Elham Abd El-Razik[1]

ABSTRACT

There are several variable selection methods for deciding which variables to include in discriminant analysis. The purpose of variable selection techniques is to choose a suitable subset of variables. There are three common methods are usually referred to as forward selection, backward elimination and stepwise method.

The linear discriminant analysis has long been known since Fisher, 1936 and can be used not only to examine multivariate differences between groups, but also to determine which variables are the most useful for discriminating between groups.

In this paper, a new approach will be introduced to select the most important variables in discriminant function using mathematical programming (MP). The mathematical programming approach used to discriminate between two or more than two groups. The new selection approach can be applied directly to discriminant function with respect to their parameters.

The idea of the suggested approach depends on "indicator variable" which helped to select the desirable number of variable like the variable selection techniques. This variable used to be binary variable equal one when the variable was selected otherwise equal zero when another variable wasn't selected. The suggested mathematical programming approach doesn't have the assumptions of the multivariate statistical techniques. The suggested approach has been used successfully in a number of applications that are briefly described.

Keywords: Discriminant Analysis, Variable Selection, Mathematical Integer Programming.

1. INTRODUCTION

In the most classification procedures, the number of unknown parameters grows more than linearly with dimension of the data. It may be desirable to apply a method of variable selection for a meaningful reduction of the set of used variables for the classification problem.

The purpose of variable selection techniques is to choose a suitable subset (although not necessarily the best), with considerably less computing than the variable that is required for all possible variables. The variable selection methods are identified sequentially by adding or deleting (forward or backward method), depending on the method. The variable selection techniques have become the focus of much research in areas of application for which data sets with ten or hundreds of thousands of variables are available.

reduced set of variables, to analyze the contribution of the original variables to the separation, and to evaluate the degree of separation. Fisher (1936) developed the technique to create a linear discriminant function to establish maximum separation among three species of iris flowers based upon four measurements (Timm, 2002).

The rest of the paper is organized as follows. In next section (2) summarizes the selection variable in discriminant analysis with classical method. Section (3) presents review for some works in variable selection for discrimination with mathematical programming. Section (4) presents the formulation of the suggested non linear mixed integer programming approach. Section (5) illustrates the application of the suggested approach on a sample problem.

2. VARIABLE SELECTION IN DISCRIMINANT WITH CLASSICAL METHOD

The discriminant analysis has long been known and can be used not only to examine multivariate differences between groups, but also to determine, which variables are the most useful for discriminating between groups, and which groups are similar and which are different (Leotta, 2004). Whether one sub class of variables works as well as another, (Szepannek, 2000). As with all multivariate statistical techniques random sampling, adequate sample size, reliable measures and correct model specification are generally assumed to be present. As regards discriminant analysis specifically, linearity, homogeneity of variance - covariance matrices, univariate - multivariate normality, and a low collinearity among independent variables are assumed (Brown, 2004). Selection of variable in discriminant analysis has become the focus of much research in areas of application for which data sets with tens or hundreds of thousands of variables are available (Guyon, 2003). In (Rawlings, 1988) variable selection procedures have been used in different settings. Among them, the regression area has been investigated extensively. Other regression techniques for variable selection are described in (Jobsan, 1991). (Back et al., 1996) Presented three alternative techniques that can be used to select predictors for failure prediction purposes. The selected techniques have all different assumptions about the relationships between the independent variables. Linear discriminant analysis based on linear combination of independent variables, logit analysis uses the logistic cumulative probability function and genetic algorithms is global search procedure based on the mechanics of natural selection and natural genetics. (Feldesman, 2002) Introduced a non parametric alternative approach to

method is introduced to shrink the many variables into a smaller subset of variables with zero mean, unit variance, and zero correlation coefficient between variables in (Roe, 2003). (Bouchard, 2004) Purposed a new criterion "Bayesian Entropy Criterion". This approach provided an interesting alternative to the cross validated error rate which is highly time consuming. (Gomaa, 2004) grouped the classical methods into three categories, namely: stepwise, canonical variate and All – Subset. The study showed the advantages and disadvantages of the different methods. (Kim et al., 2005) Robust Fisher linear discriminant analysis can systematically alleviate the sensitivity problem by explicitly incorporating a model of data uncertainty in a classification problem and optimizing for the worst- case scenario under this model. In (Ebdon et al., 1998) the paper determined the effectiveness of discriminant analysis in distinguishing water conserving Kentucky bluegrass on the basis of canopy resistance and leaf area from a population of 61 Kentucky bluegrass cultivars.

