1. (a) 
$$\overline{\mathbf{X}} = \begin{bmatrix} 37.6\\ 62.5\\ 5.8 \end{bmatrix}$$
  $\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 97.84 & 92.70 & 14.92\\ 92.70 & 275.05 & 27.80\\ 14.92 & 27.80 & 4.96 \end{bmatrix}$ 

(b) 
$$S = \frac{10}{9}\Sigma = \begin{bmatrix} 108.71 & 103 & 16.58\\ 103 & 305.61 & 30.89\\ 16.58 & 30.89 & 5.51 \end{bmatrix}$$

(c) You can compute the maximum likelihood estimate for the correlation coefficient from either  $\sum_{\Sigma}^{\Lambda}$  or S. Then

$$r = \frac{92.70}{\sqrt{(97.84)(275.05)}} = 0.5651$$
 and  $t = \frac{(0.5651)\sqrt{10-2}}{\sqrt{1-(.5651)^2}} = 1.9373$ 

with df = 10 - 2 = 8

(d) 
$$z = \frac{1}{2} \log \left( \frac{1+r}{1-r} \right) = \frac{1}{2} \log \left( \frac{1+.5651}{1-.5651} \right) = 0.6403$$
  
 $Z_{L} = 0.6403 - (1.96) \sqrt{\frac{1}{10-3}} = -0.1005$   
 $Z_{u} = 0.6403 + (1.96) \sqrt{\frac{1}{10-3}} = 1.3811$ 

and an approximate 95% confidence interval for the correlation coefficient is

$$\left(\frac{e^{2(-.1005)} - 1}{e^{2(-.1005)} + 1}, \frac{e^{2(1.3811)} - 1}{e^{2(1.3811)} + 1}\right) = (-0.1002, 0.8812)$$

(e) An estimate of the generalized variance is |S| = 42403.59, Answers may vary due to rounding off numbers in the computation of S or using the maximum likelihood estimator instead of S..

(f) An estimate of total variance is

trace(S) = 
$$108.71 + 305.61 + 5.51 = 419.83$$

(g) An approximate 95% confidence interval for the correlation coefficient is (0.0831, 0.9162)

(h) 
$$\mathbf{r}_{13,2} = \frac{\mathbf{r}_{13} \cdot \mathbf{r}_{12} \mathbf{r}_{23}}{\sqrt{(1 - \mathbf{r}_{12}^2)(1 - \mathbf{r}_{23}^2)}} = 0.4638$$
  
 $z = \frac{1}{2} \log \left(\frac{1 + r}{1 - r}\right) = \frac{1}{2} \log \left(\frac{1 + .4638}{1 - .4638}\right) = 0.5021$ 

and an approximate 95% confidence interval for the partial correlation coefficient is (-0.2894, 0.8623)

(i) The partial correlation for  $X_1$  and  $X_3$  given  $X_2$  could be zero since the 95% CI includes 0. This partial correlation is the correlation between levels of aspartate aminotransferase and glutamate dehydrogenase for any subpopulation of patients defined by a particular level of alanine aminotransferase.

(j) 
$$t = \frac{(0.4638)\sqrt{10-3}}{\sqrt{1-(.4638)^2}} = 1.384$$
 with df = 10-3 = 7 and p-value = 0.104

There is insufficient evidence to reject the null hypothesis  $\rho_{_{13.2}} = 0$ 

2. (a) Write the joint likelihood function in the form

$$L(\mu, \Sigma) = \frac{1}{(2\pi)^{np/2} |\Sigma|^{n/2}} e^{\frac{-1}{2} tr(\Sigma^{-1}A)} e^{\frac{-1}{2} (\bar{X} - \mu) \Sigma^{-1} (\bar{X} - \mu)}$$

where  $A = \sum_{j=1}^{n} (X_j - \overline{X}) (X_j - \overline{X})$ . To get the log-likelihood corresponding to the

null hypothesis, substitute

$$\Sigma = \sigma^2 I$$
,  $|\Sigma| = |\sigma^2 I| = (\sigma^2)^{np/2}$ , and  $tr(\Sigma^{-1}A) = \frac{1}{\sigma^2} tr(A)$ ,

Then, the log-likelihood for the null hypothesis is

$$\ell(\mu, \Sigma) = -\frac{np}{2}\log(2\pi) - \frac{np}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}tr(A) - \frac{n}{2\sigma^2}(\bar{X} - \mu)'(\bar{X} - \mu)$$

