Chapter 6

6.1 Eigenvalues and eigenvectors of S_d are:

$$\lambda_1 = 449.778, \quad \underset{\sim}{e}_1' = [.333, \ .943]$$
$$\lambda_2 = 168.082, \quad \underset{\sim}{e}_2' = [.943, \ -.333]$$

Ellipse centered at $\bar{\underset{\sim}{d}}' = [-9.36, \ 13.27]$. Half length of major axis is 20.57 units. Half length of minor axis is 12.58 units. Major and minor axes lie in $\underset{\sim}{e}_1$ and $\underset{\sim}{e}_2$ directions, respectively.

Yes, the test answers the question: Is $\underset{\sim}{\delta} = \underset{\sim}{0}$ inside the 95% confidence ellipse?

6.2 Using a critical value $t_{n-1}(\alpha/2p) = t_{10}(0.0125) = 2.6338$,

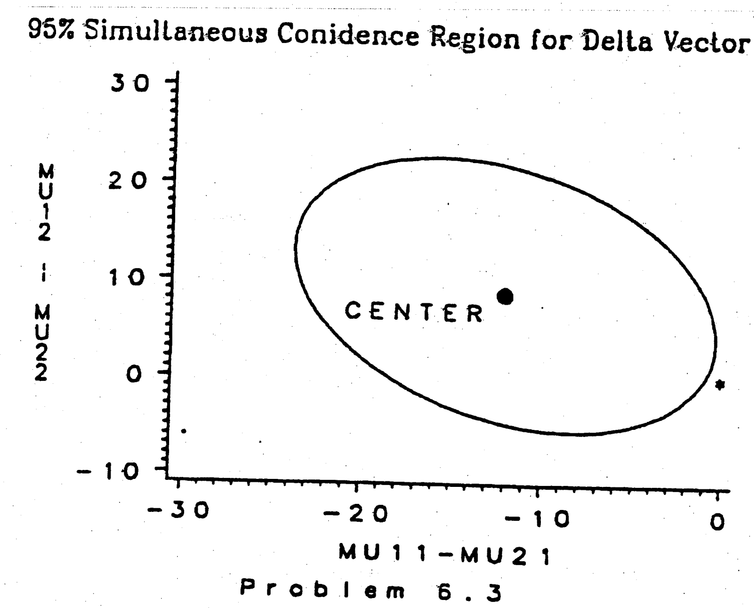
	LOWER	UPPER
Bonferroni C. I.:	-20.57	1.85
	-2.97	29.52
Simultaneous C. I.:	-22.45	3.73
	-5.70	32.25

Simultaneous confidence intervals are larger than Bonferroni's confidence intervals.

6.3 The 95% Bonferroni intervals are

	LOWER	UPPER
Bonferroni C. I.:	-21.92	-2.08
	-3.36	20.56
Simultaneous C. I.:	-23.70	-0.30
	-5.50	22.70

Since the hypothesized vector $\delta = 0$ (denoted as $*$ in the plot) is outside the joint confidence region, we reject $H_0 : \delta = 0$. Bonferroni C.I. are consistent with this result. After the elimination of the outlier, the difference between pairs became significant.



6.4

(a). Hotelling's $T^2 = 10.215$. Since the critical point with $\alpha = 0.05$ is 9.459, we reject $H_0 : \underset{\sim}{\delta} = \underset{\sim}{0}$.

(b).

	Lower	Upper
Bonferroni C. I.:	−1.09	−0.02
	−0.04	0.64
T^2 Simultaneous C. I.:	−1.18	0.07
	−0.10	0.69

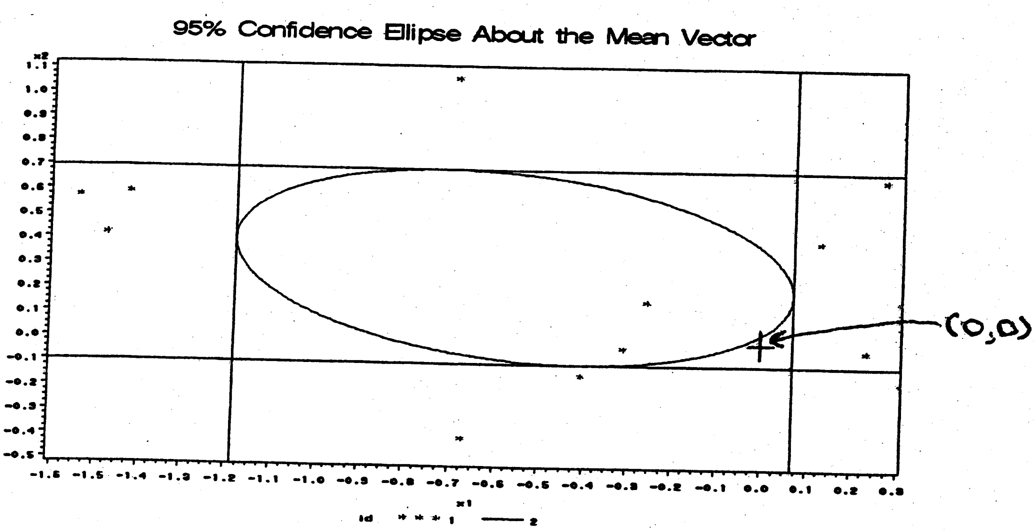
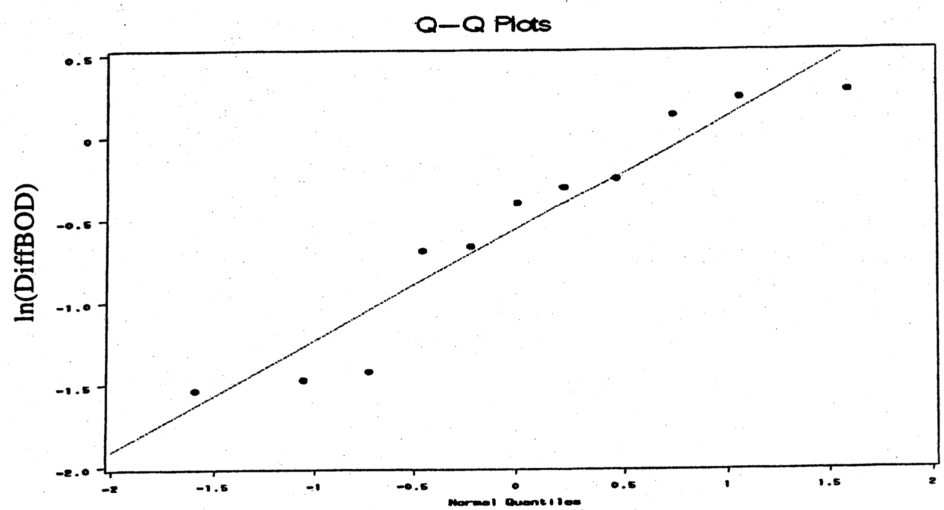
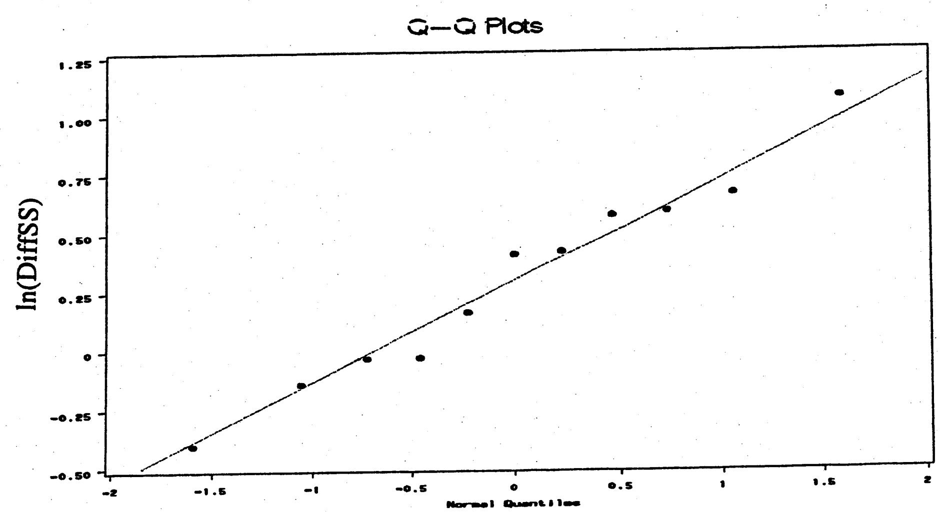
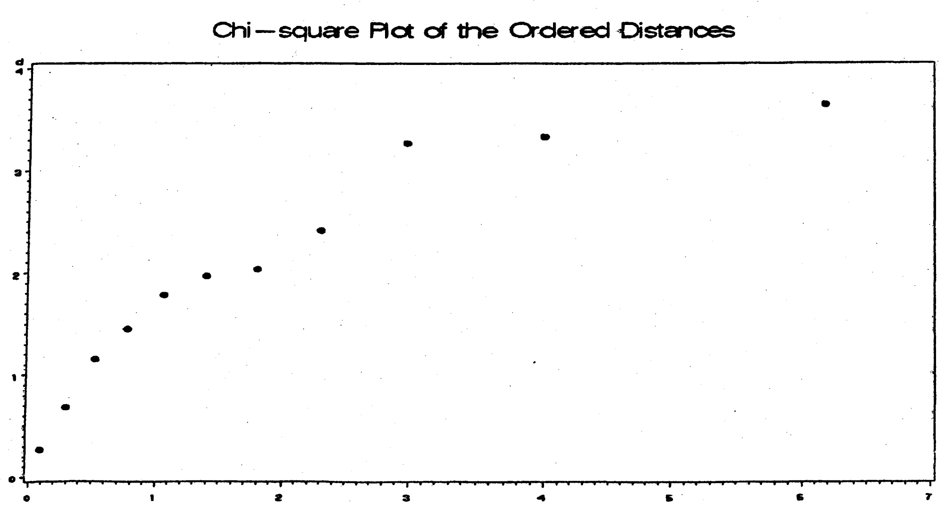


Figure 1: 95% Confidence Ellipse and Simultaneous T^2 Intervals for the Mean Difference

(c) The Q-Q plots for ln(DiffBOD) and ln(DiffSS) are shown below. Marginal normality cannot be rejected for either variable. The χ^2 plot is not straight (with at least one apparent bivariate outlier) and, although the sample size (n =11) is small, it is difficult to argue for bivariate normality.