3. VARIABLE SELECTION IN DISCRIMINANT WITH MATHEMATICAL PROGRAMMING

Mathematical programming methods don't have the assumptions of the multivariate statistical techniques. There are several mathematical programming methods for variable selection in discriminant analysis to discriminate between two groups. Since Fisher's seminal paper appeared in 1936, linear discriminants of various flavors have been developed by statisticians and applied in numerous fields. The introduction of discrimination by linear programming to the Operations Research literature can perhaps be credited to Mangasarian 1965, although the fundamental ideas can be traced a few years earlier Minnick 1961, Charnes 1964. The subject did not receive much academic attention for over a decade, but a paper published in 1981 by Freed and Glover's triggered a rich vein literature on the subject (Rawlings, 1988). Joachimsthaler and Stam 1990 suggested that the mathematical programming approach is the best approach for the classification problem in many situations. Nath and Jones in 1988 developed a linear programming method based on the Jackknife technique. They applied the method using the objective of minimization of the sum of deviation to discriminate between two groups. This method involves running the mathematical programming model with one observation excluded in turn so the computation load may be prohibitive (Gomaa, 2004). Glen 1999, 2003, and 2004 used MILP models to extend the scope of discriminant analysis used for classification. Topics covered: normalization and variable selection; using classification accuracy as the separation criterion when there are a large number

model which is developed precisely for these situations is introduced. The new models apply the chance-constrained goal programming approach to solve the classification problem in discriminant analysis. Other real life situations require discriminating between more than two groups. Two new methods namely, the compound and the single function methods, are presented. In (Gomaa, 2004) the paper introduced a new model. The new model used to select the most important variables in discriminant function using mathematical programming to discriminate between more than two groups. The advantages for the suggested model, is that by modified some constrains the model can select a certain variable(s), these model normalized for invariance under origin shift. The suggested approach of these new applications of mathematical integer programming is illustrated using published data.

4. THE SUGGESTED APPROACH FOR SELECTION OF VARIABLE IN DISCRIMINANT ANALYSIS

Mathematical programming methods also have certain advantages over the classical methods, for example, varied objectives and more complex problem formulations are easily accommodated, and some mathematical programming methods especially linear programming lend themselves to sensitivity analysis. This section *first* gives a description of the new method suggested by El-Hefnawy 1999 for selection of variables in discriminant analysis using mathematical programming in more than two groups. The model as follows:

$$Min\, Z = d_{12}^{\backslash}\Psi_1 + d_{21}^{\backslash}\Psi_2 + d_{22}^{\backslash}\Psi_2 + d_{31}^{\backslash}\Psi_3 + d_{32}^{\backslash}\Psi_3 + \cdots + d_{I1}^{\backslash}\Psi_I$$

$$X_1 A - d_{12} \le C_1\Psi_1 \quad (1)$$

$$X_2 A + d_{21} \ge C_1\Psi_2 \quad (2)$$

$$X_2 A - d_{22} \le C_2\Psi_2 \quad (3)$$

$$X_3 A + d_{31} \ge C_2\Psi_3 \quad (4)$$

$$X_3 A - d_{32} \le C_3\Psi_3 \quad (5)$$

.

.

$$X_I A - d_{I1} \ge C_I\Psi_I \quad (6)$$

$$C_i - C_{i-1} \ge S_1, i = 2,3,\cdots,I \quad (7)$$

$$A^{\backslash}\Psi_0 + C_1 + C_2 + \cdots + C_I = S \quad (8)$$

where

A is a $(K \times 1)$ vector of decision variables and is unrestricted on sign.

Ψ_0 is a $(K \times 1)$ column vector of ones.

The previous model introduced with an advantage since the classical assumptions cannot be satisfied in real world classification problem.

