CS-BIGS 2(1): 56-76 http://www.bentley.edu/csbigs/vol2-1/desbois.pdf

Introduction to Scoring Methods: Financial Problems of Farm Holdings

Dominique Desbois
INRA-SAE2 & SCEES, France

The goal of this paper is to provide a methodological introduction to "scoring" techniques on the basis of a case study about the financial problems of farm holdings. This case study is used as a teaching support for academic and vocational training in statistical data analysis.

First, this paper presents the estimation problem of financial risk, the constraints that in such a context the accounting data acquisition imposes, and the battery of the micro-economic criteria selected to measure the degree of insolvency of farm holdings. Secondly, the information provided by this battery of financial ratios is analyzed by way of multidimensional statistical techniques such as principal component analysis (PCA), and discriminant analysis. The framework provided by multidimensional tools such as PCA, discriminant analysis, and logistic regression procedures make it possible to show the methodological improvement carried out by such tools for this kind of micro-economic study. The results obtained are interpreted directly on the basis of outputs from the software used. Last, a measurement based on the ROC curve is provided in order to compare the classifying performances of linear discriminant analysis and logistic regression. This appraisal shows that the two methods are almost equivalent for individual prediction. A subset of the original data and SPSS code instructions are provided for training purposes.

This case study aims to offer a methodological introduction to the detection of financial risks applicable to farm holdings for analysts from public agencies (Regional Directorates) or professional and technical offices in charge of agriculture, as well as professional services specialized in technical and financial management (Chambers of Agriculture, Centers for Rural Economics). This case study is also currently used into the curriculum of the AgroParisTech engineering school for students completing their master degree.

KEY WORDS: scoring methods, financial risks, farm holding, principal component analysis, discriminant analysis, logistic regression, ROC curve.

Financial failures in agriculture: some issues

In a transition context resulting from Common Agricultural Policy (CAP) reforms (2000 Agenda, 2003 Reform) and the widening of the European Union to countries of Eastern and Central Europe, the problem of estimating financial risks in agriculture will gain an interest fostered by the multiplication of agricultural crises (for example dioxin, bovine spongiform encephalopathy, foot-and-mouth disease).

Indeed, one could note that during the 1980s the tightening of economic constraints in this sector caused a multiplication of financial failures. Hence, the assistance

Climatic Changes & Agricultural Commodities: Is it the time now, to take such risks?
Sources: Chicago Board of Trade, and CNRS

procedures accompanying the reorganization that arose from this trend constituted an important component of French national and European agricultural policies. The experience of this decade showed that financial recovery measures can be effective provided that preventive actions are quickly started, hence the need for having a method for early and fast detection of financial risks in agriculture (Colson & al., 1993).

The analysis of financial risks in agriculture is based on the concept of farm holding "viability" proposed by public authorities and professional agricultural organizations since the implementation of public policies oriented towards modernization then regulation of agricultural production. After payment of investment subsidies, price complements for some products and natural handicap allowances decided within the framework of public policies supporting agriculture, a viable farm must ensure the farmer an income equivalent to that of other socio-professional categories.

Viability can also be defined in a negative way by identification of the non-viable groups of farm holdings. From this point of view, in the context of family farming, studies undertaken in France during previous years make it possible to split schematically non-viable farm holdings into two groups. On the one hand, the first group brings together farm holdings of small economic size, little modernized thus little involved in debt but whose results are insufficient to ensure a correct remuneration of family work. On the other hand, in the second group, one finds farm holdings of intermediate and even large economic size, committed to a modernization process with a considerable debt, but without the results being sufficient to face at the same time financial liabilities and the private amount deducted for the support of the holder family. Thus the diagnosis of the financial risk on the one hand requires having criteria of viability but on the other hand also supposes the delimitation of subpopulations "at risk" within the universe of farms, two dimensions which will be encountered during this study.

From farm holding accounts, the "credit scoring" method aims at diagnosing in a preventive way the financial problems of the holdings. The basic idea is to select accounting ratios that in the short and medium term are predictive of financial problems. Once this selection of ratios and reference values are carried out, one can seek to combine the various judgments put forth according to the comparison of the farm holding results with established reference thresholds. Then by incorporating these judgments expressed in the form of marks via various processes of summation within a function, generally noted Z and denoting the "score function", the Z "score" as a synthetic indicator is a measure of financial risk.

Introduction to financial risk assessment

Several phenomena can contribute to destabilizing farms on the financial level:

- the decline in prices of the agricultural products;
- the rise of intermediate consumption costs;
- the increase in credit and the modification of financing rules.

This financial weakening of the farm holdings results particularly in:

- a growth in financial expenses;
- a fall in turnover;
- a recrudescence of incidents and delays in payments.

