# SAS Examples of Chapter 3

## Example 3.1

Perspiration from 20 healthy females was analyzed. Three components, $X_1$ = sweat rate, $X_2$ = sodium content, and $X_3$ = potassium content, were measured, and the results are presented in the following table.

| Individual | $X_1$ | $X_2$ | $X_3$ | Individual | $X_1$ | $X_2$ | $X_3$ |
|---|---|---|---|---|---|---|---|
| 1 | 3.7 | 48.5 | 9.3 | 11 | 3.9 | 36.9 | 12.7 |
| 2 | 5.7 | 65.1 | 8.0 | 12 | 4.5 | 58.8 | 12.3 |
| 3 | 3.8 | 47.2 | 10.9 | 13 | 3.5 | 27.8 | 9.8 |
| 4 | 3.2 | 53.2 | 12.0 | 14 | 4.5 | 40.2 | 8.4 |
| 5 | 3.1 | 55.5 | 9.7 | 15 | 1.5 | 13.5 | 10.1 |
| 6 | 4.6 | 36.1 | 7.9 | 16 | 8.5 | 56.4 | 7.1 |
| 7 | 2.4 | 24.8 | 14.0 | 17 | 4.5 | 71.6 | 8.2 |
| 8 | 7.2 | 33.1 | 7.6 | 18 | 6.5 | 52.8 | 10.9 |
| 9 | 6.7 | 47.4 | 8.5 | 19 | 4.1 | 44.1 | 11.2 |
| 10 | 5.4 | 54.1 | 11.3 | 20 | 5.5 | 40.9 | 9.4 |

(a) Set $\alpha = 0.10$, perform a test for each of the following hypothesis:

(i) $H_0 : \mu_1 = 4$ vs $H_1 : \mu_1 \neq 4$,

(ii) $H_0 : \mu_2 = 50$ vs $H_1 : \mu_2 \neq 50$,

(iii) $H_0 : \mu_3 = 10$ vs $H_1 : \mu_3 \neq 10$.

(b) Set $\alpha = 0.10$, perform a test for $H_0 : \boldsymbol{\mu} = (4, 50, 10)'$ vs $H_1 : \boldsymbol{\mu} \neq (4, 50, 10)'$.

(c) Construct 90% Bonferroni confidence intervals for $\mu_1, \mu_2, \mu_3$.

(d) Construct 90% Scheffé confidence intervals for $\mu_1, \mu_2, \mu_3$.

**Solution**

(a) From the observations, we have

$$\bar{\boldsymbol{x}} = \begin{pmatrix} 4.64 \\ 45.4 \\ 9.9650 \end{pmatrix} \quad \text{and} \quad \boldsymbol{S} = \begin{pmatrix} 2.8794 & 10.0101 & -1.8091 \\ 10.0100 & 199.7884 & -5.6400 \\ -1.8091 & -5.6400 & 3.6277 \end{pmatrix}.$$

(i) To test $H_0 : \mu_1 = 4$ vs $H_1 : \mu_1 \neq 4$, the computed $t$-statistic is

$$t = \frac{4.64 - 4}{\sqrt{2.8794/20}} = 1.6867$$

and its $p$-value is $0.1080 > 0.10$. Thus, the null hypothesis cannot be rejected at 10% significance level.

(ii) To test $H_0 : \mu_2 = 50$ vs $H_1 : \mu_2 \neq 50$, the computed $t$-statistic is

$$t = \frac{45.4 - 50}{\sqrt{199.7884/20}} = -1.4554$$

and its $p$-value is $0.1619 > 0.10$. Thus, the null hypothesis cannot be rejected at 10% significance level.
(iii) Test $H_0 : \mu_3 = 10$ vs $H_1 : \mu_3 \neq 10$, the computed $t$-statistic is

$$t = \frac{9.9650 - 10}{\sqrt{3.6277/20}} = -0.0822$$

and its $p$-value is $0.9354 > 0.10$. Thus, the null hypothesis cannot be rejected at 10% significance level.

