0$ is not inside the ellipse, we conclude that the two processes produce different results. There appears to be no big difference in mildness ($X_2$), but soaps made with the second process produce more lather.

Confidence Intervals

- As before, we can obtain simultaneous confidence intervals for any linear combination of the components of $\mu_1 - \mu_2$.

- In particular, in the case of $p$ variables, we might be interested in

$$a'_i(\mu_1 - \mu_2) = \begin{bmatrix} 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_{11} - \mu_{21} \\ \mu_{12} - \mu_{22} \\ \vdots \\ \mu_{1p} - \mu_{2p} \end{bmatrix} = \mu_{1i} - \mu_{2i},$$

where the vector $a_i$ has zeros everywhere except for the one in the $i$th position.

- Typically, we would be interested in $p$ such comparisons.

Confidence Intervals (cont’d)

- In general,

$$a'(\bar{x}_1 - \bar{x}_2) \pm t_{n_1 + n_2 - 2} (\alpha / 2) \sqrt{a'(\frac{1}{n_1} + \frac{1}{n_2}) S_{pool} a}$$

will simultaneously cover the true values of $a'(\mu_1 - \mu_2)$ with probability of at least $(1 - \alpha)\%$.

- One-at-a-time $t$ intervals would be computed as

$$a'(\bar{x}_1 - \bar{x}_2) \pm t_{n_1 + n_2 - 2} (\alpha / 2) \sqrt{a'(\frac{1}{n_1} + \frac{1}{n_2}) S_{pool} a}$$

and have less than $(1 - \alpha)\%$ simultaneous probability of coverage unless only one comparison is made. To apply the Bonferroni method, divide $\alpha$ by the number $m$ of comparisons of interest.

Heterogeneous covariance matrices

- Life gets difficult when we cannot assume that $\Sigma_1 = \Sigma_2$.

- We must modify the standardized distance measure, and the modification will not exactly be a multiple of an $F$-distribution when the null hypothesis of equal mean vectors is true.

- How to decide whether the assumption of equal covariance matrices is reasonable? Tests such as Bartlett’s test are sensitive to departures from normality.

- A crude rule of thumb is the following: if $\sigma_{1,ii} \geq 4 \sigma_{2,ii}$ or $\sigma_{2,ii} \geq 4 \sigma_{1,ii}$, then it is likely that $\Sigma_1 \neq \Sigma_2$. 


Bartlett’s test for equality of $k$ covariance matrices

- Tests for testing homogeneity of variances are touchy in that they are sensitive to the assumption of multivariate normality. They tend to reject homogeneity of covariance matrices too often when the samples are not selected from multivariate normal distributions.

- Bartlett’s test is a likelihood ratio test (independent random samples from multivariate normal distributions) for testing

$$H_0 : \Sigma_1 = \Sigma_2 = \ldots = \Sigma_k = \Sigma$$

versus the alternative where at least two $\Sigma_i$ are different.

Bartlett’s test (cont’d)

- When all $k$ samples come from multivariate normal populations and when $n_i - p$ is relatively large for all $i = 1, \ldots, k$, it has been shown that

$$MC^{-1} \sim \chi^2_{df}, \quad df = \frac{1}{2}(k-1)(p+1)p,$$

where the scale factor $C^{-1}$ is given by

$$C^{-1} = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \left( \sum_i \frac{1}{(n_i - 1)} - \frac{1}{\sum_i (n_i - 1)} \right).$$

- The null hypothesis $H_0 : \Sigma_1 = \Sigma_2 = \ldots = \Sigma_k = \Sigma$ is rejected at level $\alpha$ if $MC^{-1} \geq \chi^2_{df}(\alpha)$, with degrees of freedom as defined above.

Example: soap manufacturing

- We test the hypothesis $H_0 : \Sigma_1 = \Sigma_2 = \Sigma$. Here, $k = 2$ and $p = 2$.