The availability of annual micro-economic data on European farms offered by the Farm Accounting Data Network (FADN) together with the use of multivariate statistical analysis techniques make it possible to create a global indicator for the financial position of the farm holdings.

According to Dietsch (1989), the indicator must make it possible to measure this financial position correctly, i.e.:

- to detect among the farms subjected to a diagnosis being carried out on the basis of standardized "clinical" picture, the symptoms of vulnerability observed on the holdings in difficulty;
- to detect financial difficulties during the current fiscal year but also before the failure in order to be able to recommend assistance or rectification measures;
- to preserve a good predictive capacity in the short-medium range (3 to 5 years);
- to organize a follow-up of the financial problems of farm holdings, starting from the accounting and financial data of the FADN to be estimated annually.

The use of a "score function" to diagnose financial problems goes back to Beaver's (univariate method,

1966) and Altman's (multivariate method, 1968) pioneering work applying the techniques of discriminant analysis to a battery of financial ratios in order to judge company viability. This approach consists in discriminating the companies in bankruptcy from other companies by selecting factors which reflect financial problems in order to propose a detection system of bankruptcy risk (Bardos, 1985). The construction of a discriminant function by linear combination of these determinants makes it possible to classify each farm according to its probability of membership of the failing farm group and thus to estimate the vulnerability or the financial good health of a farm. By applying this methodology to FADN accounting and financial data, one can assign a score to each farm, obtain from it a financial risk level to be allocated to each farm, and then estimate the size of the various risk groups within the population of farm holdings at the national or European level.

To judge the financial health of farm holdings or their relative vulnerability, it is necessary to have objective criteria. For business or industrial enterprises one typically uses their position defined by legal criteria: liquidation of goods, bankruptcy proceedings or selling off the business. However, compulsory liquidations for farms relate to only one small share of business sell-offs, the greatest part being realized under private agreements between the farmer and his main creditors. Thus, the legal position does not provide a precise measurement of financial crisis situations. In addition, selling off the business is not always the expression of a financial crisis.

It is thus advisable to substitute for the strictly legal position of economic bankruptcy that of insolvency, thus separating "healthy" farms from "failing" ones by a similar criterion or "proxy". Insolvency is defined as the situation in which a farm holding is unable to honor the obligations generated by existing debt, namely the payment of interests and the payment of loans. Indeed, current financial and banking practices compare the situation of insolvency to a situation of financial problems: even if the farm is not declared bankrupt, it cannot face its liabilities.

Once having retained this criterion for farm classification, two distinct topics can be explored concerning the measurement of financial problems:

- What are the main features which make it possible to distinguish insolvent farm holdings from solvent ones?
- To what extend is insolvency the expression of financial problems within the holding?

The answer to the first question can be addressed by analyzing economic and financial ratios over the current year when insolvency appears as well as in the years immediately preceding this appearance. Answering the second question supposes being able to evaluate the degree of permanence for the financial problems striking farms, and thus to have an accounting and financial follow-up of these farm holdings over a period of several years. Indeed, if some insolvent farm holdings temporarily experience financial problems, these situations do not lead necessarily to bankruptcy.

The sample which has been collected makes it possible to answer the first question by trying to identify the characteristics of farm holdings in difficulty on the basis of accounting ratio and compare them to those of healthy farm holdings. In so doing, it will be possible to financially characterize the situation of insolvency for these farm holdings. In order to answer the second question, it would be necessary to have a cohort over several years allowing a comparison of accounting and financial trajectories followed by these two subpopulations.

The data collection methodology

The collection of data for this type of studies proves to be very expensive, for technical reasons related to the harmonization of accounting practices of farm holdings in difficulty but also because of strict confidentiality rules imposed by the nature of accounting and financial information. Farm holdings in financial difficulty were identified on the basis of exhaustive investigations with the main creditors of the farmers such as local banks, the Agricultural Social Insurance Benefit Fund and agricultural cooperatives. The financial and accounting data result from the books of farm holdings held by specialized accounting centres of the studied counties, in particular the Centers for Rural Economics (CNCER, 1996). The selection process was carried out on the basis of available and reliable accounting data for accounting periods distant by at least two years (Blogowski & al., 1992).

The process of sampling being ad hoc rather than randomly stratified, one endeavored in each county to balance the constitution of the samples in order to have roughly the same structural characteristics in economic size and product orientation for the two groups studied: on one hand, farm holdings being considered as "healthy" (no cash-flow problems) and on the other hand, farm holdings being considered as "failing" (some cash-flow incidents). Thus, this sample can be regarded as reasonably representative of modernized professional farm holdings. However, as for all studies on farmers in difficulty, this sample suffers from a selection bias relative to the farm holding population of low economic size for which, below a contractually fixed threshold of turnover, there is no legal requirement to keep regular accounts. This selection bias is reinforced by the delay in the agricultural sector to use accounting tools and the frequent disaffection of accounting centers towards farms with badly degraded financial situations.