6.5 a) H_0: $C\underset{\sim}{\mu} = \underset{\sim}{0}$ where $C = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$, $\underset{\sim}{\mu}' = [\mu_1, \mu_2, \mu_3]$.

$$C\underset{\sim}{\bar{x}} = \begin{bmatrix} -11.2 \\ 6.9 \end{bmatrix}, \quad CSC' = \begin{bmatrix} 55.5 & -32.6 \\ -32.6 & 66.4 \end{bmatrix}$$

$$T^2 = n(C\underset{\sim}{\bar{x}})'(CSC')^{-1}(C\underset{\sim}{\bar{x}}) = 90.4; \quad n = 40; \quad q = 3$$

$$\frac{(n-1)(q-1)}{(n-q+1)} F_{q-1,n-q+1}(.05) = \frac{(39)2}{38}(3.25) = 6.67$$

Since $T^2 = 90.4 > 6.67$ reject H_0: $C\underset{\sim}{\mu} = \underset{\sim}{0}$

b) 95% simultaneous confidence intervals:

$$\mu_1 - \mu_2: (46.1 - 57.3) \pm \sqrt{6.67} \sqrt{\frac{55.5}{40}} = -11.2 \pm 3.0$$

$$\mu_2 - \mu_3: 6.9 \pm 3.3$$

$$\mu_1 - \mu_3: -4.3 \pm 3.3$$

The means are all different from one another.

6.6 a) Treatment 2: Sample mean vector $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$; sample covariance matrix

$$\begin{bmatrix} 1 & -3/2 \\ -3/2 & 3 \end{bmatrix}$$

Treatment 3: Sample mean vector $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$; sample covariance matrix

$$\begin{bmatrix} 2 & -4/3 \\ -4/3 & 4/3 \end{bmatrix}$$

$$S_{pooled} = \begin{bmatrix} 1.6 & \\ -1.4 & \end{bmatrix}$$

b) $$T^2 = [2-3,\ 4-2] \left[\left(\frac{1}{3} + \frac{1}{4}\right) \begin{bmatrix} 1.6 & -1.4 \\ -1.4 & 2 \end{bmatrix} \right]^{-1} \begin{bmatrix} 2-3 \\ 4-2 \end{bmatrix} = 3.88$$

$$\frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)} F_{p,n_1+n_2-p-1}(.01) = \frac{(5)2}{4}(18) = 45$$

Since $T^2 = 3.88 < 45$ do not reject $H_0: \underset{\sim}{\mu}_2 - \underset{\sim}{\mu}_3 = \underset{\sim}{0}$ at the $\alpha = .01$ level.

c). 99% simultaneous confidence intervals:

$$\mu_{21} - \mu_{31}: (2-3) \pm \sqrt{45}\ \sqrt{\left(\frac{1}{3}+\frac{1}{4}\right)1.6} = -1 \pm 6.5$$

$$\mu_{22} - \mu_{32}: 2 \pm 7.2$$

6.7 $$T^2 = [74.4 \quad 201.6] \left[\left(\frac{1}{45} + \frac{1}{55}\right) \begin{bmatrix} 10963.7 & 21505.5 \\ 21505.5 & 63661.3 \end{bmatrix} \right]^{-1} \begin{bmatrix} 74.4 \\ 201.6 \end{bmatrix} = 16.1$$

$$\frac{(n_1+n_2-2)p}{n_1+n_2-p-1} F_{p,n_1+n_2-p-1}(.05) = 6.26$$

Since $T^2 = 16.1 > 6.26$ reject $H_0: \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 = \underset{\sim}{0}$ at the $\alpha = .05$ level.

$$\hat{\underset{\sim}{a}} \propto S^{-1}_{pooled}(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2) = \begin{bmatrix} .0017 \\ .0026 \end{bmatrix}$$

6.8 a) For first variable:

observation = mean + treatment effect + residual

$$\begin{bmatrix} 6 & 5 & 8 & 4 & 7 \\ 3 & 1 & 2 & & \\ 2 & 5 & 3 & 2 & \end{bmatrix} = \begin{bmatrix} 4 & 4 & 4 & 4 & 4 \\ 4 & 4 & 4 & & \\ 4 & 4 & 4 & 4 & \end{bmatrix} + \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ -2 & -2 & -2 & & \\ -1 & -1 & -1 & -1 & \end{bmatrix} + \begin{bmatrix} 0 & -1 & 2 & -2 & 1 \\ 1 & -1 & 0 & & \\ -1 & 2 & 0 & -1 & \end{bmatrix}$$

$SS_{obs} = 246$ $\quad SS_{mean} = 192$ $\quad SS_{tr} = 36$ $\quad SS_{res} = 18$

For second variable:

$$\begin{bmatrix} 7 & 9 & 6 & 9 & 9 \\ 3 & 6 & 3 & & \\ 3 & 1 & 1 & 3 & \end{bmatrix} = \begin{bmatrix} 5 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & & \\ 5 & 5 & 5 & 5 & \end{bmatrix} + \begin{bmatrix} 3 & 3 & 3 & 3 & 3 \\ -1 & -1 & -1 & & \\ -3 & -3 & -3 & -3 & \end{bmatrix} + \begin{bmatrix} -1 & 1 & -2 & 1 & 1 \\ -1 & 2 & -1 & & \\ 1 & -1 & -1 & 1 & \end{bmatrix}$$

$SS_{obs} = 402$ $\quad SS_{mean} = 300$ $\quad SS_{tr} = 84$ $\quad SS_{res} = 18$

Cross product contributions:

275 $\quad$ 240 $\quad$ 48 $\quad$ -13

b) MANOVA table:

Source of Variation	SSP	d.f.
Treatment	$B = \begin{bmatrix} 36 & 48 \\ 48 & 84 \end{bmatrix}$	3 - 1 = 2
Residual	$W = \begin{bmatrix} 18 & -13 \\ -13 & 18 \end{bmatrix}$	5 + 3 + 4 - 3 = 9
Total (corrected)	$\begin{bmatrix} 54 & 35 \\ 35 & 102 \end{bmatrix}$	11

c) $\Lambda^* = \frac{|W|}{|B+W|} = \frac{155}{4283} = .0362$

Using Table 6.3 with $p = 2$ and $g = 3$

$$\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right)\left(\frac{\Sigma n_\ell - g - 1}{g-1}\right) = 17.02 .$$

Since $F_{4,16}(.01) = 4.77$ we conclude that treatment differences exist at $\alpha = .01$ level.

Alternatively, using Bartlett's procedure,

$$-\left(n-1-\frac{(p+g)}{2}\right)\ell n\ \Lambda^* = -(12-1-\tfrac{5}{2})\ell n(.0362) = 28.209$$

Since $\chi_4^2(.01) = 13.28$ we again conclude treatment differences exist at $\alpha = .01$ level.

6.9 For any matrix C

$$\underset{\sim}{\bar{d}} = \frac{1}{n}\Sigma\ \underset{\sim}{d}_j = C(\frac{1}{n}\Sigma\ \underset{\sim}{x}_j) = C\underset{\sim}{\bar{x}}$$

and $\underset{\sim}{d}_j - \underset{\sim}{\bar{d}} = C(\underset{\sim}{x}_j - \underset{\sim}{\bar{x}})$

so $S_d = \frac{1}{n-1}\Sigma(\underset{\sim}{d}_j-\underset{\sim}{\bar{d}})(\underset{\sim}{d}_j-\underset{\sim}{\bar{d}})' = C(\frac{1}{n-1}\Sigma(\underset{\sim}{x}_j-\underset{\sim}{\bar{x}})(\underset{\sim}{x}_j-\underset{\sim}{\bar{x}})')C' = CSC'$

6.10 $(\bar{x}\ \underset{\sim}{1})'[(\bar{x}_1-\bar{x})\underset{\sim}{u}_1 + \dots + (\bar{x}_g-\bar{x})\underset{\sim}{u}_g]$

$$= \bar{x}[(\bar{x}_1-\bar{x})n_1 + \dots + (\bar{x}_g-\bar{x})n_g]$$

$$= \bar{x}[n_1\bar{x}_1 + \dots + n_g\bar{x}_g-\bar{x}(n_1 + \dots + n_g)]$$

$$= \bar{x}[(n_1 + \dots + n_g)\bar{x} - \bar{x}(n_1 + \dots + n_g)] = 0$$

6.11 $L(\underset{\sim}{\mu}_1, \underset{\sim}{\mu}_2, \Sigma) = L(\underset{\sim}{\mu}_1, \Sigma) L(\underset{\sim}{\mu}_2, \Sigma)$

$$= \left[\frac{1}{(2\pi)^{\frac{(n_1+n_2)p}{2}} |\Sigma|^{\frac{n_1+n_2}{2}}} \right] \exp\left\{ -\frac{1}{2} \left(\text{tr } \Sigma^{-1}[(n_1-1)S_1 + (n_2-1)S_2] \right. \right.$$

$$\left. \left. + n_1(\bar{\underset{\sim}{x}}_1 - \underset{\sim}{\mu}_1)' \Sigma^{-1} (\bar{\underset{\sim}{x}} - \underset{\sim}{\mu}_1) + n_2(\bar{\underset{\sim}{x}}_2 - \underset{\sim}{\mu}_2)' \Sigma^{-1} (\bar{\underset{\sim}{x}}_2 - \underset{\sim}{\mu}_2) \right) \right\}$$

using (4-16) and (4-17). The likelihood is maximized with respect to $\underset{\sim}{\mu}_1$ and $\underset{\sim}{\mu}_2$ at $\hat{\underset{\sim}{\mu}}_1 = \bar{\underset{\sim}{x}}_1$ and $\hat{\underset{\sim}{\mu}}_2 = \bar{\underset{\sim}{x}}_2$ respectively and with respect to Σ at

$$\hat{\Sigma} = \frac{1}{n_1+n_2} [(n_1 - 1)S_1 + (n_2 - 2)S_2] = \left(\frac{n_1+n_2-2}{n_1+n_2} \right) S_{\text{pooled}}$$

(For the maximization with respect to Σ see Result 4.10 with $b = \frac{n_1+n_2}{2}$ and $B = (n_1 - 1)S_1 + (n_2 - 2)S_2$)

6.13 a) and b) For first variable:

$$\underset{\text{Observation}}{\begin{bmatrix} 6 & 4 & 8 & 2 \\ 3 & -3 & 4 & -4 \\ -3 & -4 & 3 & -4 \end{bmatrix}} = \underset{\text{mean}}{\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}} + \underset{\text{factor 1 effect}}{\begin{bmatrix} 4 & 4 & 4 & 4 \\ -1 & -1 & -1 & -1 \\ -3 & -3 & -3 & -3 \end{bmatrix}} + \underset{\text{factor 2 effect}}{\begin{bmatrix} 1 & -2 & 4 & -3 \\ 1 & -2 & 4 & -3 \\ 1 & -2 & 4 & -3 \end{bmatrix}} + \underset{\text{residual}}{\begin{bmatrix} 0 & 1 & -1 & 0 \\ 2 & -1 & 0 & -1 \\ -2 & 0 & 1 & 1 \end{bmatrix}}$$