- From earlier results, we have:

$$|S_{pool}| = 9, \quad |S_1| = 11, \quad |S_2| = 7.$$

- Then

$$M = 98 \times \ln(9) - 49 \times \ln(11) - 49 \times \ln(7) = 2.4794$$

$$C^{-1} = 1 - \frac{2(2^2) + 3(2) - 1}{6(3)(1)} \left[ \frac{1}{49} + \frac{1}{49} - \frac{1}{98} \right] = 1 - 0.0221 = 0.9779$$

$$df = \frac{1}{2}(1)(3)(2) = 3.$$
Example: soap manufacturing (cont’d)

• We reject the null hypothesis if

\[ MC^{-1} = 2.4794 \times 0.9779 = 2.4246 \geq \chi^2(0.05) = 7.82. \]

• In this case, we fail to reject the null hypothesis and conclude that \( \Sigma_1 \) is similar to \( \Sigma_2 \). Thus, pooling the samples to obtain a single estimate of the common population variance is reasonable.

What to do when \( \Sigma_1 \neq \Sigma_2 \)

• Suppose that we reject \( H_0 \), so that \( \Sigma_1 \neq \Sigma_2 \). For \( n_1 - p \) and \( n_2 - p \) large, an approximate \( 100(1 - \alpha)\% \) CR for \( \mu_1 - \mu_2 \) is given by all \( \mu_1 - \mu_2 \) satisfying

\[ [\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)]' \left[ \frac{1}{n_1} S_1 + \frac{1}{n_2} S_2 \right]^{-1} [\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)] \leq \chi^2_p(\alpha). \]

• For large samples, an approximate test of \( H_0 : \mu_1 = \mu_2 \) is obtained by rejecting \( H_0 \) at level \( \alpha \) if

\[ T^2 = [\bar{x}_1 - \bar{x}_2]' \left[ \frac{1}{n_1} S_1 + \frac{1}{n_2} S_2 \right]^{-1} [\bar{x}_1 - \bar{x}_2] \geq \chi^2_p(\alpha). \]

Heterogeneous covariance matrices (cont’d)

• Note that if \( n_1 = n_2 = n \), then

\[
\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2 = \frac{1}{n}(S_1 + S_2) = \frac{(n - 1)S_1 + (n - 1)S_2}{n + n - 2} \left( \frac{1}{n} + \frac{1}{n} \right) = S_{\text{pool}} \left( \frac{1}{n} + \frac{1}{n} \right).
\]

• With equal sample sizes, the test statistic is the same as the statistic for homogeneous covariance matrices. Thus, the effect of heterogeneous matrices is less pronounced when sample sizes are equal.

/* This program computes a two-sample Hotelling T-squared test to compare mean yield point (X1) and mean ultimate strength (X2) of steel rolled at two different temperatures. It also tests for homogeneity of covariance matrices. It is posted on the course web page as steel.sas */

DATA SET1;
INFILE "c:\stat501\data\steel.dat";
INPUT TEMP X1 X2;
LABEL TEMP = TEMPERATURE
   X1 = YIELD POINT
   X2 = ULTIMATE STRENGTH;
run;
ods rtf file="steel.rtf";
ods graphics on;

/* Check normality assumption for each variable within each of the two treatment groups */

PROC SORT DATA=SET1; BY TEMP; run;
PROC UNIVARIATE DATA=SET1 NORMAL;
   BY TEMP;
   VAR X1 X2; qqplot;
run;

The UNIVARIATE Procedure
TEMPERATURE=1

Variable: X1 (YIELD POINT)

Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>W 0.989977</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D 0.158847</td>
<td>&gt;0.1500</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq 0.020761</td>
<td>&gt;0.2500</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq 0.147744</td>
<td>&gt;0.2500</td>
</tr>
</tbody>
</table>

Variable: X2 (ULTIMATE STRENGTH)

Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>W 0.952351</td>
<td>&gt;0.0500</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D 0.179821</td>
<td>&gt;0.1500</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq 0.031987</td>
<td>&gt;0.2500</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq 0.206799</td>
<td>&gt;0.2500</td>
</tr>
</tbody>
</table>

The UNIVARIATE Procedure
TEMPERATURE=2

Variable: X1 (YIELD POINT)

Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>W 0.958652</td>
<td>&gt;0.0500</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D 0.18436</td>
<td>&gt;0.1500</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq 0.030976</td>
<td>&gt;0.2500</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq 0.203144</td>
<td>&gt;0.2500</td>
</tr>
</tbody>
</table>

Variable: X2 (ULTIMATE STRENGTH)

Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>W 0.918311</td>
<td>&gt;0.0500</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D 0.269042</td>
<td>&gt;0.1500</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq 0.063681</td>
<td>&gt;0.2500</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq 0.37356</td>
<td>&gt;0.2500</td>
</tr>
</tbody>
</table>
/* Check homogeneity of covariance matrices */

PROC DISCRIM DATA=SET1 POOL=TEST SLPOOL=.01 WCOV PCOV;
  CLASS TEMP;
  VAR X1-X2; run;

Effect of Rolling Temperature on Strength of Steel

The DISCRIM Procedure

Total Sample Size  12 DF Total  11
Variables  2 DF Within Classes  10
Classes  2 DF Between Classes  1

Class Level Information

<table>
<thead>
<tr>
<th>TEMP</th>
<th>Variable Name</th>
<th>Frequency</th>
<th>Weight</th>
<th>Proportion</th>
<th>Prior Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>_1</td>
<td>5</td>
<td>5.0000</td>
<td>0.416667</td>
<td>0.500000</td>
</tr>
<tr>
<td>2</td>
<td>_2</td>
<td>7</td>
<td>7.0000</td>
<td>0.583333</td>
<td>0.500000</td>
</tr>
</tbody>
</table>

The DISCRIM Procedure

Within-Class Covariance Matrices

<table>
<thead>
<tr>
<th>TEMP</th>
<th>DF</th>
<th>X1</th>
<th>X2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>7.300000000</td>
<td>4.200000000</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>8.333333333</td>
<td>6.666666667</td>
</tr>
</tbody>
</table>

The GLM Procedure

Pooled Within-Class Covariance Matrix, DF = 10

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>X1</th>
<th>X2</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>YIELD POINT</td>
<td>7.9200000000</td>
<td>5.6800000000</td>
</tr>
<tr>
<td>X2</td>
<td>ULTIMATE STRENGTH</td>
<td>5.6800000000</td>
<td>6.291428571</td>
</tr>
</tbody>
</table>

Within Covariance Matrix Information

<table>
<thead>
<tr>
<th>TEMP</th>
<th>Covariance Matrix Rank</th>
<th>Natural Log of the Determinant of the Covariance Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>2.62104</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2.94694</td>
</tr>
<tr>
<td>Pooled</td>
<td>2</td>
<td>2.86595</td>
</tr>
</tbody>
</table>

Chi-Square DF Pr > ChiSq

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
<td>0.380775</td>
<td>3</td>
<td>0.9442</td>
</tr>
</tbody>
</table>

/* Compute 2-sample Hotelling T-squared test */

PROC GLM DATA=SET1;
  CLASS TEMP;
  MODEL X1 X2 = TEMP / P SOLUTION;
  MANOVA H=TEMP / PRINTH PRINTE;
  LSMEANS TEMP / PDIFF STDERR;
run;
Effect of Rolling Temperature on Strength of Steel

The GLM Procedure

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1</td>
<td>19.71666667</td>
<td>19.71666667</td>
<td>2.49</td>
<td>0.1457</td>
</tr>
<tr>
<td>Error</td>
<td>10</td>
<td>79.20000000</td>
<td>7.92000000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>11</td>
<td>98.91666667</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEMP</td>
<td>1</td>
<td>19.71666667</td>
<td>19.71666667</td>
<td>2.49</td>
<td>0.1457</td>
</tr>
</tbody>
</table>