The measurement of insolvency: micro-economic and financial criteria

The sample used for this study comprises 1,260 farm holdings located in the counties *Eure* (27) - 348 farm holdings, *Nord* (59) - 282 farm holdings, *Orne* (61) - 333 farm holdings, and *Seine-Maritime-* (76) - 297 farm holdings, that are comparable in their dominant agricultural productions because for the majority they are specialized in field crops. The sub-samples are balanced in their size and the observation carried out covers the period from 1988 to 1994 (each farm holding is observed only during one year).

The structure of these samples is described by the following variables:

CNTY	county index;
DIFF	payment incident (1= healthy ; 2 = failing);
STATUS	legal status (1 = individual holder ; 0 = company);
HECTARE	agricultural area used – AAU (in hectares);
ToF	type of farming index;
OWNLAND	owned land (O = Yes ; N = No);
AGE	holder’s age (or youngest holder’s age);
HARVEST	harvest year concerned.

The battery of economic and financial criteria comprises 22 ratios selected according to the following topics. Note that EBITDA refers to Earnings Before Interest, Taxes, Depreciation and Amortization.

Capitalization

r1 total debt / total assets;
r2 stockholders' equity / invested capital;
r3 short term debt / total debt;
r4 short term debt / total assets;
r5 long and medium term debt / total assets;

Weight of the debt

r6 total debt / gross product;
r7 long and medium term debt / gross product;
r8 short term debt / gross product;

Liquidity

r11 working capital / gross product;
r12 working capital / (real inputs - financial expenses);
r14 short term debt / circulating assets;

Debt servicing

r17 financial expenses / total debt;
r18 financial expenses / gross product;
r19 (financial expenses+ refunding of long and medium term capital) / gross product;
r21 financial expenses / EBITDA;
r22 (financial expenses + refunding of long and medium term capital)/EBITDA;

Capital profitability

r24 EBITDA /total assets;

Earnings

r28 EBITDA / gross product;
r30 available income / gross product;
r32 (EBITDA - financial expenses) / gross product;

Productive activity

r36 immobilized assets / gross product;
r37 gross product / total assets.

Principal Component Analysis of financial ratios

In order to synthesize the information contained in this battery of financial ratios, a principal components analysis (PCA) was carried out on the basis of standardized variables (standardized or normalized PCA). Correlations between our variables and the two first principal components are displayed in Figure 1.

The *F1* axis represents nearly *47 %* of the accounting ratio dispersion. The active variables best correlated with the positive side of this axis are:

- two ratios representing the weight of the debt, the total debt / gross product ratio [*r6*] and the current liability / gross product ratio [*r8*];
- two ratios representing the debt service, the financial expenses / gross product ratio [*r18*] and the financial expenses / EBITDA ratio [*r21*];
- a ratio representing the capitalization, the total debt / total asset ratio [*r1*] constituting what is called the total debt ratio (TDR);
- a ratio expressing the need for liquidity, the current liability / circulating asset ratio [*r14*].

On the negative side of the *F1* axis, the best-correlated active variables are:

- two ratios concerning results, the available income / gross product ratio [*r30*] and the (EBITDA - financial expenses) / gross product ratio [*r32*];
- a ratio representing the weight of the debt, the stockholders' equity / invested capital ratio [*r2*];
- two working capital ratios, the working capital / gross product ratio [*r11*] and the working capital / (real inputs - financial expenses) ratio [*r12*].

The active variables least correlated with this axis are:

- two ratios of productive activity, the immobilized assets / gross product ratio [*r36*] and the gross product / total asset ratio [*r37*];
- the current liability / total debt ratio [r3].

The quantitative additional variables describing the structure of the sample, such as the holder's age [*AGE*] or the agricultural area used by the holding [*HECTARE*] are correlated at an insufficient level with the first two factors so that their projections on the first factorial plane cannot be interpreted.

The same applies to the set of additional logical variables coding the categories of the qualitative variables which describe the structure of the sample; we however note that the logical variables [*LDIFF0*] and [*LDIFF1*], coding each category of the qualitative variable indicating the occurrence or not of financial problems, are near the correlation circle (see Figure 1). Their correlation with the factorial axes, in particular the first, being sufficiently high, we can interpret their respective localizations on the *F1* axis like absence [*LDIFF0*] or presence [*LDIFF1*] indicators of financial problems for the farm holdings being located in the same area of the plan (*F1<0* semi-axis, respectively *F1>0* semi-axis).

