Sep 03, 07 4:53	**x7_4.sas**	Page 1/1

```
/* x7_4.sas */
/* Purpose:  */

title1 'Example 7.4.';

data a (type=cov);
  length _type_ $ 8 _name_ $ 8;
  input pop$ _type_ @;
  if _type_='COV' or _type_='CORR' then input _name_ @;
  input x1 x2 x3;
  cards;
Gregaria COV  X1 4.7350 0.5622 1.4685
Gregaria COV  X2 0.5622 0.1413 0.2174
Gregaria COV  X3 1.4685 0.2174 0.5702
Gregaria MEAN    25.80  7.81   10.77
Gregaria N       20     20     20
Intermed COV  X1 4.7350 0.5622 1.4685
Intermed COV  X2 0.5622 0.1413 0.2174
Intermed COV  X3 1.4685 0.2174 0.5702
Intermed MEAN    28.35  7.41   10.75
Intermed N       72     72     72
;

proc print;
title2 'List of data. NB special SAS-dataset (type=cov).';

proc candisc data=a distance pcov simple;
  class pop;
  var x1 x2 x3;

data b;
  input x1 x2 x3;
cards;
27.06 8.03 11.36
;

proc print data=b;
title2 'List of testdata.';

proc discrim data=a testdata=b testlist pool=yes
    bcov pcov tcov wcov bcorr pcorr tcorr wcorr simple;
  class pop;
  var x1 x2 x3;
title2 'Linear discriminant analysis - equal prior probabilities.';

run;
```

Example 7.4. 1

List of data. NB special SAS-dataset (type=cov).

Obs	_type_	_name_	pop	x1	x2	x3
1	COV	X1	Gregaria	4.7350	0.5622	1.4685
2	COV	X2	Gregaria	0.5622	0.1413	0.2174
3	COV	X3	Gregaria	1.4685	0.2174	0.5702
4	MEAN		Gregaria	25.8000	7.8100	10.7700
5	N		Gregaria	20.0000	20.0000	20.0000
6	COV	X1	Intermed	4.7350	0.5622	1.4685
7	COV	X2	Intermed	0.5622	0.1413	0.2174
8	COV	X3	Intermed	1.4685	0.2174	0.5702
9	MEAN		Intermed	28.3500	7.4100	10.7500
10	N		Intermed	72.0000	72.0000	72.0000

```
/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<
```

Example 7.4. 2

List of data. NB special SAS-dataset (type=cov).

The CANDISC Procedure

Observations	92	DF Total	91
Variables	3	DF Within Classes	90
Classes	2	DF Between Classes	1

Class Level Information

pop	Variable Name	Frequency	Weight	Proportion
Gregaria	Gregaria	20	20.0000	0.217391
Intermed	Intermed	72	72.0000	0.782609

```
/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<
```

Example 7.4. 3

List of data. NB special SAS-dataset (type=cov).

The CANDISC Procedure

Pooled Within-Class Covariance Matrix, DF = 90

Variable	x1	x2	x3
x1	4.735000000	0.562200000	1.468500000
x2	0.562200000	0.141300000	0.217400000
x3	1.468500000	0.217400000	0.570200000

Example 7.4. 4

List of data. NB special SAS-dataset (type=cov).

The CANDISC Procedure

Simple Statistics

Total-Sample

Variable	N	Sum	Mean	Variance	Standard Deviation
x1	92	2557	27.79565	5.80141	2.4086
x2	92	689.72000	7.49696	0.16727	0.4090
x3	92	989.40000	10.75435	0.56400	0.7510

---

pop = Gregaria

Variable	N	Sum	Mean	Variance	Standard Deviation
x1	20	516.00000	25.80000	4.73500	2.1760
x2	20	156.20000	7.81000	0.14130	0.3759
x3	20	215.40000	10.77000	0.57020	0.7551

---

pop = Intermed

Variable	N	Sum	Mean	Variance	Standard Deviation
x1	72	2041	28.35000	4.73500	2.1760
x2	72	533.52000	7.41000	0.14130	0.3759
x3	72	774.00000	10.75000	0.57020	0.7551

Example 7.4. 5

List of data. NB special SAS-dataset (type=cov).

The CANDISC Procedure

Pairwise Squared Distances Between Groups

$$D^2(i|j) = (\bar{X}_i - \bar{X}_j)' \, COV^{-1} \, (\bar{X}_i - \bar{X}_j)$$

Squared Distance to pop

From pop	Gregaria	Intermed
Gregaria	0	9.82348
Intermed	9.82348	0

F Statistics, NDF=3, DDF=88 for Squared Distance to pop

From pop	Gregaria	Intermed
Gregaria	0	50.11398
Intermed	50.11398	0

Prob > Mahalanobis Distance for Squared Distance to pop

From pop	Gregaria	Intermed
Gregaria	1.0000	<.0001
Intermed	<.0001	1.0000

/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<

Example 7.4. 6

List of data. NB special SAS-dataset (type=cov).