Effect of Rolling Temperature on Strength of Steel

The GLM Procedure

Dependent Variable: X2  ULTIMATE STRENGTH

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1</td>
<td>13.75238095</td>
<td>13.75238095</td>
<td>2.19</td>
<td>0.1701</td>
</tr>
<tr>
<td>Error</td>
<td>10</td>
<td>62.91428571</td>
<td>6.29142857</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>11</td>
<td>76.66666667</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEMP</td>
<td>1</td>
<td>13.75238095</td>
<td>13.75238095</td>
<td>2.19</td>
<td>0.1701</td>
</tr>
</tbody>
</table>
\textbf{The GLM Procedure} \\
\textit{Multivariate Analysis of Variance}

\begin{align*}
E &= \text{Error SSCP Matrix} \\
| & \begin{array}{cc}
X1 & X2 \\
X1 & 79.2 & 56.8 \\
X2 & 56.8 & 62.914285714
\end{array}
\end{align*}

\textbf{Partial Correlation Coefficients from the Error SSCP Matrix} \\
\text{DF} = 10 \\
\text{X1} & \begin{array}{c}
1.000000 \\
0.804658
\end{array} & \begin{array}{c}
0.0028
\end{array} \\
\text{X2} & \begin{array}{c}
0.804658 \\
0.0028
\end{array} & \begin{array}{c}
1.000000
\end{array}

\textbf{Characteristic Roots and Vectors of: E Inverse * H, where} \\
E = \text{Error SSCP Matrix} \\
H = \text{Type III SSCP Matrix for TEMP} \\
\begin{align*}
\text{Characteristic Root} & \quad \text{Percent} \\
& \begin{array}{r}
2.39117057 \\
0.0000000
\end{array} & \begin{array}{r}
100.00 \\
0.00
\end{array} \\
\text{Characteristic Vector} & \quad V'V=1 \\
& \begin{array}{cc}
X1 & X2 \\
-0.18039499 & 0.00000000 \\
0.20098179 & 0.05722038 \\
0.00000000 & 0.06851388
\end{array}
\end{align*}

\textbf{MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall TEMP Effect} \\
E = \text{Error SSCP Matrix} \\
H = \text{Type III SSCP Matrix for TEMP} \\
S=1 \quad M=0 \quad N=3.5 \\
\begin{align*}
\text{Statistic} & \quad \text{Value} & \text{F Value} & \text{Num DF} & \text{Den DF} & \text{Pr > F} \\
\text{Wilks' Lambda} & 0.29488343 & 10.76 & 2 & 9 & 0.0041 \\
\text{Pillai's Trace} & 0.70511657 & 10.76 & 2 & 9 & 0.0041 \\
\text{Hotelling-Lawley Trace} & 2.39117057 & 10.76 & 2 & 9 & 0.0041 \\
\text{Roy's Greatest Root} & 2.39117057 & 10.76 & 2 & 9 & 0.0041
\end{align*}

\textbf{Least Squares Means}

\textbf{TEMP} \quad \text{X1 LSMEAN} \quad \text{Standard Error} \quad \text{H0:LSMEA N=0} \quad \text{Pr > |t|} \quad \text{H0:LSMean1=LSMean2} \quad \text{Pr > |t|}
\begin{align*}
\text{TEMP} & \quad \text{X1} & \text{Standard Error} & \text{H0:LSMEA N=0} & \text{Pr > |t|} & \text{H0:LSMean1=LSMean2} & \text{Pr > |t|} \\
1 & 36.4000000 & 1.2585706 & <.0001 & 0.1457 \\
2 & 39.0000000 & 1.0636863 & <.0001 \\
\text{TEMP} & \quad \text{X2 LSMEAN} \quad \text{Standard Error} \quad \text{H0:LSMEA N=0} \quad \text{Pr > |t|} \quad \text{H0:LSMean1=LSMean2} \quad \text{Pr > |t|}
\begin{align*}
\text{TEMP} & \quad \text{X2} & \text{Standard Error} & \text{H0:LSMEA N=0} & \text{Pr > |t|} & \text{H0:LSMean1=LSMean2} & \text{Pr > |t|} \\
1 & 62.6000000 & 1.1217334 & <.0001 & 0.1701 \\
2 & 60.4285714 & 0.9480377 & <.0001
\end{align*}