This interpretation is confirmed by the use of the financial problem indicator [*DIFF*] as an illustrative variable in the projection of the individual observations on the first factorial plane: each farm holding is then located by the value coding the category (label '2' if there are financial problems, label '1' if not). Thus, one obtains a density graph (see Figure 2), which by the respective localization of the labels confirms the discriminatory power of the first factorial axis between the "healthy farm" group (label '1') and the "failing farm" group (label '2').

Note that the correlation computation between the *F1* axis and the coding indicators of the other qualitative variables does not show evidence of relationship with another structure variable, for the farm holdings. The projection of the other qualitative variables describing the structure of the farms does not show any segregation of the categories according to the first axis, thus validating the effort to obtain a balanced composition of the sample in terms of structure variables for the two groups studied.

Thus, the *F1* axis consists of an opposition between a group of ratios expressing debt importance in terms of capitalization, debt burden and liquidity (*F1>0*) and a group of ratios translating the importance of results and co-varying in opposite direction (*F1<0*). Hence, farm holdings having financial problems are characterized by an important debt with high values for their corresponding ratios and low values for the income ratios.

The *F2* axis represents nearly 17% of the ratio variability and is interpreted in terms of debt structure and productive efficiency. Indeed, the ratios best correlated with the *F2* axis are, in terms of productive efficiency:

- in the *F2>0* half-plane, the productive activity ratio [*r37*], gross product divided by the total asset, and the capital profitability ratio [*r24*], EBITDA divided by the total asset;
- in the opposite *F2<0* half-plane the [*r36*] ratio, immobilization divided by gross product, the quasi inverse function of [*r37*], the previous productive activity ratio;

and in terms of debt structure:

- the capitalization ratios in the *F2>0* half-plane, namely [*r3*], current liability / total debt, and [*r4*], current liability / total asset;
- opposed to mainly the [*r7*] ratio, the burden of long and medium term loans, and to the [*r12*] and [*r11*] ratios expressing the relative importance of working capital.

Thus, the *F2* axis represents the relationship between a greater or smaller productive efficiency and a higher or smaller amount of current liability compared to the total debt. For the most intensive farm holdings, which have more important requirements for working capital than more extensive farm holdings, the absence of available equities generates a relatively more important debt in the short term to meet this need.

The cloud of the "failing" farm holdings appears more dispersed according to the *F2* axis than the cloud of the "healthy" farm holdings in the first factorial plane; this consolidates this interpretation.