Accordingly, the suggested approach is an extension of the previous work which presented by El-Hefnawy 1999. *Second* this section gives a description of the new method suggested by Gomaa 2004 for selection of the most important variables in discriminant function using mathematical integer programming in more than two groups. The model as follows:

$$Minimize \sum_{k=1}^{m} \sum_{i=1}^{n_k} d_{ki}$$

Subject to:

$$\sum_{j=1}^{p} X_{1ij}(a_j^{+} - a_j^{-}) - d_{12} \le u_1 \Psi_1 \qquad ;i = 1,2,.....n_1 \quad (9)$$

$$\sum_{j=1}^{p} X_{2ij}(a_j^{+} - a_j^{-}) - d_{21} \le u_1 \Psi_2 \qquad ;i = 1,2,.....n_2 \quad (10)$$

$$\sum_{j=1}^{p} X_{2ij}(a_j^{+} - a_j^{-}) - d_{22} \le u_2 \Psi_2 \qquad ;i = 1,2,.....n_2 \quad (11)$$

$$\sum_{j=1}^{p} X_{3ij}(a_j^{+} - a_j^{-}) - d_{31} \le u_2 \Psi_3 \qquad ;i = 1,2,.....n_3 \quad (12)$$

$$\sum_{j=1}^{p} X_{3ij}(a_j^{+} - a_j^{-}) - d_{32} \le u_3 \Psi_3 \qquad ;i = 1,2,.....n_3 \quad (13)$$

.

.

.

$$\sum_{j=1}^{p} X_{mij}(a_j^{+} - a_j^{-}) - d_{m1} \le u_m \Psi_m \qquad ;i = 1,2,.....n_m \quad (14)$$

$$\sum_{j=1}^{p} (a_j^{+} - a_j^{-}) = S \qquad (15)$$

$$u_k - u_{k-1} \ge S_1 \qquad ;k = 2,3,.....,m \quad (16)$$

$$a_j^{+} - \varepsilon \delta_j \ge 0 \qquad (17)$$

$$a_j^{+} - \delta_j \le 0 \qquad (18)$$

$$a_j^{-} - \varepsilon \gamma_j \ge 0 \qquad (19)$$

$$a_j^{-} - \gamma_j \le 0 \qquad (20)$$

$$\delta_j + \gamma_j \le 1 \qquad (21)$$

$u_1, u_2, \ldots\ldots, u_m$ are decision variables and unrestricted on sign, S, S_1 are positive constants, ψ_j is a $(n_k * 1)$ column vector of ones, $a_j^+, a_j^- \geq 0,\ \delta_j, \gamma_j = 0,1, d_{ki} \geq 0, k = 1,2,3; i = 1,2,\ldots\ldots n_k, and j = 1,2,\ldots\ldots, p.$ r=the number of the most important variables.

The following approach is an extension of the work presented by El-Hefnawy 1999 which concentrates on the "indicator variable". The advantage for this indicator variable is to select the variables in discriminant analysis using mathematical programming in two or more than two groups. This variable used to be binary variable equal one when the variable was selected otherwise equal zero when another variable wasn't selected. This approach helped in applications in which observations consist of a large number of variables, to select a limited number of variables in order to simplify the model.

The suggested approach formulation as follows:

$$MinZ = d_{12}^{\backslash}\Psi_1 + d_{21}^{\backslash}\Psi_2 + d_{22}^{\backslash}\Psi_2 + d_{31}^{\backslash}\Psi_3 + d_{32}^{\backslash}\Psi_3 + \cdots + d_{m1}^{\backslash}\Psi_m$$

$$X_1 M_1 A - d_{12} \leq U_1 \Psi_1 \qquad i = 1,2,\ldots\ldots n_1 \qquad (23)$$

$$X_2 M_2 A + d_{21} \geq U_1 \Psi_2 \qquad i = 1,2,\ldots\ldots n_2 \qquad (24)$$

$$X_2 M_2 A - d_{22} \leq U_2 \Psi_2 \qquad i = 1,2,\ldots\ldots n_2 \qquad (25)$$

$$X_3 M_3 A + d_{31} \geq U_2 \Psi_3 \qquad i = 1,2,\ldots\ldots n_3 \qquad (26)$$

$$X_3 M_3 A - d_{32} \leq U_3 \Psi_3 \qquad i = 1,2,\ldots\ldots n_3 \qquad (27)$$

.

.

$$X_m M_m A - d_{m1} \geq U_m \Psi_I \qquad i = 1,2,\ldots\ldots n_m \qquad (28)$$

$$U_k - U_{k-1} \geq S_1, \qquad k = 2,3,\cdots, m \qquad (29)$$

$$\sum_k U_k = S \qquad (30)$$

$$\sum_p M_{p-r1} = r_2 \qquad r_1 + r_2 = p = number of \text{ var} iable \qquad (31)$$

where

A is a $(K \times 1)$ vector of decision variables and is unrestricted on sign.

d_{mI} is a $(N_i \times 1)$ non-negative vector of decision variables.