/* Plot the data */

TITLE1 H=2.0 F=swiss "Effect of Rolling Temperature"
TITLE2 H=2.0 F=swiss "on Strength of Steel"
proc sgplot data=set1;
  scatter x=x1 y=x2 / group=temp
    markerattrs=(size=12 symbol=circlefilled);
  yaxis label="Ultimate Strength"
    labelattrs=(size=17) valueattrs=(size=15);
  xaxis label="Yield Point"
    labelattrs=(size=17) valueattrs=(size=15);
run;
ods graphics off;
ods rtf close;

# Code for computing a two-sample Hotelling T-squared test. The code is posted as steel.R
# Data are read into a data frame from the file steel.dat

steel<-read.table("c:/stat501/data/steel.dat",
  header=F, col.names=c("temp", "x1", "x2"))

```
steel
   temp x1 x2
1    1 33 60
2    1 36 61
3    1 35 64
4    1 38 63
5    1 40 65
6    2 35 57
7    2 36 59
8    2 38 59
9    2 39 61
10   2 41 63
11   2 43 65
12   2 41 59
```
# Compute sample mean vector and sample covariance matrix for each temperature

```r
xbar1 <- sapply(steel[steel$temp==1, 2:3], mean)
xbar1
x1  x2
36.4 62.6

xvar1 <- var(steel[steel$temp==1 ,2:3])
xvar1
x1  x2
x1 7.3 4.2
x2 4.2 4.3
```

```r
xbar2 <- sapply(steel[steel$temp==2, 2:3], mean)
xbar2
x1  x2
39.00000 60.42857

xvar2 <- var(steel[steel$temp==2 ,2:3])
xvar2
x1  x2
x1 8.333333 6.666667
x2 6.666667 7.619048
```

# Display scatterplots

```r
par(mfrow=c(2,2), pch=5)

plot(steel[steel$temp==1 ,2], steel[steel$temp==1 ,3],
     xlab = "Yield Point",
     ylab = "Ultimate Strength",
     main="Rolling Temperature 1", cex=1.5)
abline(lm(steel[steel$temp==1 ,3] ~ steel[steel$temp==1 ,2]))

plot(steel[steel$temp==2 ,2], steel[steel$temp==2 ,3],
     xlab = "Yield Point",
     ylab = "Ultimate Strength",
     main="Rolling Temperature 2", cex=1.5)
abline(lm(steel[steel$temp==2 ,3] ~ steel[steel$temp==2 ,2]))
```
# Test for univariate normality of the variables for
# the first rolling temperature
apply(steel[steel$temperature == 1, -1], 2, shapiro.test)
$yield

  Shapiro-Wilk normality test
  data: newX[, i]
  W = 0.99, p-value = 0.9796

$strength

  Shapiro-Wilk normality test
  data: newX[, i]
  W = 0.9524, p-value = 0.754

# Test for univariate normality of the variables for
# the second rolling temperature
apply(steel[steel$temperature == 2, -1], 2, shapiro.test)
$yield

  Shapiro-Wilk normality test
  data: newX[, i]
  W = 0.9587, p-value = 0.8071

$strength

  Shapiro-Wilk normality test
  data: newX[, i]
  W = 0.9183, p-value = 0.4564

# Check for multivariate normality at each temperature.
# First load the energy library
library(energy)
mvnorm.etest(x = steel[steel$temperature == 1, -1], R=10000)