$SS_{tot} = 220$ $SS_{mean} = 12$ $SS_{fac\,1} = 104$ $SS_{fac\,2} = 90$ $SS_{res} = 14$

For second variable:

$$\begin{bmatrix} 8 & 6 & 12 & 6 \\ 8 & 2 & 3 & 3 \\ 2 & -5 & -3 & -6 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \end{bmatrix} + \begin{bmatrix} 5 & 5 & 5 & 5 \\ 1 & 1 & 1 & 1 \\ -6 & -6 & -6 & -6 \end{bmatrix} + \begin{bmatrix} 3 & -2 & 1 & -2 \\ 3 & -2 & 1 & -2 \\ 3 & -2 & 1 & -2 \end{bmatrix} + \begin{bmatrix} -3 & 0 & 3 & 0 \\ 1 & 0 & -2 & 1 \\ 2 & 0 & -1 & -1 \end{bmatrix}$$

$SS_{tot} = 440$ $SS_{mean} = 108$ $SS_{fac\,1} = 248$ $SS_{fac\,2} = 54$ $SS_{res} = 30$

Sum of cross products:

$$SCP_{tot} = SCP_{mean} + SCP_{fac\,1} + SCP_{fac\,2} + SCP_{res}$$

$$227 = 36 + 148 + 51 - 8$$

c) MANOVA table:

Source of Variation	SSP	d.f.
Factor 1	$\begin{bmatrix} 104 & 148 \\ 148 & 248 \end{bmatrix}$	$g-1=3-1=2$
Factor 2	$\begin{bmatrix} 90 & 51 \\ 51 & 54 \end{bmatrix}$	$b-1=4-1=3$
Residual	$\begin{bmatrix} 14 & -8 \\ -8 & 30 \end{bmatrix}$	$(g-1)(b-1) = 6$
Total (Corrected)	$\begin{bmatrix} 208 & 191 \\ 191 & 332 \end{bmatrix}$	$gb-1 = 11$

d) We reject $H_0: \underset{\sim}{\tau}_1 = \underset{\sim}{\tau}_2 = \underset{\sim}{\tau}_3 = \underset{\sim}{0}$ at $\alpha = .05$ level since

$$-[(g-1)(b-1) - (\frac{p+1-(g-1)}{2})]\ell n\Lambda^* = -[6 - \frac{3-2}{2}]\ell n\left(\frac{|SS_{res}|}{|SSP_{fac\,1} + SSP_{res}|}\right)$$

$$= -5.5\,\ell n\left(\frac{356}{13204}\right) = 19.87 > \chi_4^2(.05) = 9.49$$

and conclude there are factor 1 effects.

We also reject $H_0: \underset{\sim}{\beta}_1 = \underset{\sim}{\beta}_2 = \underset{\sim}{\beta}_3 = \underset{\sim}{\beta}_4 = \underset{\sim}{0}$ at the $\alpha = .05$ level since

$$-[(g-1)(b-1) - (\frac{p+1-(b-1)}{2})\ln\Lambda^* = -[6 - \frac{3-3}{2}]\ln\left(\frac{|SSP_{res}|}{|SSP_{fac\,2} + SSP_{res}|}\right)$$

$$= -6\ln\left(\frac{356}{6887}\right) = 17.77 > \chi_6^2(.05) = 12.59$$

and conclude there are factor 2 effects.

6.14 b) MANOVA Table:

Source of Variation	SSP	d.f.
Factor 1	$\begin{bmatrix} 496 & 184 \\ 184 & 208 \end{bmatrix}$	2
Factor 2	$\begin{bmatrix} 36 & 24 \\ 24 & 36 \end{bmatrix}$	3
Interaction	$\begin{bmatrix} 32 & 0 \\ 0 & 44 \end{bmatrix}$	6
Residual	$\begin{bmatrix} 312 & -84 \\ -84 & 400 \end{bmatrix}$	12
Total (Corrected)	$\begin{bmatrix} 876 & 124 \\ 124 & 688 \end{bmatrix}$	23

c) Since $-[gb(n-1) - (p+1-(g-1)(b-1))/2]\ln\Lambda^* = -13.5\ln\left(\frac{|SSP_{res}|}{|SSP_{int} + SSP_{res}|}\right)$

$= -13.5\ln(.808) = 2.88 < \chi_{12}^2(.05) = 21.03$ we <u>do not reject</u>

$H_0: \underset{\sim}{\gamma}_{11} = \underset{\sim}{\gamma}_{12} = \cdots = \underset{\sim}{\gamma}_{34} = \underset{\sim}{0}$ (no interaction effects) at the

$\alpha = .05$ level.

Since

$$-[gb(n-1)-(p+1-(g-1))/2]\ell n\Lambda^* = -11.5\ell n\left(\frac{|SSP_{res}|}{|SSP_{fac\,1} + SSP_{res}|}\right)$$

$= -11.5\ell n(.2447) = 16.19 > \chi^2_4(.05) = 9.49$ we <u>reject</u>

$H_0: \underset{\sim}{\tau}_1 = \underset{\sim}{\tau}_2 = \underset{\sim}{\tau}_3 = \underset{\sim}{0}$ (no factor 1 effects) at the $\alpha = .05$ level.

Since

$$-[gb(n-1)-(p+1-(b-1))/2]\ell n\Lambda^* = -12\ell n\left(\frac{|SSP_{res}|}{|SSP_{fac\,2} + SSP_{res}|}\right)$$

$= -12\ell n(.7949) = 2.76 < \chi^2_6(.05) = 12.59$ we <u>do not reject</u>

$H_0: \underset{\sim}{\beta}_1 = \underset{\sim}{\beta}_2 = \underset{\sim}{\beta}_3 = \underset{\sim}{\beta}_4 = \underset{\sim}{0}$ (no factor 2 effects) at the $\alpha = .05$ level.

6.15 Example 6.11, $g = b = 2$, $n = 5$;

a) For $H_0: \underset{\sim}{\tau}_1 = \underset{\sim}{\tau}_2 = \underset{\sim}{0}$, $\Lambda^* = .3819$

Since

$$-[gb(n-1)-(p+1-(g-1))/2]\ell n\ \Lambda^* = -14.5\ell n(.3819) =$$

$$= 13.96 > \chi_3^2\ (.05) = 7.81\ ,$$

we reject H_0 at $\alpha = .05$ level. For $H_0: \underset{\sim}{\beta}_1 = \underset{\sim}{\beta}_2 = \underset{\sim}{0}$. $\Lambda^* = .5230$ and $-14.5\ell n\ (.5230) = 9.40$. Again we reject H_0 at $\alpha = .05$ level. These results are consistent with the exact F tests.

6.16 $H_0: C\underset{\sim}{\mu} = \underset{\sim}{0}$; $H_1: C\underset{\sim}{\mu} \neq \underset{\sim}{0}$ where $C = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$.

Summary statistics:

$$\bar{\underset{\sim}{x}} = \begin{bmatrix} 1906.1 \\ 1749.5 \\ 1509.1 \\ 1725.0 \end{bmatrix}; \qquad S = \begin{bmatrix} 105625 & 94759 & 87249 & 94268 \\ & 101761 & 76166 & 81193 \\ & & 91809 & 90333 \\ & & & 104329 \end{bmatrix}$$

$$T^2 = n(C\bar{\underset{\sim}{x}})'(CSC')^{-1}(C\bar{\underset{\sim}{x}}) = 254.7$$

$$\frac{(n-1)(q-1)}{(n-q+1)} F_{q-1,n-q+1}(\alpha) = \frac{(30-1)(4-1)}{(30-4+1)} F_{3,27}(.05) = 9.54$$

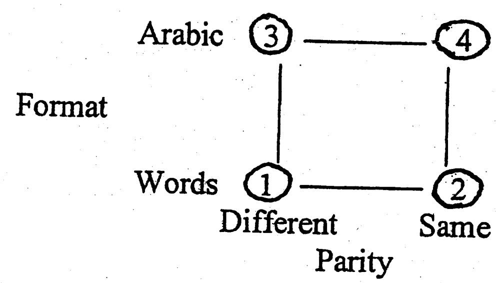
Since $T^2 = 254.7 > 9.54$ we reject H_0 at $\alpha = .05$ level.

95% simultaneous confidence interval for "dynamic" versus "static" means $(\mu_1 + \mu_2) - (\mu_3 + \mu_4)$ is, with $\underset{\sim}{c}' = [1 \quad 1 \quad -1 \quad -1]$,

$$\underset{\sim}{c}'\bar{\underset{\sim}{x}} \pm \sqrt{\frac{(n-1)(q-1)}{(n-q+1)} F_{q-1,n-q+1}(\alpha)}\ \sqrt{\frac{\underset{\sim}{c}'S\underset{\sim}{c}}{n}}$$

$$= 421.5 \pm 174.5 \longrightarrow (247,\ 596)$$

6.17 (a)



<u>Effects</u>	<u>Contrast</u>
Parity main:	$(\mu_2+\mu_4)-(\mu_1+\mu_3)$
Format main:	$(\mu_3+\mu_4)-(\mu_1+\mu_2)$
Interaction:	$(\mu_2+\mu_3)-(\mu_1+\mu_4)$

Contrast matrix:

$$C=\begin{pmatrix}-1 & 1 & -1 & 1\\ -1 & -1 & 1 & 1\\ -1 & 1 & 1 & -1\end{pmatrix}$$

Since $T^2=135.9>\frac{31(3)}{29}(2.93)=9.40$, reject $H_0:\mathbf{C}\mu=\mathbf{0}$ (no treatment effects) at the 5% level.

