**56:295 – 001**
**Multivariate Statistics and Advanced Quality Control**
**Fall 05**

HW6 Due: November 9 (Wednesday), 6:15pm

**Solution**

1. Refer to Examples 8.10 and 8.11 of J&W and the data in Table 5.8, page 240, of J&W (the data file is posted on web). Add another variable $x_6$ = regular overtime hours whose values are
   6187
   7336
   6988
   6964
   8425
   6778
   5922
   7307
   7679
   8259
   10954
   9353
   6291
   4969
   4825
   6019

and redo Examples 8.10 and 8.11. Remember to use centered observations to get correct principal component scores. You can use Calc→Standardize menu of Minitab to center the observations (Choose "Subtract mean" option in the dialog box). You can also use other software that you are familiar with.

**Ans.** The six normalized eigenvectors and eigenvalues of the sample covariance matrix as follows are corresponding to the six principal components. The first two sample principal components explain 82.6% of the total variance.

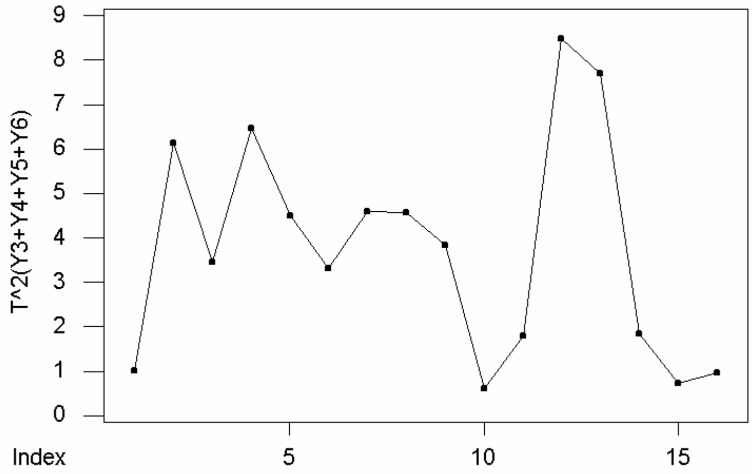
```
Eigenanalysis of the Covariance Matrix

Eigenvalue   4045922   2265079    761592    288919    181437     94303
Proportion     0.530     0.297     0.100     0.038     0.024     0.012
Cumulative     0.530     0.826     0.926     0.964     0.988     1.000

Variable         PC1       PC2       PC3       PC4       PC5       PC6
X1             0.001    -0.057    -0.516    -0.612     0.431    -0.413
X2             0.309    -0.554     0.562    -0.493    -0.180    -0.081
X3             0.482     0.386    -0.327    -0.340    -0.570     0.267
X4            -0.368    -0.641    -0.490     0.064    -0.431     0.154
X5             0.154     0.036    -0.032     0.307    -0.406    -0.845
X6             0.716    -0.358    -0.266     0.409     0.327     0.117
```

The following table lists the principal component scores for each observation. Using the first two sample principal components, we can calculate the statistic $T^2 = \frac{\hat{y}_1^2}{\hat{\lambda}_1} + \frac{\hat{y}_2^2}{\hat{\lambda}_2}$

| Period | y1 | y2 | y3 | y4 | y5 | y6 | T^2 (first two components) | T^2 (other four components) |
|---|---|---|---|---|---|---|---|---|
| 1 | -1745.42 | -1479.26 | 618.71 | -222.571 | 7.231 | 178.12 | 1.71905 | 1.01082 |
| 2 | 1096.6 | 2011.84 | 652.46 | 69.537 | 636.922 | 560.194 | 2.08413 | 6.13932 |
| 3 | -210.59 | 490.61 | 365.8 | 899.839 | -293.523 | -15.226 | 0.11723 | 3.45556 |
| 4 | 1360.05 | 1448.11 | 420.09 | -523.494 | -972.227 | 88.499 | 1.38299 | 6.47296 |
| 5 | 1255.88 | 502.07 | -422.36 | 893.807 | 359.916 | -273.731 | 0.50112 | 4.50785 |
| 6 | -971.6 | 284.69 | -316.91 | 942.849 | -83.53 | -70.131 | 0.26911 | 3.29935 |
| 7 | -1118.52 | 123.69 | 572.92 | -319.948 | -60.786 | -598.549 | 0.31597 | 4.60471 |
| 8 | 1151.59 | 1752 | -1322.06 | -700.232 | -242.194 | -158.813 | 1.68292 | 4.58283 |
| 9 | 497.26 | -593 | 209.51 | 149.237 | 101.623 | -586.24 | 0.21636 | 3.83603 |
| 10 | 2397.07 | 1819.6 | -9.51 | 147.577 | -109.861 | 207.818 | 2.88191 | 0.59999 |
| 11 | 3931.95 | -3715.69 | 924.1 | -35.086 | -274.25 | 152.913 | 9.91649 | 1.78804 |
| 12 | 1392.44 | -1688.02 | -2285.12 | -372.056 | 444.011 | 85.21 | 1.73719 | 8.49907 |
| 13 | -326.76 | 650.83 | 1251.62 | -728.777 | 809.54 | -139.973 | 0.2134 | 7.71501 |
| 14 | -3371.44 | -379.07 | -499.86 | 114.615 | -324.336 | 286.933 | 2.87284 | 1.82638 |
| 15 | -3076.56 | -199.06 | -105.72 | -419.802 | -122.302 | 3.406 | 2.35694 | 0.70721 |
| 16 | -2261.93 | -1029.35 | -53.67 | 104.505 | 123.767 | 279.569 | 1.73234 | 0.95481 |