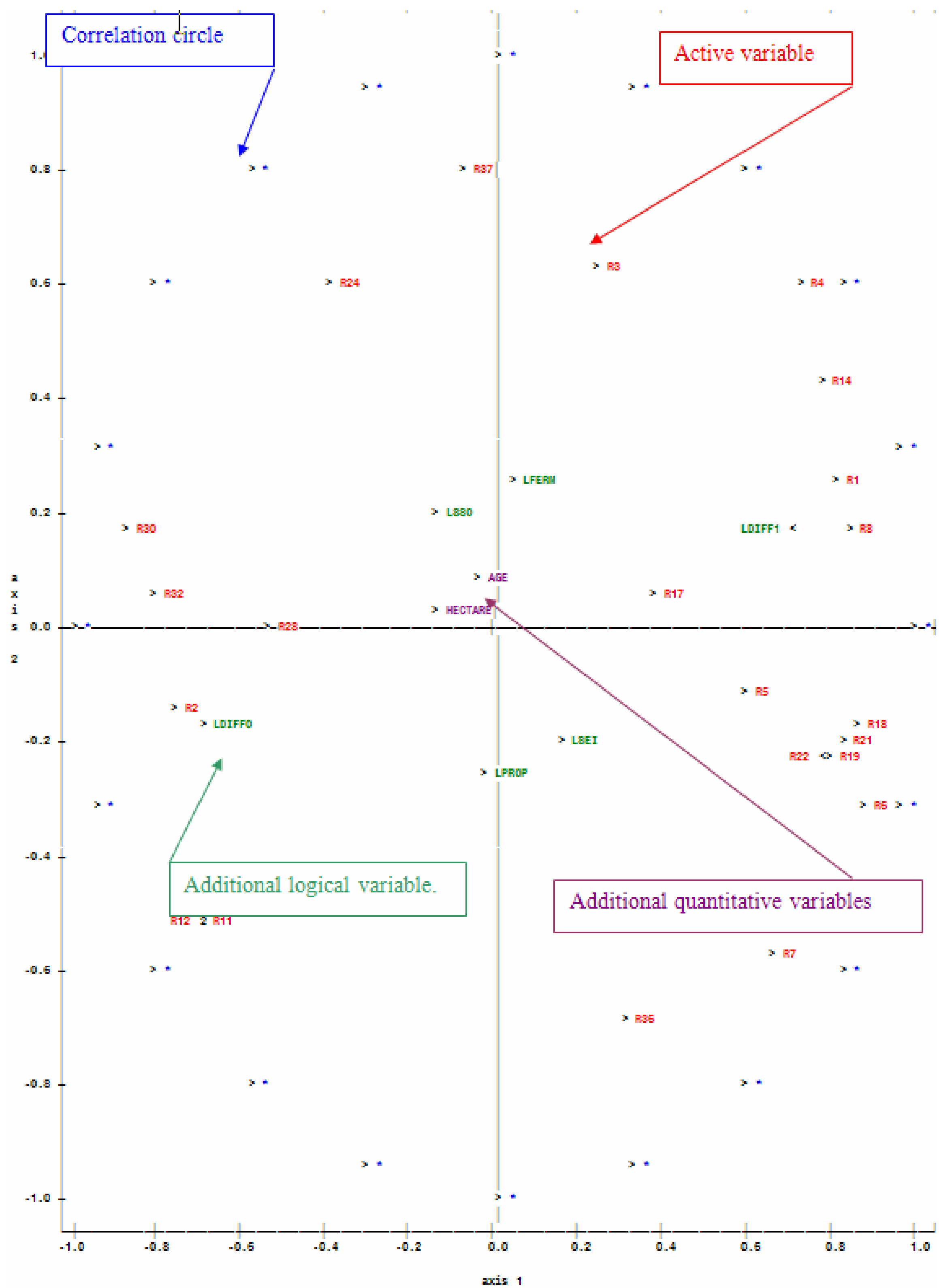

Figure 1. Projection of the financial ratios on the first factorial plane of the normalized PCA.