(b) 95% simultaneous T^2 intervals for the contrasts:

Parity main effect: $-206.4\pm\sqrt{9.40}\sqrt{\frac{20{,}598.6}{32}}\rightarrow(-280.3,\ -125.1)$

Format main effect: $-307\pm\sqrt{9.40}\sqrt{\frac{42{,}939.5}{32}}\rightarrow(-411.4,\ -186.9)$

Interaction effect: $22.4\pm\sqrt{9.40}\sqrt{\frac{9{,}818.5}{32}}\rightarrow(-32.3,\ 75.0)$

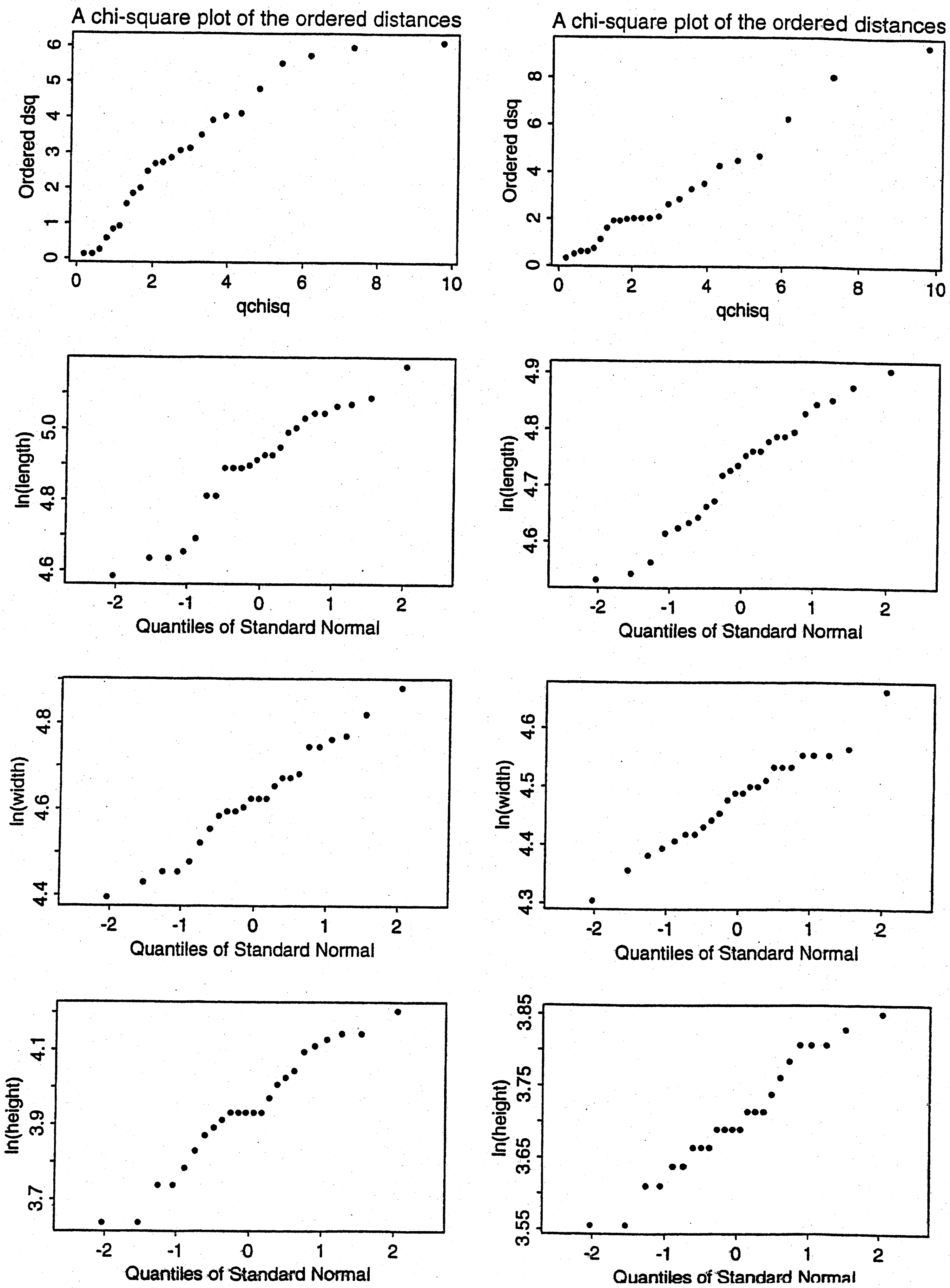
No interaction effect. Parity effect—"different" responses slower than "same" responses. Format effect—"words" slower than "Arabic".

(c) The M model of numerical cognition is a reasonable population model for the scores.

(d) The multivariate normal model is a reasonable model for the scores corresponding to the parity contrast, the format contrast and the interaction contrast.

6.18

Female turtle | Male turtle



mean vector for females:

X1BAR
4.9006593
4.6229089
3.9402858

mean vector for males:

X2BAR
4.7254436
4.4775738
3.7031858

SPOOLED			
	0.0187388	0.0140655	0.0165386
	0.0140655	0.0113036	0.0127148
	0.0165386	0.0127148	0.0158563

TSQ	CVTSQ	F	CVF	PVALUE
85.052001	8.833461	27.118029	2.8164658	4.355E-10

linear combination most responsible for rejection

of HO has coefficient vector:

COEFFVEC
-43.72677
-8.710687
67.546415

95% simultaneous CI for the difference

in female and male means

LOWER	UPPER
0.0577676	0.2926638
0.0541167	0.2365537
0.1290622	0.3451377

Bonferroni CI

LOWER	UPPER
0.0768599	0.2735714
0.0689451	0.2217252
0.1466248	0.3275751

6.19

a) $\bar{\mathbf{x}}_1 = \begin{bmatrix} 12.219 \\ 8.113 \\ 9.590 \end{bmatrix}$; $\bar{\mathbf{x}}_2 = \begin{bmatrix} 10.106 \\ 10.762 \\ 18.168 \end{bmatrix}$;

$$S_1 = \begin{bmatrix} 223.0134 & 12.3664 & 2.9066 \\ & 17.5441 & 4.7731 \\ & & 13.9633 \end{bmatrix}$$

$$S_2 = \begin{bmatrix} 4.3623 & .7599 & 2.3621 \\ & 25.8512 & 7.6857 \\ & & 46.6543 \end{bmatrix};$$

$$S_{pooled} = \begin{bmatrix} 15.8112 & 7.8550 & 2.6959 \\ & 20.7458 & 5.8960 \\ & & 26.5750 \end{bmatrix}$$

$$[(\frac{1}{n_1} + \frac{1}{n_2})S_{pooled}]^{-1} = \begin{bmatrix} 1.0939 & -.4084 & -.0203 \\ & .8745 & -.1525 \\ & & .5640 \end{bmatrix}$$

H_0: $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0}$

Since $T^2 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'[(\frac{1}{n_1} + \frac{1}{n_2})S_{pooled}]^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) = 50.92$

$$> \frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)} F_{p,n_1+n_2-p-1}(.01) = \frac{(57)(3)}{55} F_{3,55}(.01) = 13,$$

we reject H_0 at the $\alpha = .01$ level. There is a difference in the (mean) cost vectors between gasoline trucks and diesel trucks.

b) $\hat{\mathbf{a}} = S^{-1}_{pooled}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) = \begin{bmatrix} 3.58 \\ -1.88 \\ -4.48 \end{bmatrix}$

c) 99% simultaneous confidence intervals are:

$$\mu_{11} - \mu_{21}: \; 2.113 \pm 3.790$$

$$\mu_{12} - \mu_{22}: \; -2.650 \pm 4.341$$

$$\mu_{13} - \mu_{23}: \; -8.578 \pm 4.913$$

d) Assumption $\Sigma_1 = \Sigma_2$.

Since S_1 and S_2 are quite different, it may not be reasonable to pool. However, using "large sample" theory ($n_1 = 36$, $n_2 = 23$) we have, by Result 6.4,

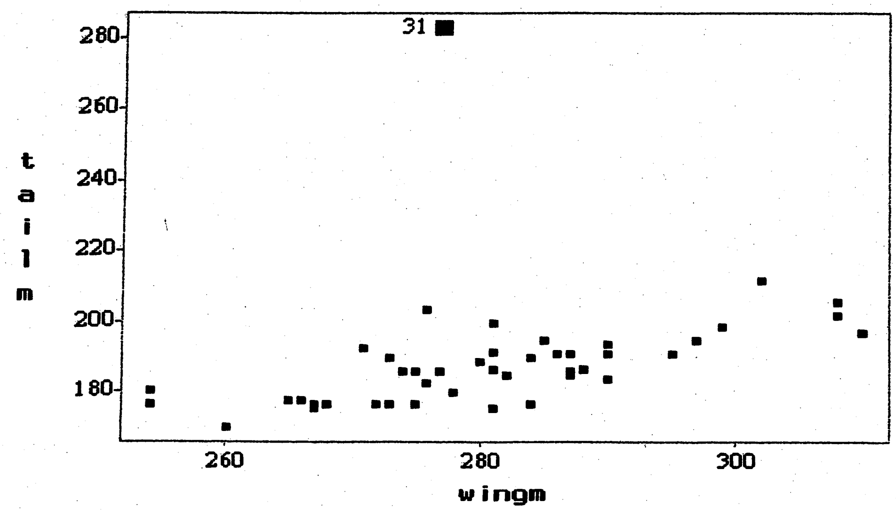
$$(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2))'[\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2]^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)) \sim \chi_p^2$$

Since

$$(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'[\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2]^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) = 43.15 > \chi_3^2(.01) = 11.34$$

we reject $H_0: \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0}$ at the $\alpha = .01$ level. This is consistent with the result in part (a).

6.20 (a)



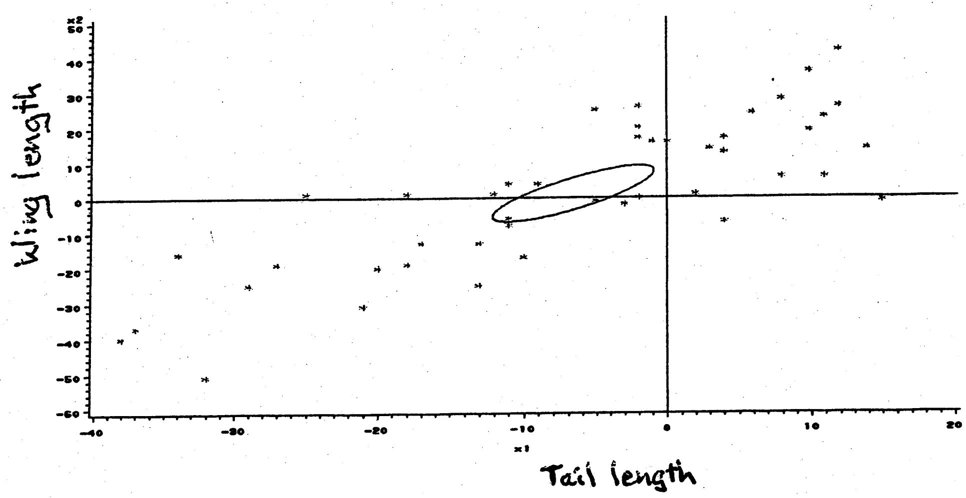
(b) The output below shows that the analysis does not differ when we delete the observation 31 or when we consider it equals 184. Both tests reject the null hypothesis of equal mean difference. The most critical linear combination leading to the rejection of H_0 has coefficient vector $[-3.490238; 2.07955]'$ and the the linear combination most responsible for the rejection of H_0 is the Tail difference.