(b) Regardless of the value of  $\sigma^2$ , under the null hypothesis the log-likelihood is maximized when  $\stackrel{\wedge}{\mu} = \overline{x}$ . Then, the last term in the log-likelihood is zero, and the mle for  $\sigma^2$  can be derived by maximizing

$$\ell(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = -\frac{np}{2}\log(2\pi) - \frac{np}{2}\log(\boldsymbol{\sigma}^2) - \frac{1}{2\boldsymbol{\sigma}^2}\operatorname{tr}(\mathbf{A})$$

Setting the first partial derivative equal to zero yields the equation

$$\frac{\partial \ell(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial \boldsymbol{\sigma}^2} = - \frac{np}{2\boldsymbol{\sigma}^2} + \frac{1}{2\boldsymbol{\sigma}^4} \operatorname{tr}(A) = 0$$

The solution to this equation yields the maximum likelihood estimate

$$\hat{\sigma}^{2} = \frac{1}{np} \operatorname{tr}(A) = \frac{1}{p} \operatorname{tr}(\hat{\Sigma}) = \frac{n-1}{np} \operatorname{tr}(S) = \frac{1}{np} \sum_{j=1}^{n} (X_{j} - \overline{X})'(X_{j} - \overline{X})$$

which is simply the average of the estimated variances for the p responses obtained from the units in the sample..

(c) There are various ways to write the same formula, but the simplest form is

$$-2\ln(\Lambda) = (np)\log(\frac{1}{p} tr(\hat{\Sigma})) - (n)\log(|\hat{\Sigma}|)$$

- (d) d.f. =  $\left(\frac{p(p+1)}{2} + p\right)$  (p+1) =  $\frac{p(p+1)}{2}$  1 =  $\frac{(p+2)(p-1)}{2}$
- (e) Answers may vary if you do the calculations with a calculator instead of SAS or S- plus, depending on how you rounded off the numbers in the estimated covariance matrix.

 $-2\ln(\Lambda) = 41.687$  with 5 d.f. and p - value = 0

The null hypothesis is rejected which implies all correlation can't be zero.

Bartlett's correction to  $-2\ln(\Lambda) = 36.36$ , d.f = 5, p-value<0.0001

The correction had no effect on the conclusion that the correlation is not zero.

3.	Correlation:	0.8	-3.16	0.0	0.80	1.0
	Generalized variance:	9.0	9.0	9.0	5.76	0.0
	Total variance:	10.0	7.0	6.0	10.0	10.0

The first three cases exhibit different levels of correlation, but the same generalized variance. The total variance for these cases becomes smaller as the correlation gets closer to zero. When there is a perfect correlation, as in the last case, the generalized variance is zero because there is only variation along a onedimensional line, not in two-dimensional space. The total variance is positive because it only accounts for the sum of the variances of the univariate marginal distributions.

4. (a). 
$$S = \begin{bmatrix} 10963.69 & 21505.42 \\ 21505.42 & 63661.31 \end{bmatrix}$$
 with d.f.=98

(b).

$$M = \sum_{i=1}^{2} (n_{i} - 1) \log(|S|) - \sum_{i=1}^{2} (n_{i} - 1) \log(|S_{i}|) = 19.3662$$
$$C^{-1} = 1 - \frac{(2)2^{-2} + (2)(2) - 1}{6(2 + 1)(2 - 1)} \left(\frac{1}{45 - 1} + \frac{1}{55 - 1} - \frac{1}{44 + 54}\right) = 0.9776$$
$$MC^{-1} = 18.95 > \chi^{2}_{3,.005} = 12.84 \quad \text{and} \quad \text{p-value} = .0003 < .005$$

The covariance matrices are not the same. Homes with air conditioning exhibit larger variances for both on peak and off peak use of electricity.