$U_1, U_2, \cdots, U_k$ are decision variables and are unrestricted on sign.

S, S_1 are positive constant.

Ψ_i is a $(N_i \times 1)$ column vector of ones.

M_p is a binary variable =0,1.

programming approach to select the variables in discriminant analysis in two or more than two groups. The first numerical application will be comparison between the suggested approach, the suggested approach by Gomaa 2004 and statistical method. The second numerical application will be comparison between the suggested approach, the suggested approach by Gomaa 2004 when the number of variables differed. The third numerical application will be used to estimate the discriminant function when more than two group.

The suggested mathematical programming approach comparison between the previous methods which select the most important variables in discriminant function to discriminate between two or more groups are listed at the end of the study.

The first

The study introduced a program by using GAMS 2.25 statistical package to determine the efficiency of the suggested mathematical programming approach to select the most important variables in discriminant function to discriminate between two groups. The suggested program will be listed at the end of study

The following section presents the results of the comparison between the suggested two approaches and the statistical method. The data a random sample of 32 loan application from a bank. Data are available on each applicant's total family assets, total family income total debt outstanding, family size, number of years with present employer for household head, and a qualitative variable that equals 1 if the applicant has repaid the loan and 0 if he or she has not repaid the loan. The data and the analyzed for the data using SPSS program derived from (Aczel, 2003).

The second

The study introduced a program by using GAMS 2.25 statistical package to determine the efficiency of the suggested mathematical programming approach by Gomaa 2004 to select the most important variables in discriminant function to discriminate between two or more than two groups. The suggested program will be listed at the end of study. To compare between the previous methods, the data were reported as an application in (Aczel, 2003).

The following section presents the results of the comparison between the suggested two approaches when the number of variables differed. The present study used the previous examples to analyze the same data using the two suggested approaches and Compare between the study suggested approach and

The third

Finally the study will be introduces the suggested mathematical programming approach to select the most important variables in discriminant function when discriminate between more than two groups.

The following section presents the suggested approach to select the most important variables in discriminant function when discriminate between more than two groups. The data a random sample of 46 loan application from a bank. Data are available on each applicant's total family assets, total family income total debt outstanding, family size, number of years with present employer for household head, and a qualitative variable that equals 1 if the applicant has repaid the loan , 0 if he or she has not repaid the loan and 2 if the people have some difficulties. The suggested program with data will be listed at the end of the study.

6. RESULTS

The first case:

As an illustration, consider the data which taken from (Aczel 2003). The suggested approach will be applied to estimate the discriminant function for the two groups (m=2) for the first case, (S_1=any positive number), (S= any positive number), and (r_2=the desirable number of variable) where (p=number of variable =5) for the two mathematical programming.

The estimated discriminant function which calculated in (Aczel 2003) by using SPSS was:

D=-0.995 -0.0352 Assets + 0.0429 Debt + .483 Size

The estimated discriminant function which calculated for the suggested approach by Gomaa 2004 was:

D=.015Assets + 0.026 Debt + .536 Size

Z=2.302

From the pervious results for the estimated discriminant function which was calculated by the three methods under consideration, the suggested approach selected the same variables as the three most discriminating variables between the two groups. The objective function (z) value is smaller for the suggested approach by the researcher than the value which computed by by Gomaa 2004. The study suggested two steps at the end of the program to calculate the "highest probability" or "effectiveness of prediction" to helped us to calculate the correct classification proportion. After the results comparing by the "actual group" two groups the correct classification proportion 71.8% was calculated in the pervious paper. The correct classification proportion 92.8% was calculated for Gomaa 2004. But the correct classification proportion 100% by using the suggested approach.

The second case

The suggested approach will be applied to estimate the discriminant function for the two groups when the number of variables differed.

The estimated discriminant function when number of variables=2 which calculated for the suggested approach by Gomaa 2004 was:

D=.018Assets +.037 income

Z=2.763

Whereas, the estimated discriminant function when number of variables=2 were computed by the researcher using the suggested mathematical programming approach was:

D=.002 Assets +.004 Income

Z=.996

The estimated discriminant function when number of variables=4 which calculated for the suggested approach by Gomaa 2004 was:

D=.012 Assets +.025 Income +.021debt+.429size

Z=1.842

Whereas, the estimated discriminant function when number of variables=4

The estimated discriminant function when number of variables=5 which calculated for the suggested approach by Gomaa 2004 was:

D=.009 Assets +.019 Income +.016debt+.322size -.048job

Z=1.381

Whereas, the estimated discriminant function when number of variables=5 were computed by the researcher using the suggested mathematical programming approach was:

D=.0007465 Assets +.002 Income +.001debt+.027size -.004job

Z=.460

From the pervious results for the estimated discriminant function which was calculated by the two suggested methods under consideration, the suggested approach selected the same variables as the most discriminating variables between the two groups. The objective function (z) value is smaller for the suggested approach by the researcher than the value which computed by Gomaa 2004.