(c) Results below.

```
Comparing Mean Vectors from Two Populations

              [Obs. 31 Deleted]

          T2          C
     25.005014 5.9914645

     Reject HO. There is mean difference

 95% simultaneous confidence intervals:

 LABELCI              LICIMD    LSCIMD

 Mean Diff. 1:    -11.76436 -1.161905  (Tail difference)
 Mean Diff. 2:    -5.985685 8.3392202  (Wing difference)

 RESULT                    COEF

 Coefficient Vector:    -3.490238
                         2.07955
```

```
Comparing Mean Vectors from Two Populations
                     Tail
             [Obs. 31 = 184]

            T2          C
    25.662531 5.9914645

       Reject H0. There is mean difference

     95% simultaneous confidence intervals:

     LABELCI             LICIMD     LSCIMD

     Mean Diff. 1:    -11.78669  -1.27998
     Mean Diff. 2:    -6.003431 8.1812088

     RESULT                      COEF

     Coefficient Vector:    -3.574268
                             2.1220203
```

95% Confidence Ellipse About the Mean Vector



(d) Female birds are generally larger, since the confidence interval bounds for difference in Tails (Male - Female) are negative and the confidence interval for difference in Wings includes zero, indicating no significance difference.

6.21 (a) The (4,2) and (4,4) entries in S_1 and S_2 differ considerably. However, $n_1 = n_2$ so the large sample approximation amounts to pooling.

(b) H_0: $\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 = \underset{\sim}{0}$ and H_1: $\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 \neq \underset{\sim}{0}$

$$T^2 = 15.830 > \frac{(38)(4)}{35} F_{4,35}(.05) = 11.47$$

so we reject H_0 at the $\alpha = .05$ level.

(c) $$\hat{\underset{\sim}{a}} \propto S^{-1}_{\text{pooled}}(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2) = \begin{bmatrix} -.24 \\ .16 \\ -3.74 \\ .01 \end{bmatrix}$$

(d) Looking at the coefficients $\hat{a}_i\sqrt{s_{ii,\text{pooled}}}$, which apply to the standardized variables, we see that X_2: long term interest rate has the largest coefficient and therefore might be useful in classifying a bond as "high" or "medium" quality.

(e) From (b), $T^2 = 15.830$. Have $p = 4$ and $\nu = \dfrac{4+16}{.53556} = 37.344$ so, at the 5% level, the critical value is

$$\frac{\nu p}{\nu - p + 1} F_{p,\nu-p+1}(.05) = \frac{37.344(4)}{37.344-4+1} F_{4,37.344-4+1}(.05) = \frac{149.376}{34.344}(2.647) = 11.513$$

Since $T^2 = 15.830 > 11.513$, reject $H_0 : \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0}$, the same conclusion reached in (b). Notice the critical value here is only slightly larger than the critical value in (b).

6.22 (a) The sample means for female and male are :

$$\bar{x}_F = \begin{bmatrix} 0.3136 \\ 5.1788 \\ 2.3152 \\ 38.1548 \end{bmatrix}, \quad \bar{x}_M = \begin{bmatrix} 0.3972 \\ 5.3296 \\ 3.6876 \\ 49.3404 \end{bmatrix}.$$

The Hotelling's $T^2 = 96.487 > 11.00$ where 11.00 is a critical point corresponding to $\alpha = 0.05$. Therefore, we reject $H_0 : \mu_1 - \mu_2 = 0$. The coefficient of the linear combination of most responsible for rejection is $(-95.600, 6.145, 5.737, -0.762)'$.

(b) The 95% simultaneous C. I. for female mean $-$male mean:

$$\begin{bmatrix} -0.1697234, & 0.00252336 \\ -1.4650835, & 1.16348346 \\ -1.8760572, & -0.8687428 \\ -17.032834, & -5.3383659 \end{bmatrix}$$

(c) We cannot extend the obtained result to the population of persons in their mid-twenties. Firstly this was a self selected sample of volunteers (friends) and is not even a random sample of graduate students. Further, graduate students are probably more sedentary than the typical persons of their age.

6.23 $n_1 = n_2 = n_3 = 50$; $p = 2$, $g = 3$ (sepal width and petal width responses only!)

$$\underset{\sim}{\bar{x}}_1 = \begin{bmatrix} 3.428 \\ .306 \end{bmatrix}; \quad S_1 = \begin{bmatrix} .14364 & -.00474 \\ & .18576 \end{bmatrix}$$

$$\underset{\sim}{\bar{x}}_2 = \begin{bmatrix} 2.770 \\ 1.326 \end{bmatrix}; \quad S_2 = \begin{bmatrix} .09860 & .04128 \\ & .03920 \end{bmatrix}$$

$$\underset{\sim}{\bar{x}}_3 = \begin{bmatrix} 2.974 \\ 2.026 \end{bmatrix}; \quad S_3 = \begin{bmatrix} .10368 & .04764 \\ & .07563 \end{bmatrix}$$

MANOVA Table:

Source	SSP	d.f.
Treatment	$B = \begin{bmatrix} 11.344 & -21.820 \\ & 75.352 \end{bmatrix}$	2
Residual	$W = \begin{bmatrix} 16.950 & 4.125 \\ & 14.729 \end{bmatrix}$	147
Total	$B+W = \begin{bmatrix} 28.294 & -17.695 \\ & 90.081 \end{bmatrix}$	149

$$\Lambda^* = \frac{|W|}{|B+W|} = \frac{232.64}{2235.64} = .104$$

Since $\left(\frac{\Sigma n_\ell - p - 2}{p}\right)\left(\frac{1 - \sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) = 153.3 > 2.37 \doteq F_{4,292}(.05)$

we reject H_0: $\underset{\sim}{\tau}_1 = \underset{\sim}{\tau}_2 = \underset{\sim}{\tau}_3$ at the $\alpha = .05$ level.

6.24 Wilks' lambda: $\Lambda^* = .8301$. Since $g = 3$, $\left(\frac{90-4-2}{4}\right)\left(\frac{1-\sqrt{.8301}}{\sqrt{.8301}}\right) = 2.049$ is an F value with 8 and 168 degrees of freedom. Since p-value $= P(F > 2.049) = .044$, we would just reject the null hypothesis $H_0 : \underset{\sim}{\tau}_1 = \underset{\sim}{\tau}_2 = \underset{\sim}{\tau}_3 = \underset{\sim}{0}$ at the 5% level implying there is a time period effect.

F statistics and p-values for ANOVA's:

	F	p-value
MaxBrth:	3.66	.030
BasHght:	0.47	.629
BasLgth:	3.84	.025
NasHght:	0.10	.901

Any differences over time periods are probably due to changes in maximum breath of skull (MaxBrth) and basialveolar length of skull (BasLgth).

95% Bonferroni simultaneous intervals: $m = pg(g-1)/2 = 12$, $t_{87}(.05/24) = 2.94$

BasBrth

$$\tau_{11} - \tau_{21} : \quad -1 \pm 2.94\sqrt{\frac{1785.4}{87}\left(\frac{1}{30}+\frac{1}{30}\right)} \rightarrow -1 \pm 3.44$$

$$\tau_{11} - \tau_{31} : \quad -3.1 \pm 3.44$$

$$\tau_{21} - \tau_{31} : \quad -2.1 \pm 3.44$$

BasHght

$$\tau_{12} - \tau_{22} : \quad 0.9 \pm 2.94\sqrt{\frac{1924.3}{87}\left(\frac{1}{30}+\frac{1}{30}\right)} \rightarrow 0.9 \pm 3.57$$

$$\tau_{12} - \tau_{32} : \quad -0.2 \pm 3.57$$

$$\tau_{22} - \tau_{32} : \quad -1.1 \pm 3.57$$

BasLgth

$$\tau_{13} - \tau_{23} : \quad 0.10 \pm 2.94\sqrt{\frac{2153}{87}\left(\frac{1}{30}+\frac{1}{30}\right)} \rightarrow 0.10 \pm 3.78$$

$$\tau_{13} - \tau_{33} : \quad 3.14 \pm 3.78$$

$$\tau_{23} - \tau_{33} : \quad 3.03 \pm 3.78$$

NasHgth

$$\tau_{14} - \tau_{24} : \quad 0.30 \pm 2.94\sqrt{\frac{840.2}{87}\left(\frac{1}{30}+\frac{1}{30}\right)} \rightarrow 0.30 \pm 2.36$$

$$\tau_{14} - \tau_{34} : \quad -0.03 \pm 2.36$$

$$\tau_{24} - \tau_{34} : \quad -0.33 \pm 2.36$$

All the simultaneous intervals include 0. Evidence for changes in skull size over time is marginal. If changes exist, then these changes might be in maximum breath and basialveolar length of skull from time periods 1 to 3.

The usual MANOVA assumptions appear to be satisfied for these data.

6.25

Without transforming the data, $\Lambda^* = \frac{|W|}{|B + W|} = .1159$ and F = 18.98.

After transformation, $\Lambda^* = .1198$ and F = 18.52 $> F_{10,98}(.05) = 1.93$

There is a clear need for transforming the data to make the hypothesis tenable.

6.26 To test for parallelism, consider H_{01}: $C\underset{\sim}{\mu}_1 = C\underset{\sim}{\mu}_2$ with C given by (6-61).

$$C(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2) = \begin{bmatrix} -.413 \\ -.167 \\ -.036 \end{bmatrix}; \quad (CS_{pooles}C')^{-1} = \begin{bmatrix} 1.674 & .947 & .616 \\ & 2.014 & 1.144 \\ & & 2.341 \end{bmatrix}$$

$T^2 = 9.58 > c^2 = 8.0$, we reject H_0 at the $\alpha = .05$ level. The excess electrical usage of the test group was much lower than that of the control group for the 11 A.M., 1 P.M. and 3 P.M. hours. The similar 9 A.M. usage for the two groups contradicts the parallelism hypothesis.

6.27 a) Plots of the husband and wife profiles look similar but seem disparate for the level of "companionate love that you feel for your partner".

b) Parallelism hypothesis H_0: $C\underset{\sim}{\mu}_1 = C\underset{\sim}{\mu}_2$ with C given by (6-61).

$$C(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2) = \begin{bmatrix} -.13 \\ -.17 \\ .33 \end{bmatrix}; \quad CS_{pooled}C' = \begin{bmatrix} .685 & .733 & .029 \\ & .870 & -.028 \\ & & .095 \end{bmatrix}$$

For $\alpha = .05$, $c^2 = 8.7$ (see (6-62)). Since

$T^2 = 19.58 > c^2 = 8.7$ we reject H_0 at the $\alpha = .05$ level.

6.28 $T^2 = 106.13 > 16.59$. We reject $H_0 : \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 = \underset{\sim}{0}$ at 5% significance level. There is a significant difference in the two species.