(c). 
$$H_0: \rho_1 = \rho_2 \text{ vs. } H_a: \rho_1 \neq \rho_2$$
  
 $Z_1 = \frac{1}{2} \log \left( \frac{1+r_1}{1-r_1} \right) \approx N \left( \frac{1}{2} \log \left( \frac{1+p_1}{1-p_1} \right), \frac{1}{n_1 - 3} \right)$   
 $Z_2 = \frac{1}{2} \log \left( \frac{1+r_2}{1-r_2} \right) \approx N \left( \frac{1}{2} \log \left( \frac{1+p_2}{1-p_2} \right), \frac{1}{n_2 - 3} \right)$   
Since the secondary are independent. Z is independent of

Since the samples are independent,  $Z_1$  is independent of  $Z_2$ . This implies that  $Var(Z_1-Z_2) = Var(Z_1) + Var(Z_2) = \pm \frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}$  and  $Z_1-Z_2 \sim N\left(\frac{1}{2}\log\left(\frac{(1+p_1)(1-p_2)}{(1+p_2)(1-p_1)}\right), \frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}\right)$ 

under Ho, => Z<sub>1</sub>-Z<sub>2</sub> ~ N
$$\left(0, \frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}\right)$$
  
$$\frac{1}{2} \left[ \ln \frac{1 + p_1}{1 - p_1} - \ln \frac{1 + p_2}{1 - p_2} \right] / \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}} \sim Z(0, 1)$$

For theses data, Z = -2.14 with p-value=0.03, and the null hypothesis of equal correlations can be rejected at the 0.05 level of significance. The correlations between total on-peak and off-peak usage are not the same for homes with or without air conditioning.

(d) 
$$T^2 = 16.0662$$

- (e) F=7.95 d.f.=2,97 p-value=<0.0001
- (f). Do not use a pooled estimate of the covariance matrix. Estimate

Var
$$(\bar{X}_1 - \bar{X}_2)$$
 as  $\frac{1}{n_1}S_1 + \frac{1}{n_2}S_2$  and compute  
 $T^2 = (\bar{X}_1 - \bar{X}_2)' (\frac{1}{n_1}S_1 + \frac{1}{n_2}S_2)^{-1} (\bar{X}_1 - \bar{X}_2)$ 

A test statistic with an approximate F-distribution when the null hypothesis is

true

is  $F = \frac{n_1 + n_2 - p - 1}{p(n_1 + n_2 - 2)}T^2 = 7.74$  with  $(p, n_1 + n_2 - p - 1)$  degrees of freedom

The p-value<0.0001 is still small in this case.

5. (a)



There are some extreme values in the plot of the results for observer 1 versus the monitor and the plot of the results for observer 2 versus the monitor that indicate some substantial disagreements between the expert observers and the monitor. The agreement between the two expert observers is quite strong.

(b)

	X1	X2	X3
Value of W	0.9175	0.9154	0.9366
p-value	4.388e-05	3.504e-05	0.0004136

The p-values of the Shapiro-Wilk statistics are very small for all three blood pressure measurements, leading to the conclusion that none of three sets of measurements ia a random sample from a normal distribution. The QQ plots also indicate substantial right skewness for each distribution.



The Chi-Square probability plot showed that the fit of a three dimensional normal model is inappropriate since it doesn't resemble a straight line through the origin having a slope of 1. The extreme observations seen in the scatter plots in part (a) are clearly evident in the upper right part of this plot.

(c) Since the QQ plots give evidence of right skewness, it may be useful to consider a power transformation with lambda <1.

Results from Shapiro-Wilk tests and QQ plots, suggest the following transformation for the three responses. To compare the mean responses for the two expert observers and the monitor, we must use the same transformation for all three responses. It appears that the inverse transformation provides the best compromise for these three responses.

	X1	X2	X3
Transformation	X1^(-1)	X2^(-1)	X3^(1/2)
Shapiro-Wilk W	0.9893	0.9822	0.9819
p-value	0.7182	0.641	0.2786

lambda		X1	X2	X3
0.25	Shapiro-Wilk	0.9613	0.9592	0.9671
0.25	p-value	0.01199	0.00881	0.02876
Log	Shapiro-Wilk	0.9716	0.9696	0.974
Log	p-value	0.05781	0.04241	0.08307
1	Shapiro-Wilk	0.9893	0.9882	0.9814
-1	p-value	0.7181	0.641	0.2595
-0.5	Shapiro-Wilk	0.9853	0.9837	0.9819
	p-value	0.4493	0.3590	0.2786

(d) Using the natural logarithm of each response, we obtain

	0.051236	0.050604	0.040198		[ 1	0.99874	0.83241
S =	0.050604	0.050107	0.039877	R =	0.99874	1	0.83501
	0.040198	0.039877	0.045516		0.83241	0.83501	1

These results suggest that the correlation (or agreement) between the log-responses for the two expert observers is stronger than the correlation between the log-responses between either expert and the monitor. Also, variablity may be lower for the monitor. Perhaps the monitor has difficulty recording high blood pressure values.