The third case:

The suggested mathematical programming approach will be applied to estimate the discriminant function and select the most important variables in discriminant function when discriminate between more than two groups.

The estimated discriminant function was computed by the researcher using the suggested mathematical programming approach was:

D= .001Assets +4.43376E-4 Income +.002debt+.020size -.004job

Z=1.544

The correct classification proportion 100% by using the suggested approach.

7. CONCLUSION

From the previous results, it can be summarized that suggested mathematical programming approach has the following advantages:

1-The suggested mathematical programming approach doesn't have the

2- The idea of the suggested approach depends on "indicator variable" which helped to select the desirable number of variable like the variable selection techniques.

3- The suggested approach estimated the discriminant function for two groups like the statistical methods. Also it can be used to select the most important variables in discriminant function to discriminate between more than two groups.

4- The suggested approach calculates the "highest probability" or "effectiveness of prediction" to helped for calculating the correct classification proportion.

5- In the suggested two mathematical integer approaches, the discriminant function with any number of variables is derived directly from the solution to the mathematical integer approaches, it can also be seen that, as expected the optimal objective function value decreases as the number of variables are increased. The suggested approach by the researcher calculates smallest objective function in all case or when the number of variables differed.

6-The suggested mathematical approach proved efficiency in correctly classifying population group members, corresponding the solution with the statistical methods, and produced the best discriminant function with correct classification proportion is 100%. The suggested mathematical approach can be applied to any classification method.

ACKNOWLEDGEMENTS

I am extremely grateful to *prof. Dr. Ramadan H. Mohamed* for his kind support in this work.

REFERENCES:

1- Aczel D.A.(2003) " Complete Business Statistics " IRWIN Homewood, Boston.

2- Back B., Laitinen T.,Sere K., Weze v.M (1996) "Choosing Bankruptcy Predictors Using Discriminant Analysis, Logit Analysis, and Genetic Algorithms" Turku Center for Computer Science.

3- Bouchard G., Celeux G.(2004) "Model Selection Supervised Classification" Theme COG – Systemes cognitifs Projects Select.

4- Brown J. (2004) "Techniques of Multivariate Data Analysis" ph.D.

5- Ebdon J.S., Petrovic A.M., Schwager S.J.(1998) " Evaluation of Discriminant Analysis in Identification of Low – and High – Water Use Kentucky Bluegrass Cultivars" Crop Sci.38:152-157.

6- El- Hefnawy A.(1999) " Mathematical Program Approach to Discriminant Analysis with Application in Demography" Cairo University, Faculty of Economics &Political Science, Department of Statistics.

7- Feldesman R.M.(2002) " Classification Trees as Alternative to Linear Discriminant Analysis" American Journal of Physical Anthropology 119:257-275.

8- Gomaa Abd El-Salam (2004) " Selection of Variables in Discriminant Analysis Using Mathematical Programming Approach" Cairo University, Faculty of Economics &Political Science, Department of Statistics.

9- Guyon I., Elisseeff (2003)" An Introduction to Variable and Feature Selection" Journal of Machine Liearning Research 3 1157-1182.

10- Jobson J.D.(1991) " Applied Multivariate Data Analysis" Springer-Verlag -New York.

11- Karam A. (2005) "Essays on Linear Discrimination" ph.D

12- Kim j.S., Magnani A., Boyd P.S.,(2005)" Robust Fisher Discriminant Analysis" Information System Laboratory – Electrical Engineering Department, Stanford University.

13- Leotta Roberto (2004) "Use of Linear Dicsriminant Analysis to Characterise Three Dairy Cattle Breeds on the Basis of Several Milk Characteristics" ITAL.J.ANIM.SCI.VOL.3,377-383.

14- Rawlings J.O. (1988) "Applied Regression Analysis A research Tool" John Wiley and sons, Inc. New York.

15- Roe B.P.(2003) " Event Selection Using an Extended Fisher Discriminant Method" PHYSTST, SLAC, Stanford, California 8-11.