(b) To test $H_0 : \boldsymbol{\mu} = (4, 50, 10)'$ vs $H_1 : \boldsymbol{\mu} \neq (4, 50, 10)'$, the computed $T^2$-statistic is

$$\begin{aligned} T^2 &= 20 \times \begin{pmatrix} 4.64 - 4 \\ 45.4 - 50 \\ 9.9650 - 10 \end{pmatrix}' \begin{pmatrix} 2.8794 & 10.0101 & -1.8091 \\ 10.0100 & 199.7884 & -5.6400 \\ -1.8091 & -5.6400 & 3.6277 \end{pmatrix}^{-1} \begin{pmatrix} 4.64 - 4 \\ 45.4 - 50 \\ 9.9650 - 10 \end{pmatrix} \\ &= 9.7388 \end{aligned}$$

which is greater than the critical value $CV = \frac{(20-1)(3)}{20-3} F_{0.10}(3, 20-3) = 8.1726$ and the corresponding $p$-value is $0.0649 < 0.10$. Thus, the null hypothesis is rejected at 10% significance level.

(c) By putting $\boldsymbol{a} = (1, 0, 0)'$, $\boldsymbol{a} = (0, 1, 0)'$ and $\boldsymbol{a} = (0, 0, 1)'$ for the confidence intervals of $\mu_1$, $\mu_2$ and $\mu_3$ respectively, the 90% Bonferroni confidence intervals are

| | Lower Bound | Upper Bound |
|---|---|---|
| $\mu_1$ | 3.7694 | 5.5106 |
| $\mu_2$ | 38.1483 | 52.6517 |
| $\mu_3$ | 8.9878 | 10.9422 |

Note that $m = 3$ in this case.

(d) By putting $\boldsymbol{a} = (1, 0, 0)'$, $\boldsymbol{a} = (0, 1, 0)'$ and $\boldsymbol{a} = (0, 0, 1)'$ for $\mu_1$, $\mu_2$ and $\mu_3$ respectively, the 90% simultaneous (Scheffé) confidence intervals are

| | Lower Bound | Upper Bound |
|---|---|---|
| $\mu_1$ | 3.5553 | 5.7247 |
| $\mu_2$ | 36.3646 | 54.4354 |
| $\mu_3$ | 8.7475 | 11.1825 |

## Example 3.2

(a) Suppose there is an ‘anthropometric law’ which states that the means of height, chest circumference and MUAC (mid-upper-arm circumference) are in the ratios 6:4:1 and we wish to test whether or not our data are consistent with this law.

| Individual | Height (cm) | Chest circumference (cm) | MUAC (cm) |
|---|---|---|---|
| 1 | 78 | 60.6 | 16.5 |
| 2 | 76 | 58.1 | 12.5 |
| 3 | 92 | 63.2 | 14.5 |
| 4 | 81 | 59.0 | 14.0 |
| 5 | 81 | 60.8 | 15.5 |
| 6 | 84 | 59.5 | 14.0 |

(b) The table given previously refers to a sample of approximately two-year-old boys from a country in Asia while the following table gives data for approximately two-year-old girls from the same country.

| Individual | Height (cm) | Chest circumference (cm) | MUAC (cm) |
|---|---|---|---|
| 1 | 80 | 58.4 | 14.0 |
| 2 | 75 | 59.2 | 15.0 |
| 3 | 78 | 60.3 | 15.0 |
| 4 | 75 | 57.4 | 13.0 |
| 5 | 79 | 59.5 | 14.0 |
| 6 | 78 | 58.1 | 14.5 |
| 7 | 75 | 58.0 | 12.5 |
| 8 | 64 | 55.5 | 11.0 |
| 9 | 80 | 59.2 | 12.5 |

Test the hypothesis that there is no difference in the means of height, chest circumference and MUAC for two-year-old boys and girls at 0.01 significance level.

**Solution**
(a) Test $H_0 : \frac{1}{6}\mu_1 = \frac{1}{4}\mu_2 = \mu_3$. Under $H_0$, the relationship between $\mu_1$ and $\mu_2$ can be represented by

$$2\mu_1 - 3\mu_2 = 0.$$

The relationship between $\mu_1$ and $\mu_3$ can be represented by

$$\mu_1 - 6\mu_3 = 0.$$

Then, the null hypothesis can be rewritten as

$$H_0 : \begin{pmatrix} 2 & -3 & 0 \\ 1 & 0 & -6 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \mathbf{0} \quad \text{or} \quad \boldsymbol{C\mu} = \mathbf{0}.$$

From the observations of boys, we can have

$$n_1 = 6, \bar{\boldsymbol{x}}_1 = \begin{pmatrix} 82.0 \\ 60.2 \\ 14.5 \end{pmatrix}, \boldsymbol{S}_1 = \begin{pmatrix} 31.600 & 8.040 & 0.500 \\ 8.040 & 3.172 & 1.310 \\ 0.500 & 1.310 & 1.900 \end{pmatrix}.$$

The computed test statistic is

$$T^2 = 47.1434 > CV = 45.$$

Also, its corresponding $p$-value is $0.0092 < 0.01$. Thus, the null hypothesis is rejected at 1% significance level.