The CANDISC Procedure

Multivariate Statistics and Exact F Statistics

S=1 M=0.5 N=43

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.36921743	50.11	3	88	<.0001
Pillai's Trace	0.63078257	50.11	3	88	<.0001
Hotelling-Lawley Trace	1.70843115	50.11	3	88	<.0001
Roy's Greatest Root	1.70843115	50.11	3	88	<.0001

Example 7.4. 7

List of data. NB special SAS-dataset (type=cov).

The CANDISC Procedure

	Canonical Correlation	Adjusted Canonical Correlation	Approximate Standard Error	Squared Canonical Correlation
1	0.794218	0.790721	0.038705	0.630783

Eigenvalues of Inv(E)*H
= CanRsq/(1-CanRsq)

	Eigenvalue	Difference	Proportion	Cumulative
1	1.7084		1.0000	1.0000

Test of H0: The canonical correlations in the
current row and all that follow are zero

	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
1	0.36921743	50.11	3	88	<.0001

NOTE: The F statistic is exact.

/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<

Example 7.4. 8

List of data. NB special SAS-dataset (type=cov).

The CANDISC Procedure

Total Canonical Structure

Variable	Can1
x1	-0.552841
x2	0.510718
x3	0.013906

Between Canonical Structure

Variable	Can1
x1	-1.000000
x2	1.000000
x3	1.000000

Pooled Within Canonical Structure

Variable	Can1
x1	-0.373893
x2	0.339513
x3	0.008451

```
                               Example 7.4.                                      9
                List of data. NB special SAS-dataset (type=cov).

                             The CANDISC Procedure

                Total-Sample Standardized Canonical Coefficients

                           Variable              Can1

                           x1             -2.110484421
                           x2              0.890684936
                           x3              1.079564687


             Pooled Within-Class Standardized Canonical Coefficients

                           Variable              Can1

                           x1             -1.906669340
                           x2              0.818633154
                           x3              1.085479487


                           Raw Canonical Coefficients

                           Variable              Can1

                           x1             -0.876224498
                           x2              2.177801277
                           x3              1.437500533


                       Class Means on Canonical Variables

                           pop                   Can1

                           Gregaria        2.452885820
                           Intermed       -0.681357172

/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<
                               Example 7.4.                                     10
                            List of testdata.

                         Obs      x1      x2       x3

                          1     27.06    8.03    11.36

/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<
                               Example 7.4.                                     11
           Linear discriminant analysis - equal prior probabilities.

                             The DISCRIM Procedure

            Observations      92          DF Total                91
            Variables          3          DF Within Classes       90
            Classes            2          DF Between Classes       1


                            Class Level Information

              Variable                                                    Prior
  pop         Name         Frequency       Weight    Proportion    Probability
```



```
  Gregaria    Gregaria          20       20.0000      0.217391       0.500000
  Intermed    Intermed          72       72.0000      0.782609       0.500000
```

```
                             Example 7.4.                                      12
         Linear discriminant analysis - equal prior probabilities.

                         The DISCRIM Procedure
                    Within-Class Covariance Matrices

                      pop = Gregaria,     DF = 19

        Variable               x1                x2                x3

        x1            4.735000000       0.562200000       1.468500000
        x2            0.562200000       0.141300000       0.217400000
        x3            1.468500000       0.217400000       0.570200000

--------------------------------------------------------------------------------


                      pop = Intermed,     DF = 71

        Variable               x1                x2                x3

        x1            4.735000000       0.562200000       1.468500000
        x2            0.562200000       0.141300000       0.217400000
        x3            1.468500000       0.217400000       0.570200000

--------------------------------------------------------------------------------

/><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<>><<
                             Example 7.4.                                      13
         Linear discriminant analysis - equal prior probabilities.

                         The DISCRIM Procedure

            Pooled Within-Class Covariance Matrix,     DF = 90

        Variable               x1                x2                x3

        x1            4.735000000       0.562200000       1.468500000
        x2            0.562200000       0.141300000       0.217400000
        x3            1.468500000       0.217400000       0.570200000


               Between-Class Covariance Matrix,     DF = 1

        Variable               x1                x2                x3

        x1            2.212570888      -0.347069943      -0.017353497
        x2           -0.347069943       0.054442344       0.002722117
        x3           -0.017353497       0.002722117       0.000136106


               Total-Sample Covariance Matrix,     DF = 91

        Variable               x1                x2                x3

        x1            5.801409460       0.380580029       1.443590540
        x2            0.380580029       0.167267559       0.216387004
        x3            1.443590540       0.216387004       0.564002867
```



```
                             Example 7.4.                                      14
         Linear discriminant analysis - equal prior probabilities.

                         The DISCRIM Procedure
       Within-Class Correlation Coefficients  /  Pr > |r|

                           pop = Gregaria

       Variable              x1            x2            x3

       x1               1.00000       0.68732       0.89372
                                       0.0008        <.0001

       x2               0.68732       1.00000       0.76590
                         0.0008                      <.0001

       x3               0.89372       0.76590       1.00000
                         <.0001        <.0001

--------------------------------------------------------------------------------


                           pop = Intermed

       Variable              x1            x2            x3

       x1               1.00000       0.68732       0.89372
                                       <.0001        <.0001

       x2               0.68732       1.00000       0.76590
                         <.0001                      <.0001

       x3               0.89372       0.76590       1.00000
                         <.0001        <.0001

--------------------------------------------------------------------------------
```

Example 7.4. 15
Linear discriminant analysis - equal prior probabilities.