16- Szepannek G.,Weihs C.(2000) "Variable Selection for Discrimination of More than Two Classes Where Data are Sparse" Fachbereich statistic Universitat Dortmund.

17- Timm N.H (2002) "Applied Multivariate Analysis" Springer-Verlag - New York.

First Case

```
sets

        i     size    /1*32/
        g      group  /1*2/
        j   variable/assets, income, debt, size ,job,def /
        n    dev   /1*2/       ;
table  dat(i,*)
    assets       income          debt          size       job       def
1   98            35              12            4          4         1
2   65            44              5             3          1         1
3   22            50              0             2          7         1
4   78            60              34            5          5         1
5   50            31              4             2          2         1
6   21            30              5             3          7         1
7   42            32              21            4          11        1
8   20            41              10            2          3         1
9   33            25              0             3          6         1
10  57            32              8             2          5         1
11  21            12              28            3          2         0
12  10            17              0             2          3         0
13  60            40              10            3          2         0
14  78            60              8             3          5         0
15  59            18              9             3          5         0
16  12            23              10            4          4         0
17  55            36              12            2          5         0
18  67            33              35            2          4         0
19  81            23              12            2          1         0
20  0             15              10            4          2         0
21  12            18              7             3          4         0
22  77            21              19            4          2         1
23  15            14              28            2          1         1
24  30            27              50            4          4         0
25  29            18              30            3          6         1
26  91            22              0             4          5         1
27  12            25              39            5          3         0
28  23            30              65            3          1         1
29  34            45              21            2          5         1
30  57            39              13            5          8         1
31  45            33              9             4          7         1
32  42            45              12            3          8         0

parameter
y(g) constant
h(i)  rusl
k(i)  rty
d(i)  fgg  ;
 y(g)=1;


binary variables m(j);
positive variable d1(g,n);
free variable z,x1,x2,x3,x4,x5,x6,u(g);
equation l1,l2(i),l3(i),l5(g),l15(i),l6(i),l7(i),l4;
l1..z=e=sum((g,n),d1(g,n)*y(g));
l2(i)$( dat(i,"def")eq 0)..(x1*dat(i,"assets")+x2* dat(i,"income")
+x3* dat(i,"debt")+x4* dat(i,"size")+x5* dat(i,"job"))*sum(j,(m(j)))-d1("1","2"
=l=u("1")*y("1");
l3(i)$(dat(i,"def")eq 1)..(x1*dat(i,"assets")+x2* dat(i,"income")+x3* dat(i,"
debt")
+x4*dat(i,"size")+x5* dat(i,"job"))*sum(j,(m(j)))+d1("2","1")=g=u("1")*y("2");
l15(i)$( dat(i,"def")eq 1)..(x1* dat(i,"assets")+x2*dat(i,"income")+x3*dat(i,"
```

```
   debt")
60 +x4* dat(i,"size")+x5* dat(i,"job"))*sum(j,(m(j)))-d1("2","2")=l=u("2")*y("2");
61
62 l5(g)..u(g)-u(g-1)=g=1.2;
63 l4..sum((g),u(g))=e=5;
64
65 l6(i)$( dat(i,"def")eq 0)..sum(j,m(j-3))=e=3;
66 l7(i)$( dat(i,"def")eq 1)..sum(j,m(j-32))=e=3;
67
68
69   model test/all/;
70
71 solve test using minlp minimizing z;
72
73  h(i)$( dat(i,"def")eq 0)=(dat(i,"assets")*x1.l +
74 dat(i,"debt")*x3.l+dat(i,"size")*x4.l)<u.l("1");
75 k(i)$( dat(i,"def")eq 1)=(dat(i,"assets")*x1.l +
76 dat(i,"debt")*x3.l+dat(i,"size")*x4.l)<u.l("2");
77
78   display x1.l,x2.l,x3.l,x4.l,x5.l,d1.l,u.l,m.l,h,k;
```