(b) From the observations of girls, we can have

$$n_2 = 9, \bar{\boldsymbol{x}}_2 = \begin{pmatrix} 76.0 \\ 58.4 \\ 13.5 \end{pmatrix}, \boldsymbol{S}_2 = \begin{pmatrix} 24.5000 & 5.6375 & 4.3125 \\ 5.6375 & 1.9700 & 1.4562 \\ 4.3125 & 1.4562 & 1.8125 \end{pmatrix}.$$

The pooled estimate of $\boldsymbol{\Sigma}$ is

$$\boldsymbol{S} = \frac{(n_1-1)\boldsymbol{S}_1 + (n_2-1)\boldsymbol{S}_2}{n_1+n_2-2} = \begin{pmatrix} 27.2308 & 6.5615 & 2.8462 \\ 6.5615 & 2.4323 & 1.4000 \\ 2.8462 & 1.4000 & 1.8462 \end{pmatrix}.$$

Now, we test $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ and the computed test statistic is

$$\begin{aligned} T^2 &= \frac{(6)(9)}{6+9} \begin{pmatrix} 82.0-76.0 \\ 60.2-58.4 \\ 14.5-13.5 \end{pmatrix}' \begin{pmatrix} 27.2308 & 6.5615 & 2.8462 \\ 6.5615 & 2.4323 & 1.4000 \\ 2.8462 & 1.4000 & 1.8462 \end{pmatrix}^{-1} \begin{pmatrix} 82.0-76.0 \\ 60.2-58.4 \\ 14.5-13.5 \end{pmatrix} \\ &= 5.3117 \end{aligned}$$

which is smaller than the critical value $CV = \frac{(6+9-2)(3)}{6+9-3-1} F_{0.01}(3, 6+9-3-1) = 22.0411$. Also, its corresponding $p$-value is $0.2693 > 0.01$. Thus, the null hypothesis cannot be rejected at 1% significance level.

## SAS Codes for Example 3.1

```
proc iml;
   x = {3.7 48.5 9.3,
        5.7 65.1 8.0,
        3.8 47.2 10.9,
        3.2 53.2 12.0,
        3.1 55.5 9.7,
        4.6 36.1 7.9,
        2.4 24.8 14.0,
        7.2 33.1 7.6,
        6.7 47.4 8.5,
        5.4 54.1 11.3,
        3.9 36.9 12.7,
        4.5 58.8 12.3,
        3.5 27.8 9.8,
        4.5 40.2 8.4,
        1.5 13.5 10.1,
        8.5 56.4 7.1,
        4.5 71.6 8.2,
        6.5 52.8 10.9,
        4.1 44.1 11.2,
        5.5 40.9 9.4 };

   n = nrow(x);
   p = ncol(x);
   xbar = x[:,]`;
   S = (x`*x - n*xbar*xbar`)/(n-1);
   print xbar[format=10.4] S[format=10.4];

   mu0 = {4, 50, 10};

   /* part (a) */
   /* (i) */
   tstat = (xbar[1]-mu0[1])/sqrt(S[1,1]/n);
   pval = 2*(1-probt(abs(tstat),n-1));
   print "Test for mu1=4",, tstat[format=10.4] pval[format=10.4];

   /* (ii) */
   tstat = (xbar[2]-mu0[2])/sqrt(S[2,2]/n);
   pval = 2*(1-probt(abs(tstat),n-1));
   print "Test for mu2=50",, tstat[format=10.4] pval[format=10.4];

   /* (iii) */
   tstat = (xbar[3]-mu0[3])/sqrt(S[3,3]/n);
   pval = 2*(1-probt(abs(tstat),n-1));
   print "Test for mu3=10",, tstat[format=10.4] pval[format=10.4];

   /* part (b)*/
   tsq = n*(xbar-mu0)`*inv(S)*(xbar-mu0);
   alpha = 0.1;
   cv = (n-1)*p/(n-p)*finv(1-alpha,p,n-p);
   pval = 1-probf((n-p)/(p*(n-1))*tsq,p,n-p);
   print tsq[format=10.4] cv[format=10.4] pval[format=10.4];