The DISCRIM Procedure

Pooled Within-Class Correlation Coefficients / Pr > |r|

Variable	x1	x2	x3
x1	1.00000	0.68732 <.0001	0.89372 <.0001
x2	0.68732 <.0001	1.00000	0.76590 <.0001
x3	0.89372 <.0001	0.76590 <.0001	1.00000

Between-Class Correlation Coefficients / Pr > |r|

Variable	x1	x2	x3
x1	1.00000 .	-1.00000 .	-1.00000 .
x2	-1.00000 .	1.00000 .	1.00000 .
x3	-1.00000 .	1.00000 .	1.00000 .

Total-Sample Correlation Coefficients / Pr > |r|

Variable	x1	x2	x3
x1	1.00000	0.38634 0.0001	0.79806 <.0001
x2	0.38634 0.0001	1.00000	0.70451 <.0001
x3	0.79806 <.0001	0.70451 <.0001	1.00000

Example 7.4. 16
Linear discriminant analysis - equal prior probabilities.

The DISCRIM Procedure
Simple Statistics

Total-Sample

Variable	N	Sum	Mean	Variance	Standard Deviation
x1	92	2557	27.79565	5.80141	2.4086
x2	92	689.72000	7.49696	0.16727	0.4090
x3	92	989.40000	10.75435	0.56400	0.7510

---

pop = Gregaria

Variable	N	Sum	Mean	Variance	Standard Deviation
x1	20	516.00000	25.80000	4.73500	2.1760
x2	20	156.20000	7.81000	0.14130	0.3759
x3	20	215.40000	10.77000	0.57020	0.7551

---

pop = Intermed

Variable	N	Sum	Mean	Variance	Standard Deviation
x1	72	2041	28.35000	4.73500	2.1760
x2	72	533.52000	7.41000	0.14130	0.3759
x3	72	774.00000	10.75000	0.57020	0.7551

Pooled Covariance Matrix Information

Covariance Matrix Rank	Natural Log of the Determinant of the Covariance Matrix
3	-3.45023

Example 7.4. 17
Linear discriminant analysis - equal prior probabilities.

The DISCRIM Procedure

Pairwise Generalized Squared Distances Between Groups

$$D^2(i|j) = (\bar{X}_i - \bar{X}_j)' \, COV^{-1} \, (\bar{X}_i - \bar{X}_j)$$

Generalized Squared Distance to pop

From pop	Gregaria	Intermed
Gregaria	0	9.82348
Intermed	9.82348	0

Linear Discriminant Function

$$\text{Constant} = -.5 \, \bar{X}_j' \, COV^{-1} \, \bar{X}_j \qquad \text{Coefficient Vector} = COV^{-1} \, \bar{X}_j$$

Linear Discriminant Function for pop

Variable	Gregaria	Intermed
Constant	-221.41089	-195.34403
x1	-2.18631	0.55999
x2	63.49325	56.66749
x3	0.31070	-4.19478

/><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><<

Example 7.4. 18
Linear discriminant analysis - equal prior probabilities.

The DISCRIM Procedure
Classification Results for Test Data: WORK.B
Classification Results using Linear Discriminant Function

Generalized Squared Distance Function

$$D_j^2(X) = (X - \bar{X}_j)' \, COV^{-1} \, (X - \bar{X}_j)$$

Posterior Probability of Membership in Each pop

$$Pr(j|X) = \exp(-.5 \, D_j^2(X)) \, / \, SUM_k \exp(-.5 \, D_k^2(X))$$

Posterior Probability of Membership in pop

Obs	Classified into pop	Gregaria	Intermed
1	Gregaria	0.9964	0.0036

Example 7.4. 19

Linear discriminant analysis - equal prior probabilities.

The DISCRIM Procedure

Classification Summary for Test Data: WORK.B

Classification Summary using Linear Discriminant Function

Generalized Squared Distance Function

$$D_j^2(X) = (X-\bar{X}_j)' \, COV^{-1} (X-\bar{X}_j)$$

Posterior Probability of Membership in Each pop

$$Pr(j|X) = \exp(-.5\, D_j^2(X)) \,/\, SUM_k \exp(-.5\, D_k^2(X))$$

Number of Observations and Percent Classified into pop

	Gregaria	Intermed	Total
Total	1	0	1
	100.00	0.00	100.00
Priors	0.5	0.5	