Sample Mean for L.torrens and L.carteri:

L.torrens	L.carteri	Difference
96.457	99.343	-2.886
42.914	43.743	-0.829
35.371	39.314	-3.943
14.514	14.657	-0.143
25.629	30.000	-4.371
9.571	9.657	-0.086
9.714	9.371	0.343

Pooled Sample Covariance Matrix:

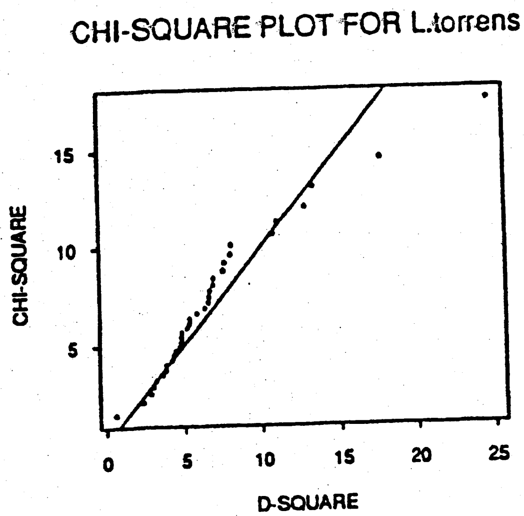
36.008	14.595	6.078	3.675	9.573	2.426	2.649
	16.639	2.764	2.992	6.101	1.053	0.934
		6.437	0.692	1.615	0.211	0.671
			3.039	2.407	0.274	0.229
				13.767	0.565	0.637
					1.213	0.914
						0.990

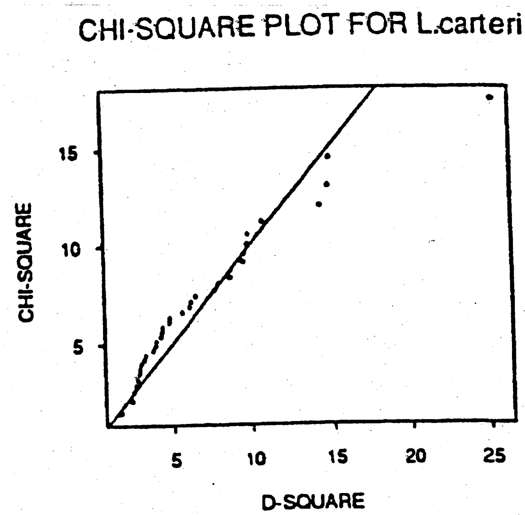
Linear Combination of most responsible for rejection
of Ho: L.torrens mean - L.carteri mean = 0 is :
(0.006, 0.151, -0.854, 0.268, -0.383, -2.187, 2.971)'

95% Simultaneous C. I. for L.torrens mean - L.carteri mean:

LOWER	UPPER
-8.73	2.96
-4.80	3.14
-6.41	-1.47
-1.84	1.55
-7.98	-0.76
-1.16	0.99
-0.63	1.31

The third and fifth components are most responsible for rejecting H_0. The χ^2 plots look fairy straight.





6.29

(a).

	XBAR	S		
Summary Statistics:	0.02548	0.00366259	0.00482862	0.00154159
	0.05784	0.00482862	0.01628931	0.00304801
	0.01056	0.00154159	0.00304801	0.00602526

Hotelling's $T^2 = 5.946$. The critical point is 9.979 and we fail to reject $H_0 : \mu_1 - \mu_2 = 0$ at 5% significance level.

(b). (c).

	LOWER	UPPER
Bonferroni C. I.:	-0.0057	0.0566
	-0.0079	0.1235
	-0.0294	0.0505
Simultaneous C. I.:	-0.0128	0.0637
	-0.0228	0.1385
	-0.0385	0.0596

6.30

```
HOTELLING T SQUARE   =   9.0218
P-VALUE    0.3616
```

	N	MEAN	STDEV	T2 INTERVAL TO		BONFERRONI TO	
x1	24	0.00012	0.04817	-.0443	.0445	-.0283	.0285
x2	24	-0.00325	0.02751	-.0286	.0221	-.0195	.0130
x3	24	-0.0072	0.1030	-.1020	.0876	-.0679	.0535
x4	24	-0.0123	0.0625	-.0701	.0455	-.0493	.0247
x5	24	0.01513	0.03074	-.0130	.0436	-.0030	.0333
x6	24	0.00017	0.04689	-.0430	.0434	-.0275	.0278

The Bonferroni intervals use t (.00417) = 2.89 and

the T intevals use the constant 4.516.

6.31 **(a)** Two-factor MANOVA of peanuts data

E = Error SS&CP Matrix

	X1	X2	X3
X1	104.205	49.365	76.48
X2	49.365	352.105	121.995
X3	76.48	121.995	94.835

H = Type III SS&CP Matrix for FACTOR1 (Location)

	X1	X2	X3
X1	0.7008333333	-10.6575	7.1291666667
X2	-10.6575	162.0675	-108.4125
X3	7.1291666667	-108.4125	72.520833333

Manova Test Criteria and Exact F Statistics for
the Hypothesis of no Overall FACTOR1 Effect
H = Type III SS&CP Matrix for FACTOR1 E = Error SS&CP Matrix

S=1 M=0.5 N=1

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.10651620	11.1843	3	4	0.0205
Pillai's Trace	0.89348380	11.1843	3	4	0.0205
Hotelling-Lawley Trace	8.38824348	11.1843	3	4	0.0205
Roy's Greatest Root	8.38824348	11.1843	3	4	0.0205

H = Type III SS&CP Matrix for FACTOR2 (Variety)

	X1	X2	X3
X1	196.115	365.1825	42.6275
X2	365.1825	1089.015	414.655
X3	42.6275	414.655	284.10166667

Manova Test Criteria and F Approximations for
the Hypothesis of no Overall FACTOR2 Effect
H = Type III SS&CP Matrix for FACTOR2 E = Error SS&CP Matrix

S=2 M=0 N=1

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.01244417	10.6191	6	8	0.0019
Pillai's Trace	1.70910921	9.7924	6	10	0.0011
Hotelling-Lawley Trace	21.37567504	10.6878	6	6	0.0055
Roy's Greatest Root	18.18761127	30.3127	3	5	0.0012

H = Type III SS&CP Matrix for FACTOR1*FACTOR2

	X1	X2	X3
X1	205.10166667	363.6675	107.78583333
X2	363.6675	780.695	254.22
X3	107.78583333	254.22	85.951666667

Manova Test Criteria and F Approximations for
the Hypothesis of no Overall FACTOR1*FACTOR2 Effect
H = Type III SS&CP Matrix for FACTOR1*FACTOR2   E = Error SS&CP Matrix

S=2   M=0   N=1

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.07429984	3.5582	6	8	0.0508
Pillai's Trace	1.29086073	3.0339	6	10	0.0587
Hotelling-Lawley Trace	7.54429038	3.7721	6	6	0.0655
Roy's Greatest Root	6.82409388	11.3735	3	5	0.0113

(b) The residuals for X_2 at location 2 for variety 5 seem large in absolute value, but Q-Q plots of residuals indicate that univariate normality cannot be rejected for all three variables.

CODE	FACTOR1	FACTOR2	PRED1	RES1	PRED2	RES2	PRED3	RES3
a	1	5	194.80	0.50	160.40	-7.30	52.55	-1.15
a	1	5	194.80	-0.50	160.40	7.30	52.55	1.15
b	2	5	185.05	4.65	130.30	9.20	49.95	5.55
b	2	5	185.05	-4.65	130.30	-9.20	49.95	-5.55
c	1	6	199.45	3.55	161.40	-4.60	47.80	2.00
c	1	6	199.45	-3.55	161.40	4.60	47.80	-2.00
d	2	6	200.15	2.55	163.95	2.15	57.25	3.15
d	2	6	200.15	-2.55	163.95	-2.15	57.25	-3.15
e	1	8	190.25	3.25	164.80	-0.30	58.20	-0.40
e	1	8	190.25	-3.25	164.80	0.30	58.20	0.40
f	2	8	200.75	0.75	170.30	-3.50	66.10	-1.10
f	2	8	200.75	-0.75	170.30	3.50	66.10	1.10

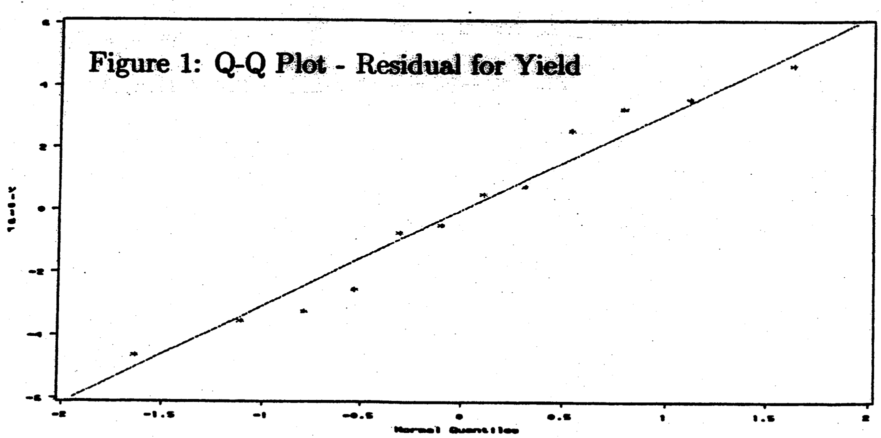


Figure 1: Q-Q Plot - Residual for Yield

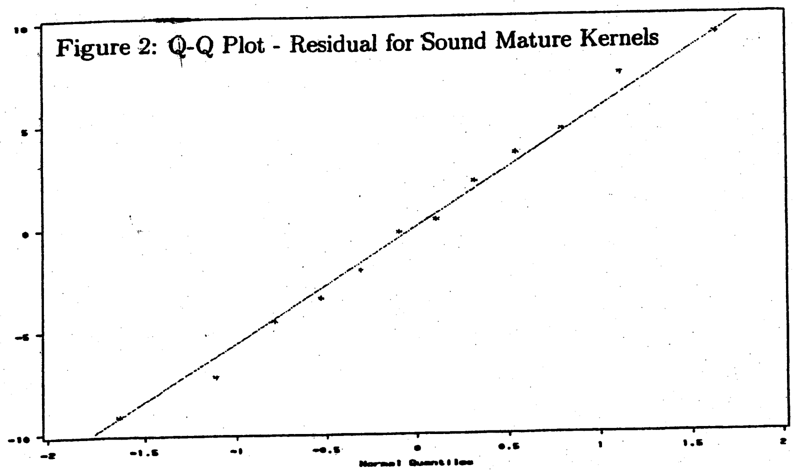


Figure 2: Q-Q Plot - Residual for Sound Mature Kernels

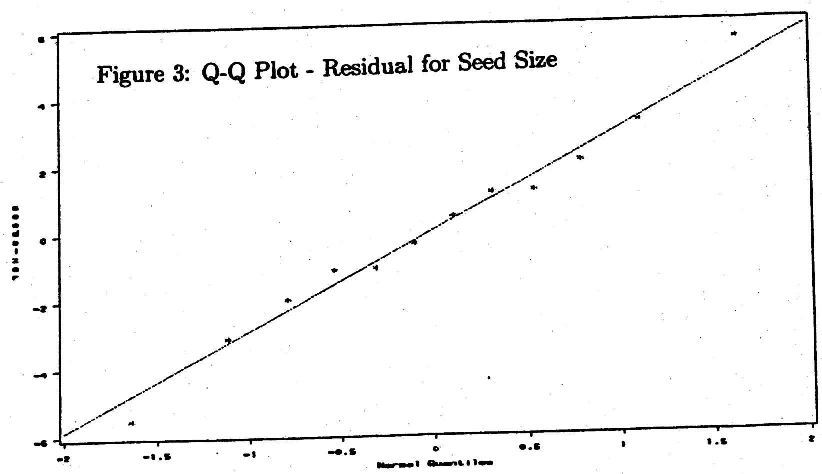


Figure 3: Q-Q Plot - Residual for Seed Size

(c) Univariate two factor ANOVAs follow. Evidence of variety effect and, for X_1 = yield and X_2 = sound mature kernel, a location*variety interaction.