C:\WINDOWS\gamsdir\dis(data).gms 24 يونيو، ٢٠٠٧ ١٢:٤٦:٥٩ ص

second case

```
sets

        i     size    /1*32/
        g      group  /1*2/
        j   variable/assets, income, debt, size ,job,def /

table  dat(i,*)
```

	assets	income	debt	size	job	def
1	98	35	12	4	4	1
2	65	44	5	3	1	1
3	22	50	0	2	7	1
4	78	60	34	5	5	1
5	50	31	4	2	2	1
6	21	30	5	3	7	1
7	42	32	21	4	11	1
8	20	41	10	2	3	1
9	33	25	0	3	6	1
10	57	32	8	2	5	1
11	21	12	28	3	2	0
12	10	17	0	2	3	0
13	60	40	10	3	2	0
14	78	60	8	3	5	0
15	59	18	9	3	5	0
16	12	23	10	4	4	0
17	55	36	12	2	5	0
18	67	33	35	2	4	0
19	81	23	12	2	1	0
20	0	15	10	4	2	0
21	12	18	7	3	4	0
22	77	21	19	4	2	1
23	15	14	28	2	1	1
24	30	27	50	4	4	0
25	29	18	30	3	6	1
26	91	22	0	4	5	1
27	12	25	39	5	3	0
28	23	30	65	3	1	1
29	34	45	21	2	5	1
30	57	39	13	5	8	1
31	45	33	9	4	7	1
32	42	45	12	3	8	0

```
parameter
y(g) constant
h(i)  rusl
k(i)  rty;

y(g)=1;

variable x(g,i,j);
binary variables m(j),n(j);
positive variable d1(g,i),a1(j),a2(j);
free variable z,x1,x2,x3,x4,x5,x6,u(g);
equation l1,l2(i),l3(i),l4,l5(g),l6(j),l7(j),l8(j),l9(j),l10(j),l11,l15(i);
l1..z=e=sum((g,i),d1(g,i));
l2(i)$( dat(i,"def")eq 0)..(x1*dat(i,"assets")+x2* dat(i,"income")
+x3* dat(i,"debt")+x4* dat(i,"size")+x5* dat(i,"job"))*sum(j,(a1(j)-a2(j)))-d1
1","2")=l=u("1")*y("1");
l3(i)$(dat(i,"def")eq 1)..(x1*dat(i,"assets")+x2* dat(i,"income")+x3* dat(i,"
debt")
+x4*dat(i,"size")+x5* dat(i,"job"))*sum(j,(a1(j)-a2(j)))+d1("2","1")=g=u("1")*
```

```
   y("2");
60 l15(i)$( dat(i,"def")eq 1)..(x1* dat(i,"assets")+x2*dat(i,"income")+x3*dat(i,"
   debt")
61 +x4* dat(i,"size")+x5* dat(i,"job"))*sum(j,(a1(j)-a2(j)))-d1("2","2")=l=u("2"
   y("2");
62 l4..sum(j,a1(j)-a2(j))=e=1.13;
63 l5(g)..u(g)-u(g-1)=g=5;
64 l6(j)..a1(j)-.001*m(j)=g=0;
65 l7(j)..a1(j)-m(j)=l=0;
66 l8(j)..a2(j)-.001*n(j)=g=0;
67 l9(j)..a2(j)-n(j)=l=0;
68 l10(j)..m(j)+ n(j)=l=1;
69 l11..sum(j,m(j)+n(j))=e=3
70
71   model test/all/;
72
73 solve test using minlp minimizing z;
74  h(i)$( dat(i,"def")eq 0)=(dat(i,"assets")*x1.l +
75 dat(i,"debt")*x3.l+dat(i,"size")*x4.l)<u.l("1");
76 k(i)$( dat(i,"def")eq 1)=(dat(i,"assets")*x1.l +
77 dat(i,"debt")*x3.l+dat(i,"size")*x4.l)<u.l("2");
78
79   display a1.l,a2.l,x1.l,x2.l,x3.l,x4.l,x5.l,d1.l,u.l,h,k;
```