For the first two principal components, we use upper control limit $\chi_2^2(0.05) = 5.99$. The control ellipse is omitted here. One point –period 11—is out of control.

We can create the chart for the other principal components using the statistic $T^2 = \frac{\hat{y}_3^2}{\hat{\lambda}_3} + \frac{\hat{y}_4^2}{\hat{\lambda}_4} + \frac{\hat{y}_5^2}{\hat{\lambda}_5} + \frac{\hat{y}_6^2}{\hat{\lambda}_6}$. The upper control limit is $\chi_4^2(0.05) = 9.488$. From the control chart, no point exceeds the control limit.




**2. Solve the following problems from the textbook (J&W)**

9.1

**Ans.**

$$L' = [.9 \quad .7 \quad .5]; \qquad LL' = \begin{bmatrix} .81 & .63 & .45 \\ .63 & .49 & .35 \\ .45 & .35 & .25 \end{bmatrix}$$

$$\text{so} \qquad \boldsymbol{\rho} = LL' + \Psi$$

9.2

**Ans.**

a) For $m = 1$ $h_1^2 = \ell_{11}^2 = .81$

$h_2^2 = \ell_{21}^2 = .49$

$h_3^2 = \ell_{31}^2 = .25$

The communalities are those parts of the variances of the variables explained by the single factor.

b) $\text{Corr}(Z_i, F_1) = \text{Cov}(Z_i, F_1)$, $i = 1,2,3$. By (9-5) $\text{Cov}(Z_i, F_1) = \ell_{i1}$. Thus $\text{Corr}(Z_1, F_1) = \ell_{11} = .9$; $\text{Corr}(Z_2, F_1) = \ell_{21} = .7$; $\text{Corr}(Z_3, F_1) = \ell_{31} = .5$. The first variable, $Z_1$, has the largest correlation with the factor and therefore will probably carry the most weight in naming the factor.

9.3

**Ans.**

a) $L = \sqrt{\lambda_1}\, \underline{e}_1 = \sqrt{1.96} \begin{bmatrix} .625 \\ .593 \\ .507 \end{bmatrix} = \begin{bmatrix} .876 \\ .831 \\ .711 \end{bmatrix}$. Slightly different from result in Exercise 9.1.

b) Proportion of total variance explained $= \dfrac{\lambda_1}{p} = \dfrac{1.96}{3} = .65$

9.12 (a)(b)(c)

**Ans.**

(a)

$\Psi_1 = s_{11} - l_{11}^2 = 11.072 \times 10^{-3} - 0.1022^2 = 6.27 \times 10^{-4}$

$\Psi_2 = s_{22} - l_{21}^2 = 6.417 \times 10^{-3} - 0.0752^2 = 7.62 \times 10^{-4}$

$\Psi_3 = s_{33} - l_{31}^2 = 6.773 \times 10^{-3} - 0.0765^2 = 9.21 \times 10^{-4}$

(b)

$\hat{h}_1^2 = l_{11}^2 = 0.0104$, $\hat{h}_2^2 = l_{21}^2 = 5.655 \times 10^{-3}$, $\hat{h}_3^2 = l_{31}^2 = 5.852 \times 10^{-3}$

(c)

Proportion of variance explained by the factors =

$$\frac{l_{11}^2 + l_{21}^2 + l_{31}^2}{s_{11} + s_{22} + s_{33}} = \frac{0.02195}{0.024262} = 0.9047 .$$