```

```
    /* part (c), calculation of Bonferroni C.I. */
    a = {1,0,0};
    merror = abs(tinv(alpha/(2*3),n-1))*sqrt(a‘*S*a/n);
    ci1 = (a‘*xbar-merror)||(a‘*xbar+merror);
    a = {0,1,0};
    merror = abs(tinv(alpha/(2*3),n-1))*sqrt(a‘*S*a/n);
    ci2 = (a‘*xbar-merror)||(a‘*xbar+merror);
    a = {0,0,1};
    merror = abs(tinv(alpha/(2*3),n-1))*sqrt(a‘*S*a/n);
    ci3 = (a‘*xbar-merror)||(a‘*xbar+merror);

    ci = ci1//ci2//ci3;
    r = {"mu1" "mu2" "mu3"};
    c = {"Lower Bound" "Upper Bound"};

    print ci[label="90% Bonferroni C.I.s" rowname=r colname=c format=10.4];

    /* part (d), calculation of Scheffe C.I. */
    a = {1,0,0};
    merror = sqrt(cv)*sqrt(a‘*S*a/n);
    ci1 = (a‘*xbar-merror)||(a‘*xbar+merror);
    a = {0,1,0};
    merror = sqrt(cv)*sqrt(a‘*S*a/n);
    ci2 = (a‘*xbar-merror)||(a‘*xbar+merror);
    a = {0,0,1};
    merror = sqrt(cv)*sqrt(a‘*S*a/n);
    ci3 = (a‘*xbar-merror)||(a‘*xbar+merror);

    ci = ci1//ci2//ci3;
    r = {"mu1" "mu2" "mu3"};
    c = {"Lower Bound" "Upper Bound"};

    print ci[label="90% Simultaneous C.I.s" rowname=r colname=c format=10.4];

quit;
```

## SAS Codes for Example 3.2

```
proc iml;
/* part (a) */
   x1 = {78 60.6 16.5,
         76 58.1 12.5,
         92 63.2 14.5,
         81 59.0 14.0,
         81 60.8 15.5,
         84 59.5 14.0 };
   n1 = nrow(x1);
   mean1 = x1[:,]`;
   W1 = (x1`*x1 - n1*mean1*mean1`);
   S1 = W1/(n1-1);
   print n1[format=5.0] mean1[format=10.4] S1[format=10.4] W1[format=10.4];

   alpha = 0.01;
   c = {2 -3 0,
        1 0 -6};
   q = nrow(c);
   tsq = n1*(c*mean1)`*inv(c*s1*c`)*(c*mean1);
   cv = (n1-1)*q/(n1-q)*finv(1-alpha,q,n1-q);
   pval = 1-probf(tsq*(n1-q)/(q*(n1-1)),q,n1-q);
   print tsq[format=10.4] cv[format=10.4] pval[format=10.4];

   /* part (b) */
   x2 = { 80 58.4 14.0,
          75 59.2 15.0,
          78 60.3 15.0,
          75 57.4 13.0,
          79 59.5 14.0,
          78 58.1 14.5,
          75 58.0 12.5,
          64 55.5 11.0,
          80 59.2 12.5};
   n2 = nrow(x2);
   mean2 = x2[:,]`;
   W2 = x2`*x2 - n2*mean2*mean2`;
   S2 = W2/(n2-1);
   print n2[format=5.0] mean2[format=10.4] S2[format=10.4] W2[format=10.4];

   S = (W1+W2)/(n1+n2-2);
   print S[format=10.4];

   alpha = 0.01;
   p = nrow(S1);
   tsq = n1*n2/(n1+n2)*(mean1-mean2)`*inv(s)*(mean1-mean2);
   m = n1+n2-p-1;
   cv = (n1+n2-2)*p/m*finv(1-alpha,p,m);
   pval = 1-probf(tsq*m/((n1+n2-2)*p),p,m);
   print tsq[format=10.4] cv[format=10.4] pval[format=10.4];

quit;
```