Dependent Variable: yield

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	401.9175000	80.3835000	4.63	0.0446
Error	6	104.2050000	17.3675000		
Corrected Total	11	506.1225000			

R-Square	Coeff Var	Root MSE	yield Mean
0.794111	2.136324	4.167433	195.0750

Source	DF	Type III SS	Mean Square	F Value	Pr > F
location	1	0.7008333	0.7008333	0.04	0.8474
variety	2	196.1150000	98.0575000	5.65	0.0418
location*variety	2	205.1016667	102.5508333	5.90	0.0382

Dependent Variable: sdmatker

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2031.777500	406.355500	6.92	0.0177
Error	6	352.105000	58.684167		
Corrected Total	11	2383.882500			

R-Square	Coeff Var	Root MSE	sdmatker Mean
0.852298	4.832398	7.660559	158.5250

Source	DF	Type III SS	Mean Square	F Value	Pr > F
location	1	162.067500	162.067500	2.76	0.1476
variety	2	1089.015000	544.507500	9.28	0.0146
location*variety	2	780.695000	390.347500	6.65	0.0300

The GLM Procedure

Dependent Variable: seedsize

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	442.5741667	88.5148333	5.60	0.0292
Error	6	94.8350000	15.8058333		
Corrected Total	11	537.4091667			

R-Square	Coeff Var	Root MSE	seedsize Mean
0.823533	7.188166	3.975655	55.30833

Source	DF	Type III SS	Mean Square	F Value	Pr > F
location	1	72.5208333	72.5208333	4.59	0.0759
variety	2	284.1016667	142.0508333	8.99	0.0157
location*variety	2	85.9516667	42.9758333	2.72	0.1443

(d) Bonferroni simultaneous comparisons of variety.
Only varieties 5 and 8 differ, and they differ only on X_3.

```
Bonferroni (Dunn) T tests for variable: X1
Alpha= 0.05  Confidence= 0.95  df= 8  MSE= 38.66333
Critical Value of T= 3.01576
Minimum Significant Difference= 13.26
Comparisons significant at the 0.05 level are indicated by '***'.
```

FACTOR2 Comparison	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit	
6 - 8	-8.960	4.300	17.560	
6 - 5	-3.385	9.875	23.135	
8 - 6	-17.560	-4.300	8.960	
8 - 5	-7.685	5.575	18.835	
5 - 6	-23.135	-9.875	3.385	
5 - 8	-18.835	-5.575	7.685	

```
Bonferroni (Dunn) T tests for variable: X2
Alpha= 0.05  Confidence= 0.95  df= 8  MSE= 141.6
Critical Value of T= 3.01576
Minimum Significant Difference= 25.375
Comparisons significant at the 0.05 level are indicated by '***'.
```

FACTOR2 Comparison	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit	
8 - 6	-20.500	4.875	30.250	
8 - 5	-3.175	22.200	47.575	
6 - 8	-30.250	-4.875	20.500	
6 - 5	-8.050	17.325	42.700	
5 - 8	-47.575	-22.200	3.175	
5 - 6	-42.700	-17.325	8.050	

```
Bonferroni (Dunn) T tests for variable: X3
Alpha= 0.05  Confidence= 0.95  df= 8  MSE= 22.59833
Critical Value of T= 3.01576
Minimum Significant Difference= 10.137
Comparisons significant at the 0.05 level are indicated by '***'.
```

FACTOR2 Comparison	Simultaneous Lower Confidence Limit	Difference Between Means	Simultaneous Upper Confidence Limit	
8 - 6	-0.512	9.625	19.762	
8 - 5	0.763	10.900	21.037	***
6 - 8	-19.762	-9.625	0.512	
6 - 5	-8.862	1.275	11.412	
5 - 8	-21.037	-10.900	-0.763	***
5 - 6	-11.412	-1.275	8.862	

6.32 (a) MANOVA for Species: Wilks' lambda $\Lambda_1^* = .00823$

$F = 5.011$; $p\text{-value} = P(F > 5.011) = .173$

$F_{4,2}(.05) = 19.25$

Do not reject H_0: No species effects

MANOVA for Nutrient: Wilks' lambda $\Lambda_2^* = .31599$

$F = 1.082$; $p\text{-value} = P(F > 1.082) = .562$

$F_{2,1}(.05) = 199.5$

Do not reject H_0: No nutrient effects

(b) Minitab output for the two-way ANOVA's:

560CM

Analysis of Variance for 560CM

Source	DF	SS	MS	F	P
Spec	2	47.476	23.738	10.06	0.090
Nutrient	1	8.260	8.260	3.50	0.202
Error	2	4.722	2.361		
Total	5	60.458			

720CM

Analysis of Variance for 720CM

Source	DF	SS	MS	F	P
Spec	2	262.239	131.119	28.82	0.034
Nutrient	1	4.489	4.489	0.99	0.425
Error	2	9.099	4.550		
Total	5	275.827			

The ANOVA results are mostly consistent with the MANOVA results. The exception is for 720CM where there appears to be Species effects. A look at the data suggests the spectral reflectance of Japanese larch (JL) at 720 nanometers is somewhat larger than the reflectance of the other two species (SS and LP) regardless of nutrient level. This difference is not as apparent at 560 nanometers.

For MANOVA, the value of Wilks' lambda statistic does not indicate Species effects. However, Pillai's trace statistic, 1.6776 with $F = 5.203$ and p-value = .07, suggests there may be Species effects. (For Nutrient, Wilks' lambda and Pillai's trace statistic give the same F value.) For larger sample sizes, Wilks' lambda and Pillai's trace statistic would give essentially the same result for all factors.

6.33 (a) MANOVA for Species: Wilks' lambda $\Lambda_1^* = .06877$

$F = 36.571$; $p\text{-value} = P(F > 36.571) = .000$

$F_{4,52}(.05) = 2.55$

Reject H_0: No species effects

MANOVA for Time: Wilks' lambda $\Lambda_2^* = .04917$

$F = 45.629$; $p\text{-value} = P(F > 45.629) = .000$

$F_{4,52}(.05) = 2.55$

Reject H_0: No time effects

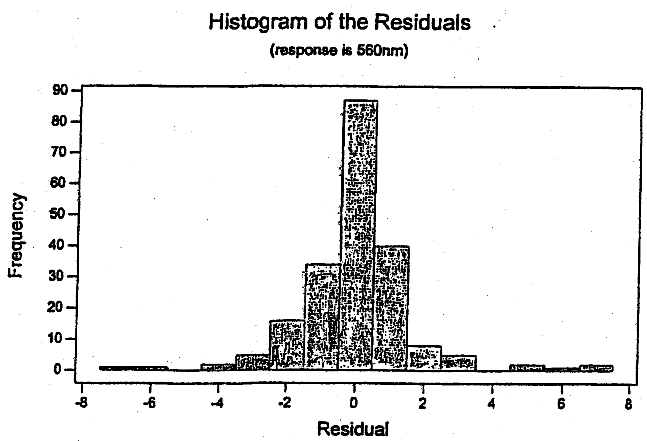
MANOVA for Species*Time: Wilks' lambda $\Lambda_{12}^* = .08707$

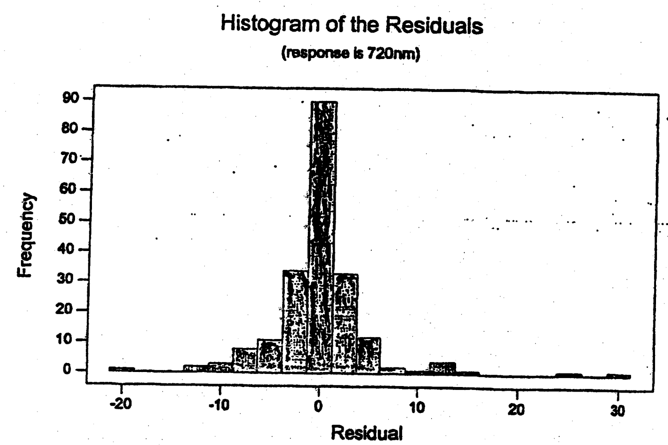
$F = 15.528$; $p\text{-value} = P(F > 15.528) = .000$

$F_{8,52}(.05) = 2.12$

Reject H_0: No interaction effects

(b) A few outliers but, in general, residuals approximately normally distributed (see histograms below). Observations are likely to be positively correlated over time. Observations are not independent.





(c) Interaction shows up for the 560nm wavelength but not for the 720nm wavelength. See the Minitab ANOVA output below.

Analysis of Variance for 560nm

Source	DF	SS	MS	F	P
Species	2	965.18	482.59	169.97	0.000
Time	2	1275.25	637.62	224.58	0.000
Species*Time	4	795.81	198.95	70.07	0.000
Error	27	76.66	2.84		
Total	35	3112.90			

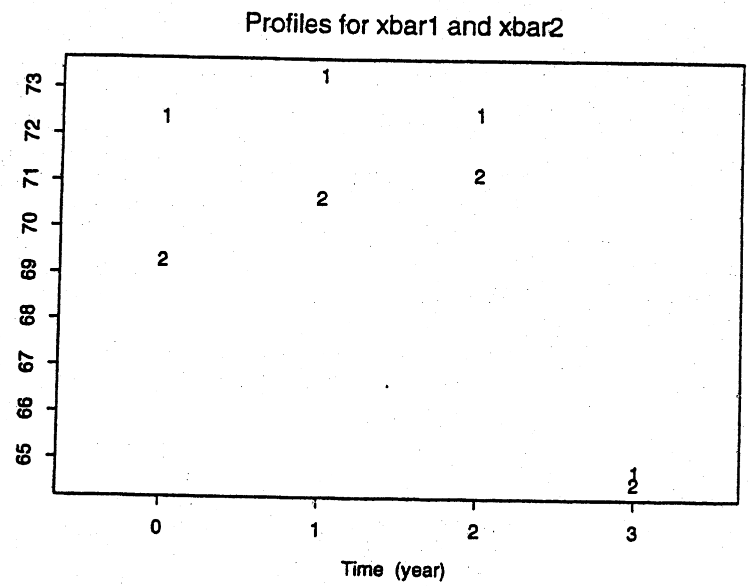
Analysis of Variance for 720nm

Source	DF	SS	MS	F	P
Species	2	2026.86	1013.43	15.46	0.000
Time	2	5573.81	2786.90	42.52	0.000
Species*Time	4	193.55	48.39	0.74	0.574
Error	27	1769.64	65.54		
Total	35	9563.85			

6.33. (Continued)

(d) The data might be analyzed using the growth curve methodology discussed in Section 6.9. The data might also be analyzed assuming species are "nested" within date. In this case, an interesting question is: Is spectral reflectance the same for all species for each date?