(e)

$$-2\ln(\Lambda) = -n \ln |R| = 610.28$$
 with  $\frac{p(p-1)}{2} = 3 \text{ d.f.}$  and  $p - \text{value} = 0$ 

Bartlett's correction,

$$-\left(n-1-\frac{2p+5}{6}\right)\ln |R| = 589.94$$
 with 3 d.f. and p-value = 0

Since the estimated correlations are positive and the logarithm is a monotone transformation, there is evidence that at least one correlation is not zero which implies that there is more than random agreement either between the two experts or between at least one of the experts and the monitor.

(f)

Bartlett correction to  $-2*\log(\text{lambda})=313.0538$  with d.f = 4 and p-value <.0001.

Reject the null hypothesis that all of the correlations are equal and all of the variances are equal for the natural logarithms of systolic blood pressure measurements for the two expert observers and the semi-automatic monitor.

[,1] [,2] [,3] [,4] [,5] [,6] rlowerb 1.0041 0.8900 0.8746 1.0870 1.0818 0.9899 rupperb 1.0430 1.3871 1.3498 1.3918 1.3925 1.0045

(g) Since the two experts and the monitor provide repeated measurements on the same subjects, you cannot assume that the estimated correlations are independent. You can use the bootstrap method to obtain approximate 95% confidence intervals for differences between (or ratios of) correlations. You must take bootstrap samples from the original sample, sampling with replacement, and compute the correlation matrix for each bootstrap sample. Then, for each bootstrap sample compute the differences (or ratios) for the three pairs of correlations. Results below are based on 5000 bootstrap samples.

Ratio	95% Bootstrap C.I.		
$\rho_{12} / \rho_{13}$	(1.087, 1.392)		
$\rho_{12} / \rho_{23}$	(1.082, 1.393)		
$\rho_{13}/\rho_{23}$	(0.990, 1.004)		

The correlation between the blood pressure measurements for the two expert observers is stronger than the correlation between the blood pressure measurements for the monitor and either expert. The correlation between the first expert and the monitor is not significantly different from the correlation between the second expert and the monitor.

If you wanted to make simultaneous confidence intervals, you could have invoked the Bonferroni procedure and constructed individual 98.33% confidence intervals for the three differences (or ratios)

(h) Use the bootstrap approach described in part (g) to construct approximate confidence intervals for ratios of variances for log-responses. Note that you cannot use F-tests to compare variances in this case, because the variance estimates are not independent.

Variance Ratio	Bootstrap C.I.
var(ln(x1))/var(ln(x2))	(1.004, 1.043)
var(ln(x1))/var(ln(x3))	(0.890, 1.387)
var(ln(x2)) /var(ln(x3))	(0.875, 1.350)

The first expert displayed more variability than the second expert, but this does not necessarily imply that the first expert is less reliable. Why? There were no significant differences between the variance for the monitor and the variance for either expert, although the confidence intervals for the variance ratios were quite wide.

(i)  $H_0: u_1 = u_2 = u_3$  is equivalent to test that u1-u2=u2-u3=0.

$$H_0: C u = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = 0$$
  
Let  $Y = CX, X \sim NID(u, \Sigma)$   
Then  $\overline{Y} = C\overline{X}$   $S_y = \frac{1}{n-1} \sum_{j=1}^n (Y - \overline{Y}) (Y - \overline{Y})^{'}$   
 $T^2 = n(\overline{Y} - 0)^{'} S_y^{-1} (\overline{Y} - 0) = 77.279$ 

$$F = \frac{n - p + 1}{(p - 1)(n - 1)}T^2 = 38.17964 \quad d.f = 2, 83 \text{ and } p\text{-value} < 0.0001$$

There is evidence that the mean blood pressure readings are not the same for the two experts or the monitor.

(j) You can do three paired t-tests, or square the paired t-tests to obtain F-tests.  $H_0: u_1 = u_2$ ,

$$T^2 = 0.005350383$$
 d.f = (1, 84) and p-value = 0.9418

There is insufficient evidence to conclude the average blood pressure measurements differ for the two experts.

 $H_0: u_2 = u_3,$ T^2 = 76.30276 d.f = (1, 84) and p-value < 0.0001

## $H_0: u_1 = u_3$ T^2 = 74.140706 d.f = (1, 84) and p-value< 0.0001

There is significant evidence that the mean systolic blood pressure measurement is lower for the semi-automatic monitor than for either of the two experts.