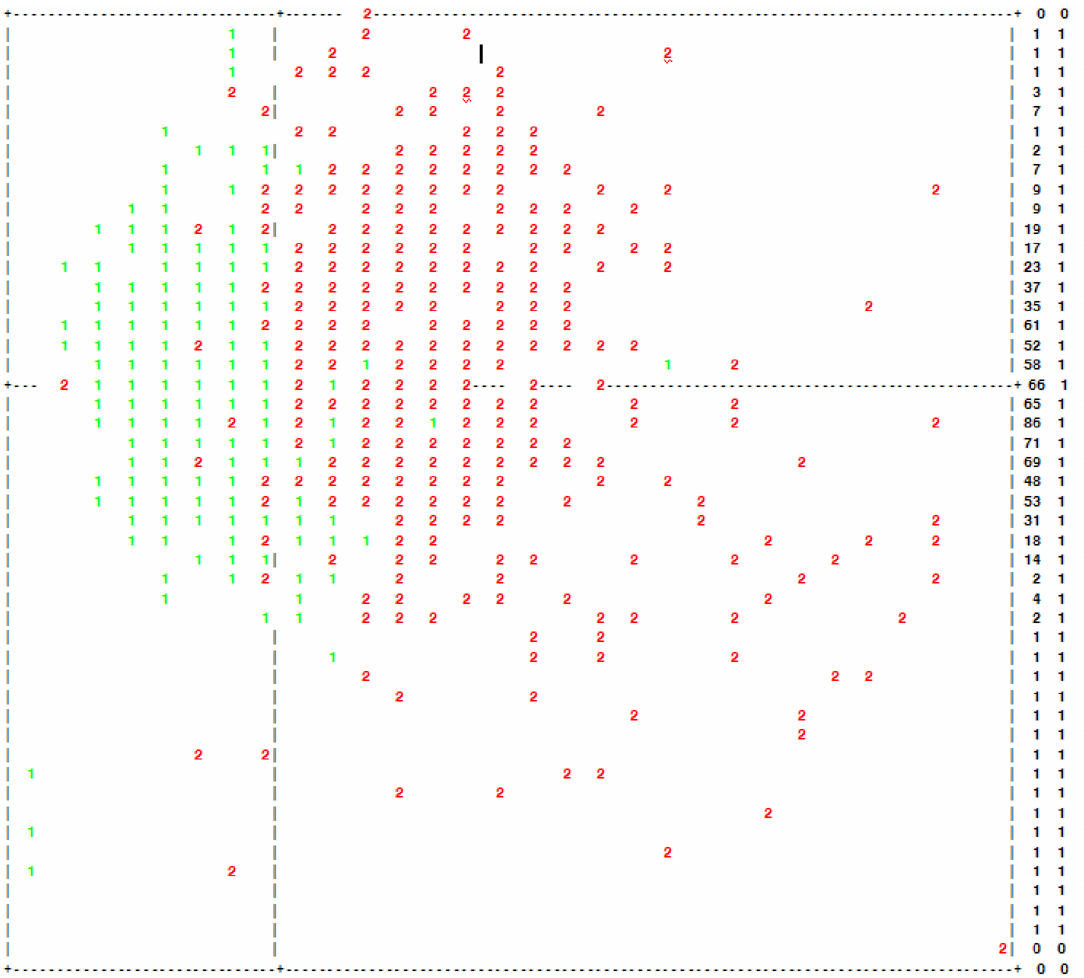

Figure 2. Plot of the farm holdings in the first factorial plane of the normalized PCA based on financial ratios with illustrative variable LDIFF (1="healthy"; 2="failing").

```
THE EIGENVALUES       VAL(1)=  10.26466
------------------------------------------------------------------------------------------
|NUM | EIGENVALUE | PERCT.| CUMUL |DELTA  |*| EIGENVALUE HISTOGRAM
------------------------------------------------------------------------------------------
|  1 |   10.26466 | 46.658| 46.658|*******|*|***************|***************|***************|***************
|  2 |    3.69294 | 16.786| 63.444| 29.871|*|***************|*******
|  3 |    2.46976 | 11.226| 74.670|  5.560|*|**************
|  4 |    1.58131 |  7.188| 81.858|  4.038|*|*********
|  5 |    1.19154 |  5.416| 87.274|  1.772|*|*******
|  6 |     .82654 |  3.757| 91.031|  1.659|*|*****
|  7 |     .59810 |  2.719| 93.749|  1.038|*|***
|  8 |     .38601 |  1.755| 95.504|   .964|*|**
|  9 |     .25027 |  1.138| 96.642|   .617|*|*
| 10 |     .16708 |   .759| 97.401|   .378|*|*
| 11 |     .13706 |   .623| 98.024|   .136|*|*
| 12 |     .12747 |   .579| 98.603|   .044|*|*
| 13 |     .09370 |   .426| 99.029|   .153|*|*
| 14 |     .05840 |   .265| 99.295|   .160|*|
| 15 |     .04450 |   .202| 99.497|   .063|*|
| 16 |     .04033 |   .183| 99.680|   .019|*|
| 17 |     .02259 |   .103| 99.783|   .081|*|
| 18 |     .01464 |   .067| 99.850|   .036|*|
| 19 |     .01202 |   .055| 99.904|   .012|*|
| 20 |     .01047 |   .048| 99.952|   .007|*|
| 21 |     .00637 |   .029| 99.981|   .019|*|
| 22 |     .00424 |   .019|100.000|   .010|*|
```

Histogram of the eigenvalues, extract from the listing of ADANCOMP procedure, principal component analysis, ADDADSAS software (cf. [Flavigny and Lebeaux, 1998])

Figure 3. Selection of eigenvalues higher than 1.

The interpretation of the higher rank axes (five axes correspond to eigenvalues greater than one, see Figure 3) proves to be more delicate:

- the F3 axis seems to characterize a sub-group of farm holdings where the productive efficiency ([*r24*] ratio of capital profitability and [*r28*], result ratio of the EBITDA to the gross product) is positively correlated with a higher debt in the medium and long term ([*r5*] ratio) and negatively with a lower debt in the short term ([*r3*] ratio);
- the *F4* axis seems specific to a sub-group characterized by an important EBITDA [*r28*] performing well at the price of strong fixed assets compared to their gross product [*r36*];
- finally, the *F5* axis appears to be specific to farm holdings which have significant financial expenses compared to the amount of their debt [*r17*]; that can be a sign of extremely jeopardized financial situations.

The eigenvalues corresponding to the axes of rank higher than 5 are lower than the unit (cf. Graph 3 of eigenvalues) and do not offer any interesting correlation (i.e. higher than 0.3 in absolute value) in terms of factorial projection. One thus chooses to regard the corresponding share of inertia (13%) as residual.

On the basis of the battery of financial ratios studied, the subspace of the first five principal components gathers more than *87* % of the total inertia and thus constitutes a good summary of the sample variability. The first factorial plane gathers more than *63* % of the total inertia. Considering the respective positions of the healthy and failing groups of farm holdings on the first factorial plane, we see that the *F1* axis could be used as an index of classification (see Figures 2 and 4).

Indeed, the average coordinate of the "healthy" group on the *F1* axis is approximately equal to $\mu_0 \approx -0.6754$ while the average coordinate of the "failing" group is approximately equal to $\mu_1 \approx 0.7266$.