Finally

```
sets

        i     size    /1*46/
        g      group  /1*3/
        j   variable/assets, income, debt, size ,job,def /
        n    dev   /1*3/       ;
table  dat(i,*)
    assets       income          debt          size       job       def
1    98            35             12            4          4         1
2    65            44             5             3          1         1
3    22            50             0             2          7         1
4    78            60             34            5          5         1
5    50            31             4             2          2         1
6    21            30             5             3          7         1
7    42            32             21            4          11        1
8    20            41             10            2          3         1
9    33            25             0             3          6         0
10   57            32             8             2          5         0
11   21            12             28            3          2         0
12   10            17             0             2          3         0
13   60            40             10            3          2         0
14   78            60             8             3          5         0
15   59            18             9             3          5         0
16   12            23             10            4          4         0
17   55            36             12            2          5         0
18   67            33             35            2          4         1
19   81            23             12            2          1         0
20   0             15             10            4          2         1
21   12            18             7             3          4         0
22   77            21             19            4          2         1
23   15            14             28            2          1         1
24   30            27             50            4          4         0
25   29            18             30            3          6         2
26   91            22             0             4          5         2
27   12            25             39            5          3         2
28   23            30             65            3          1         2
29   34            45             21            2          5         2
30   57            39             13            5          8         2
31   45            33             9             4          7         2
32   42            45             12            3          8         2
33   13            18             8             3          5         2
34   55            12             9             4          6         2
35   60            30             6             5          3         2
36   89            22             4             2          4         2
37   23            56             10            4          8         2
38   34            45             18            2          2         2
39   15            13             16            3          3         0
40   78            45             5             2          6         2
41   56            3              9             4          8         1
42   66            23             15            2          7         0
43   58            12             3             3          9         1
44   24            14             2             2          6         2
45   45            25             13            3          2         0
46   39            23             1             4          3         2

parameter
y(g) constant
h(i)  rusl
k(i)  rty
d(i)  fgg;
 y(g)=1;
```

```
62
63
64 binary variables m(j);
65 positive variable d1(g,n);
66 free variable z,x1,x2,x3,x4,x5,x6,u(g);
67 equation l1,l2(i),l3(i),l5(g),l15(i),l14(i),l16(i),l6(i),l7(i),l8(i),l4;
68 l1..z=e=sum((g,n),d1(g,n)*y(g));
69 l2(i)$( dat(i,"def")eq 0)..(x1*dat(i,"assets")+x2* dat(i,"income")
70 +x3* dat(i,"debt")+x4* dat(i,"size")+x5* dat(i,"job"))*sum(j,(m(j)))-d1("1","2"
   =l=u("1")*y("1");
71 l3(i)$(dat(i,"def")eq 1)..(x1*dat(i,"assets")+x2* dat(i,"income")+x3* dat(i,"
   debt")
72 +x4*dat(i,"size")+x5* dat(i,"job"))*sum(j,(m(j)))+d1("2","1")=g=u("1")*y("2");
73 l15(i)$( dat(i,"def")eq 1)..(x1* dat(i,"assets")+x2*dat(i,"income")+x3*dat(i,"
   debt")
74 +x4* dat(i,"size")+x5* dat(i,"job"))*sum(j,(m(j)))-d1("2","2")=l=u("2")*y("2");
75 l14(i)$(dat(i,"def")eq 2)..(x1*dat(i,"assets")+x2* dat(i,"income")+x3* dat(i,"
   debt")
76 +x4*dat(i,"size")+x5* dat(i,"job"))*sum(j,(m(j)))+d1("3","1")=g=u("2")*y("3");
77 l16(i)$( dat(i,"def")eq 2)..(x1* dat(i,"assets")+x2*dat(i,"income")+x3*dat(i,"
   debt")
78 +x4* dat(i,"size")+x5* dat(i,"job"))*sum(j,(m(j)))-d1("3","2")=l=u("3")*y("3");
79 l5(g)..u(g)-u(g-1)=g=1.13;
80 l4..sum((g),u(g))=e=5;
81 l6(i)$( dat(i,"def")eq 0)..sum(j,m(j-1))=e=5;
82 l7(i)$( dat(i,"def")eq 1)..sum(j,m(j-1))=e=5;
83 l8(i)$( dat(i,"def")eq 2)..sum(j,m(j-1))=e=5;
84
85   model test/all/;
86
87 solve test using minlp minimizing z;
88
89  h(i)$( dat(i,"def")eq 0)=(dat(i,"assets")*x1.l +dat(i,"income")*x2.l+
90 dat(i,"debt")*x3.l+dat(i,"size")*x4.l+dat(i,"job")*x5.l)<u.l("1");
91 k(i)$( dat(i,"def")eq 1)=(dat(i,"assets")*x1.l +dat(i,"income")*x2.l+
92 dat(i,"debt")*x3.l+dat(i,"size")*x4.l+dat(i,"job")*x5.l)<u.l("2");
93 d(i)$( dat(i,"def")eq 2)=(dat(i,"assets")*x1.l +dat(i,"income")*x2.l+
94 dat(i,"debt")*x3.l+dat(i,"size")*x4.l+dat(i,"job")*x5.l)<u.l("3");
95   display x1.l,x2.l,x3.l,x4.l,x5.l,d1.l,u.l,m.l,h,k,d;
```