  Energy test of multivariate normality: estimated parameters
  data: x, sample size 5, dimension 2, replicates 10000
  E-statistic = 0.5402, p-value = 0.6107

mvnorm.etest(x = steel[steel$temperature == 2, -1], R=10000)

  Energy test of multivariate normality: estimated parameters
  data: x, sample size 7, dimension 2, replicates 10000
  E-statistic = 0.5866, p-value = 0.5156

# Check for homogeneity of variance-covariance matrix assumptions.
# Box's M-test for testing homogeneity of covariance matrices
# Written by Andy Liaw (2004) converted from Matlabhelp
# Andy's note indicates that he has left the original
# Matlab comments intact.
# Slight clean-up and fix with corrected documentation
# provided by Ranjan Maitra (2012)
BoxMTest <- function(X, cl, alpha=0.05) {

  ## Multivariate Statistical Testing for the Homogeneity of
  ## Covariance Matrices by the Box's M.
## Syntax: function [MBox] = BoxMTest(X,alpha)

## Inputs:
## X = data matrix (Size of matrix must be n-by-p; 
## cl = name of the grouping variable 
## alpha = significance level (default = 0.05). 
## Output:
## MBox = the Box's M statistic. 
## Chi-square or F = the approximation test statistic. 
## df's = degrees' of freedom of the approximation 
## statistic test. 
## P = observed significance level. 
## 
## If the groups sample-size is at least 20 (sufficiently large), 
## Box's M test takes a Chi-square approximation; otherwise it take 
## an F approximation.

## Created by A. Trujillo-Ortiz and R. Hernandez-Walls
## Facultad de Ciencias Marinas
## Universidad Autonoma de Baja California
## Apdo. Postal 453
## Ensenada, Baja California
## Mexico.
## atrujo_at_uabc.mx
## And the special collaboration of the post-graduate 
## students of the 2002:2 
## Multivariate Statistics Course: Karel Castro-Morales, 
## Alejandro Espinoza-Tenorio, Andrea Guia-Ramirez, 
## Raquel Muniz-Salazar, Jose Luis Sanchez-Osorio and 
## Roberto Carmona-Pina, November 2002.

## References:
## Stevens, J. (1992), Applied Multivariate Statistics 
## for Social Sciences. 2nd. ed., New-Jersey:Lawrance 

if (alpha <= 0 || alpha >= 1)
    stop('significance level must be between 0 and 1')
g = nlevels(cl) ## Number of groups.
n = table(cl) ## Vector of groups-size.
N = nrow(X)
p = ncol(X)
bandera = 2
if (any(n >= 20)) bandera = 1