A geometrical rule of classification for a farm holding $\boldsymbol{i}_0$ which does not belong to the training sample consists in positioning the farm holding in relation to the "pivot point", i.e. the median point of the segment linking the two group barycenters on the *F1* axis, that is to say

$$\frac{\mu_0 + \mu_1}{2} = 0.0256:$$

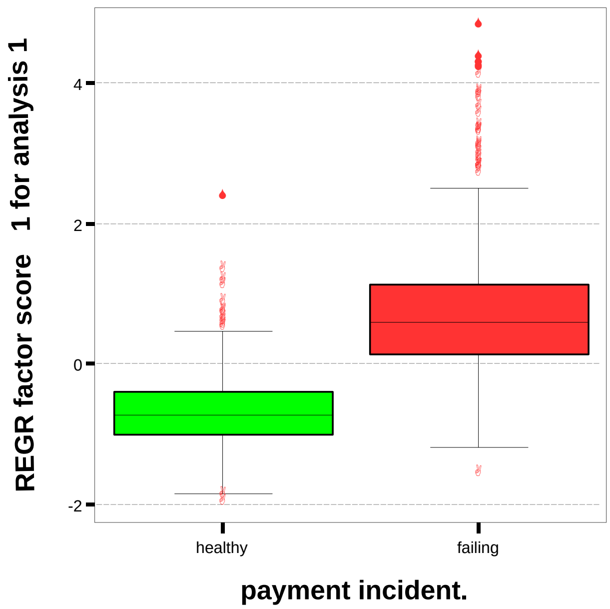

Figure 4. Distribution of the two sub-groups according to the first factorial axis.

- if $F1(\boldsymbol{i}_0) < 0.0256$, then the management of the farm holding is declared 'not at risk' (RISK=0);
- if $F1(\boldsymbol{i}_0) > 0.0256$, then the management of the farm is declared 'risky' (RISK=1);
- if $F1(\boldsymbol{i}_0) = 0.0256$, then the assignment of the farm holding to one of the two groups can be carried out on the basis of random sampling.

The application of this rule to the training sample (see Figure 5) gives us a satisfactory percentage of correctly-classified "healthy" farm holdings (*91.7* %), however it is less satisfactory for the "failing" farm holdings (*79.6* %). Even if this rule does not take account of the dispersion differences between the two groups, it is better than a rule based on a single ratio. Indeed, let us take as pilot ratio the debt ratio [r1] because, on the one hand, it is very much used by management professionals who know it as the Total Debt Rate (TDR) and, on the other hand, it is well correlated with the *F1* axis (*r=0.81*). With a *28* % TDR median for French farm holdings, we correctly classify the near total of the "failing" farm holdings (*99* %) but nearly 66 % of the "healthy" farm holdings are also classified as "indebted" (i. e. with a debt greater than the TDR median).

Using the third quartile (see Figure 6) rather than the median, that is to say a 48 % TDR bound, we improve our percentage of correctly-classified among the "healthy" ones to *74.1* %, but our percentage of classification errors climbs to *11.4* % within the group of "failing" farm holdings.

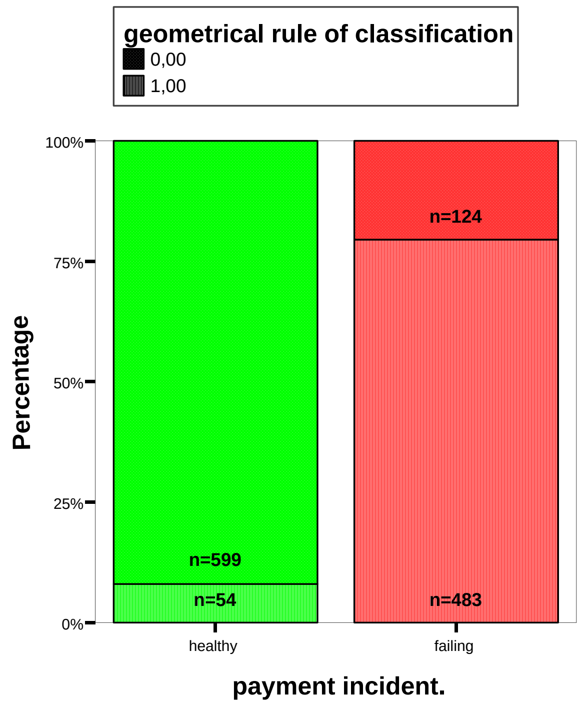

Figure 5. Classification results for each sub-group, according to the first factorial coordinate.