6.34 Fitting a linear growth curve to calcium measurements on the dominant ulna



XBAR

72.3800	69.2875
73.2933	70.6562
72.4733	71.1812
64.7867	64.5312

Grand mean

71.1939
71.8273
72.1848
65.2667

MLE of beta

73.4707	70.5049
-1.9035	-0.9818

[B'Sp^(-1)B]^(-1)

93.1313	-5.2393
-5.2393	1.2948

S1

92.1189	86.1106	73.3623	74.5890
86.1106	89.0764	72.9555	71.7728
73.3623	72.9555	71.8907	63.5918
74.5890	71.7728	63.5918	75.4441

S2

98.1745	97.0134	89.4824	86.1111
97.0134	100.5960	88.1425	88.2095
89.4824	88.1425	86.3496	80.5506
86.1111	88.2095	80.5506	81.4156

Spooled

95.2511	91.7500	81.7003	80.5487
91.7500	95.0348	80.8108	80.2745
81.7003	80.8108	79.3694	72.3636
80.5487	80.2745	72.3636	78.5328

W = (N-g)*Spooled

2762.282	2660.749	2369.308	2335.912
2660.749	2756.009	2343.514	2327.961
2369.308	2343.514	2301.714	2098.544
2335.912	2327.961	2098.544	2277.452

Estimated covariance matrix

7.1816	-0.4040	0.0000	0.0000
-0.4040	0.0998	0.0000	0.0000
0.0000	0.0000	6.7328	-0.3788
0.0000	0.0000	-0.3788	0.0936

W1

2803.839	2610.438	2271.920	2443.549
2610.438	2821.243	2464.120	2196.065
2271.920	2464.120	2531.625	1845.313
2443.549	2196.065	1845.313	2556.818

Lambda = |W|/|W1| = 0.201

Since, with $\alpha = 0.01$, $-\left[N - \frac{1}{2}(p - q + g)\right]\log(\Lambda) = 45.72 > \chi^2_{(4-1-1)2}(0.01) = 13.28$, we reject the null hypothesis of a linear fit at $\alpha = 0.01$.

6.35 Fitting a quadratic growth curve to calcium measurements on the dominant ulna, treating all 31 subjects as a single group.

```
XBAR                MLE of beta         [B'Sp^(-1)B]^(-1)
70.7839             71.6039             92.2789 -5.9783  0.0799
71.9323              3.8673             -5.9783  9.3020 -2.9033
71.8065             -1.9404              0.0799 -2.9033  1.0760
64.6548

S                                       W = (n-1)*S
94.5441 90.7962 80.0081 78.0676         2836.322 2723.886 2400.243 2342.027
90.7962 93.6616 78.9965 77.7725         2723.886 2809.848 2369.894 2333.175
80.0081 78.9965 77.1546 70.0366         2400.243 2369.894 2314.639 2101.099
78.0676 77.7725 70.0366 75.9319         2342.027 2333.175 2101.099 2277.957

Estimated covariance matrix             W2
 3.1894 -0.2066  0.0028                 2857.167 2764.522 2394.410 2369.674
-0.2066  0.3215 -0.1003                 2764.522 2889.063 2358.522 2387.070
 0.0028 -0.1003  0.0372                 2394.410 2358.522 2316.271 2093.362
                                        2369.674 2387.070 2093.362 2314.625

Lambda = |W|/|W2| = 0.7653
```

Since, with $\alpha = 0.01$, $-\left[n - \frac{1}{2}(p - q + 1)\right] \log(\Lambda) = 7.893 > \chi^2_{4-2-1}(0.01) = 6.635$, we reject the null hypothesis of a quadratic fit at $\alpha = 0.01$.

6.36 Here

$$p = 2, n_1 = 45, n_2 = 55, \ln|S_1| = 19.90948, \ln|S_2| = 18.40324, \ln|S_{pooled}| = 19.27712$$

so $$u = \left[\frac{1}{44} + \frac{1}{54} - \frac{1}{44+54}\right]\left[\frac{2(4)+3(2)-1}{6(2+1)(2-1)}\right] = .02242$$

and
$$C = (1 - .02242)(98(19.27712) - 44(19.90948) - 54(18.40324)) = 18.93$$

The chi-square degrees of freedom $\nu = \frac{1}{2}2(3)(1) = 3$ and $\chi^2_3(.05) = 7.81$. Since $C = 18.93 > \chi^2_3(.05) = 7.83$, reject $H_0 : \Sigma_1 = \Sigma_2 = \Sigma$ at the 5% level.

6.37 Here

$$p=3, n_1=24, n_2=24, \ln|S_1|=9.48091, \ln|S_2|=6.67870, \ln|S_{pooled}|=8.62718$$

so $$u=\left[\frac{1}{23}+\frac{1}{23}-\frac{1}{23+23}\right]\left[\frac{2(9)+3(3)-1}{6(3+1)(2-1)}\right]=.07065$$

and

$$C=(1-.07065)(46(8.62718)-23(9.48091)-23(6.67870))=23.40$$

The chi-square degrees of freedom $\nu=\frac{1}{2}3(4)(1)=6$ and $\chi_6^2(.05)=12.59$. Since $C=23.40>\chi_6^2(.05)=12.59$, reject $H_0:\Sigma_1=\Sigma_2=\Sigma$ at the 5% level.

6.38 Working with the transformed data, X_1 = vanadium, $X_2=\sqrt{\text{iron}}$, $X_3=\sqrt{\text{beryllium}}$, $X_4=1/[\text{saturated hydrocarbons}]$, X_5 = aromatic hydrocarbons, we have $p=5, n_1=7, n_2=11, n_3=38, \ln|S_1|=-17.81620, \ln|S_2|=-7.24900,$ $\ln|S_3|=-7.09274, \ln|S_{pooled}|=-7.11438$

so $$u=\left[\frac{1}{6}+\frac{1}{10}+\frac{1}{37}-\frac{1}{6+10+37}\right]\left[\frac{2(25)+3(5)-1}{6(5+1)(3-1)}\right]=.24429$$

and

$$C=(1-.24429)(53(-7.11438)-6(-17.81620)-10(-7.24900)-37(-7.09274))=48.94$$

The chi-square degrees of freedom $\nu=\frac{1}{2}5(6)(2)=30$ and $\chi_{30}^2(.05)=43.77$. Since $C=48.94>\chi_{30}^2(.05)=43.77$, reject $H_0:\Sigma_1=\Sigma_2=\Sigma_3=\Sigma$ at the 5% level.

6.39 **(a)** Following Example 6.5, we have $(\bar{\mathbf{x}}_F-\bar{\mathbf{x}}_M)'=(119.55, 29.97)$,

$$\left[\frac{1}{28}\mathbf{S}_F+\frac{1}{28}\mathbf{S}_M\right]^{-1}=\begin{bmatrix} .033186 & -.108533 \\ -.108533 & .423508 \end{bmatrix}$$ and $T^2=76.97$. Since $T^2=76.97>\chi_2^2(.05)=5.99$, we reject $H_0:\mu_F-\mu_M=\mathbf{0}$ at the 5% level.

(b) With equal sample sizes, the large sample procedure is essentially the same as the procedure based on the pooled covariance matrix.

(c) Here p=2, $t_{54}(.05/2(2))\approx z(.0125)=2.24$, $\left[\frac{1}{28}\mathbf{S}_F+\frac{1}{28}\mathbf{S}_M\right]=\begin{bmatrix} 186.148 & 47.705 \\ 47.705 & 14.587 \end{bmatrix}$, so

$\mu_{F1}-\mu_{M1}$: $119.55\pm 2.24\sqrt{186.148}\rightarrow(88.99,\ 150.11)$

$\mu_{F2}-\mu_{M2}$: $29.97\pm 2.24\sqrt{14.587}\rightarrow(21.41,\ 38.52)$

Female Anacondas are considerably longer and heavier than males.

6.41 Three factors: (Problem) Severity, (Problem) Complexity and (Engineer) Experience, each at two levels. Two responses: Assessment time, Implementation time. MANOVA results for significant (at the 5% level) effects.

Effect	Wilks' lambda	F	P-value
Severity	.06398	73.1	.000
Complexity	.01852	265.0	.000
Experience	.03694	130.4	.000
Severity*Complexity	.33521	9.9	.004

Individual ANOVA's for each of the two responses, Assessment time and Implementation time, show only the same three main effects and two factor interaction as significant with p-values for the appropriate F statistics less than .01 in all cases. We see that both assessment time and implementation time is affected by problem severity, problem complexity and engineer experience as well as the interaction between severity and complexity. Because of the interaction effect, the main effects severity and complexity are not additive and do not have a clear interpretation. For this reason, we do not calculate simultaneous confidence intervals for the magnitudes of the mean differences in times across the two levels of each of these main effects. There is no interaction term associated with experience however. Since there are only two levels of experience, we can calculate ordinary t intervals for the mean difference in assessment time and the mean difference in implementation time for gurus (G) and novices (N). Relevant summary statistics and calculations are given below.

$$\text{Error sum of squares and crossproducts matrix} = \begin{bmatrix} 2.222 & 1.217 \\ 1.217 & 2.667 \end{bmatrix}$$

Error deg. of freedom: 11

Assessment time: $\bar{x}_G = 3.68, \bar{x}_N = 5.39$

95% confidence interval for mean difference in experience:

$$3.68 - 5.39 \pm 2.201\sqrt{\frac{2.222}{11}\frac{2}{8}} = -1.71 \pm .49 \rightarrow (-2.20, -1.22)$$

Implementation time: $\bar{x}_G = 6.80, \bar{x}_N = 10.96$

95% confidence interval for mean difference in experience:

$$6.80 - 10.96 \pm 2.201\sqrt{\frac{2.667}{11}\frac{2}{8}} = -4.16 \pm .54 \rightarrow (-4.70, -3.62)$$

The decrease in mean assessment time for gurus relative to novices is estimated to be between 1.22 and 2.20 hours. Similarly the decrease in mean implementation time for gurus relative to novices is estimated to be between 3.62 and 4.70 hours.