## Partition of the group covariance matrices.
covList <- tapply(as.matrix(X), rep(cl, ncol(X)), 
    function(x, nc) cov(matrix(x, nc = nc), 
    ncol(X)))
deno = sum(n) - g
suma = array(0, dim=dim(covList[[1]]))
for (k in 1:g)
    suma = suma + (n[k] - 1) * covList[[k]]
Sp = suma / deno ## Pooled covariance matrix.
Falta=0
for (k in 1:g)
    Falta = Falta + ((n[k] - 1) * log(det(covList[[k]]))))
\[ MB = (\sum(n) - g) \times \log(det(Sp)) - \text{Falta} \] ## Box's M statistic.
\[ \text{suma1} = \sum(1 / (n[1:g] - 1)) \]
\[ \text{suma2} = \sum(1 / ((n[1:g] - 1)^2)) \]
\[ C = (((2 \times p^2) + (3 \times p) - 1) / (6 \times (p + 1) \times (g - 1))) \times \text{suma1} - (1 / \text{deno}) \] ## Computing of correction factor.
if (bandera == 1)
{
    \[ X2 = MB \times (1 - C) \] ## Chi-square approximation.
    \[ v = \text{as.integer}(p \times (p + 1) \times (g - 1)) / 2 \] ## Degrees of freedom.
    \[ \text{P} = \text{pchisq}(X2, v, \text{lower}=\text{FALSE}) \] ## RM: corrected to be the upper
    \[ \text{cat('------------------------------------------------
    cat(' MBox Chi-sqr. df P
    cat('------------------------------------------------
    \[ \text{cat(sprintf("%10.4f%11.4f%12.4f%13.4f\n", MB, X2, v, P))} \]
    \[ \text{cat('------------------------------------------------
}
else
{
    ## To obtain the F approximation we first define Co, which combines the before C value are used to estimate the denominator degree of freedom (v2); resulting two possible cases.
    \[ Co = (((p-1) \times (p+2)) / (6 \times (g-1))) \times (\text{suma2} - (1 / \text{deno}^2)) \]
    if (Co - (C^2) >= 0)
    {
        \[ v1 = \text{as.integer}(p \times (p + 1) \times (g - 1)) / 2 \] ## Numerator DF.
        \[ v21 = \text{as.integer}(\text{trunc}((v1 + 2) / (Co - (C^2)))) \] ## Denominator DF.
        \[ \text{F1} = MB \times (1 - C - (v1 / v21)) / v1 \] ## F approximation.
        \[ \text{P1} = \text{pf}(F1, v1, v21, \text{lower}=false) \]
        \[ \text{cat('------------------------------------------------
        cat(' MBox F df1 df2 P
        cat('------------------------------------------------
        \[ \text{cat(sprintf("%10.4f%11.4f%12.4f%13.4f\n", MB, F1, v1, v21))} \]
        \[ \text{cat('------------------------------------------------
    }\]
else
{
    \[ v1 = \text{as.integer}(p \times (p + 1) \times (g - 1)) / 2 \] ## Numerator df.
    \[ v22 = \text{as.integer}(\text{trunc}((v1 + 2) / (C^2 - Co))) \] ## Denominator df.
    \[ b = v22 / (1 - C - (2 / v22)) \]
    \[ F2 = (v22 \times MB) / (v1 * (b - MB)) \] ## F approximation.
    \[ \text{P2} = \text{pf}(F2, v1, v22, \text{lower}=false) \]
    \[ \text{cat('------------------------------------------------
    cat(' MBox F df1 df2 P
    cat('------------------------------------------------
    \[ \text{cat(sprintf("%10.4f%11.4f%12.4f%13.4f\n", MB, F2, v1, v22))} \]
    \[ \text{cat('------------------------------------------------
}
}}
if (P2 >= alpha) {
    cat('Covariance matrices are not significantly different.\n')
} else {
    cat('Covariance matrices are significantly different.\n')
}return(list(MBox=MB, F=F2, df1=v1, df2=v22, pValue=P2))
#
# This is the end of the definition of the function

cov(steel[steel$temperature == 1, -1])
yield strength
  yield    7.3  4.2
  strength 4.2  4.3
cov(steel[steel$temperature == 2, -1])
yield strength
  yield   8.333333 6.666667
  strength 6.666667 7.619048
BoxMTest(steel[, -1], cl = as.factor(steel$temperature),
        alpha=0.05)

---------------------------------
  MBox F df1 df2 P
---------------------------------
  0.4937 0.1268 3 4121 0.9443
---------------------------------
Covariance matrices are not significantly different.

$MBox
1
0.4936815

$F
1
0.1268051

$df1
[1] 3

$df2
[1] 4121

$pValue
1
0.944252
# Now perform the two-sample Hotelling T²-test
# First attach the ICSNP library

library(ICSNP)

HotellingsT2(steel[steel$temperature == 1, -1],
        steel[steel$temperature == 2, -1])

Hotelling's two sample T²-test

data:  steel[steel$temperature == 1, -1] and
        steel[steel$temperature == 2, -1]

T.2 = 10.7603, df1 = 2, df2 = 9, p-value = 0.004106
alternative hypothesis: true location difference is not
equal to c(0,0)