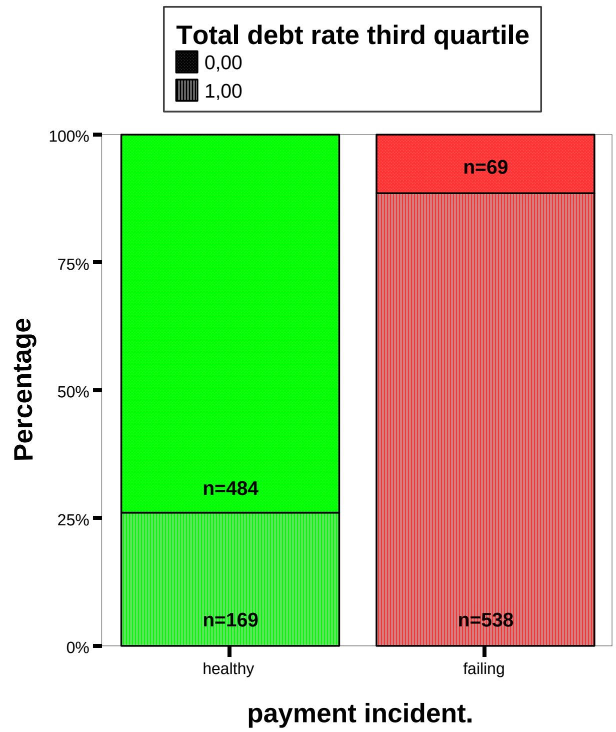

Figure 6. Classification results according to the total debt rate for each subpopulation.

Using the information provided by the sample for the TDR and the same geometrical rule of classification as previously (the pivot point rule), we obtain, with the "indebted" indicator, better correctly-classified percentages in the two groups (*86.4* % for the "healthy" ones and *77.8* % for the "failing" ones), however the classification performances remain lower than those based on the *F1* axis, which carries out a more pertinent compromise, as a linear combination of the whole set of accounting ratios with respect to the financial failure phenomenon.

Thus, depending on the institutional position (borrower, lender or regulator), there will be an interest in privileging one or another criterion. The banker will privilege the median TDR criterion because it minimizes the incurred risk of non-repayment. On the other hand such a criterion can appear abusive to a potential borrower who would have to support the cost of the classification errors made on the "healthy" farm holdings. Provided the costs are symmetrical, a regulator will privilege a criterion having similar error rates for each contracting party, even a global rate error on the whole set to classify.

However, if the first principal component proves to be an acceptable discriminating factor for a regulator privileging general interests, the use of a technique specifically designed to maximize the overall score of correctly-classified holdings can enable us to improve classification performances. A discriminant analysis makes it possible to achieve this goal:

- by seeking directions of projection, which maximize inter-group variability and minimize intra-group variability;
- by taking into account the relative variability of each group;
- by proposing a probabilistic model to minimize the risk of classification error in the assignment process of individuals.

Discriminant analysis with financial ratios

In order to discriminate as well as possible between the two groups of farm holdings indexed according to their financial problems, we will use a discriminant analysis on the basis of the most relevant financial ratios, to predict the membership of each farm holding to the group defined by the two values of the "payment incidents" categorical variable:

- if no incident of payment intervened (*DIFF='0'*), the farm holding is regarded as financially "healthy";
- on the other hand (*DIFF='1'*), the farm holding is regarded as financially "failing".

On the basis of financial ratio combinations used as exogenous (explanatory) variables in the analysis, the discriminant analysis builds discriminant functions in order to assign farm holdings to one of the preset groups, either using a geometrical rule, or using a Bayesian probabilistic rule.

In order to validate the results obtained, we use a cross-validation procedure[1], which consists in carrying out a classification for each individual in the sample, on the basis of a discriminant linear function obtained with the other individuals of the sample. This is equivalent to carrying out as many estimates as they are individuals in the sample: each classification is carried out on the basis of a training sample consisting of the $n-1$ remaining individuals. According to this procedure, each classified individual is used as a test-sample for the correctly-classified percentage computation

Table 1. Classification results for each subpopulation, estimated on the basis of a cross-validation procedure.

			Predicted group membership		
		DIFF	healthy	failing	Total
Original	Count	healthy	605	48	653
		failing	98	509	607
	%	healthy	92.6	7.4	100
		failing	16.1	83.9	100
Cross-validated	Count	healthy	605	48	653
		failing	99	508	607
	%	healthy	92.6	7.4	100
		failing	16.3	83.7	100

Notes. a. Cross-validation is carried out only for observations included in the estimation. In cross-validation, each observation is classified according to functions derived from the other observations. b. 88.4% of the original observations are correctly classified. c. 88.3% of the cross-validated observations are correctly classified.

With this classification procedure[2], one obtains better percentages of correctly-classified holdings for the "failing" farm holding group, that is to say *83.7 %* with the *D* discriminant linear function compared to *78.6 %* with the *F1* first principal component of the PCA. For the "healthy" farm holding group, the results are practically equivalent, that is to say *92.6 %* for discriminant analysis compared to *92.5 %* with the PCA.

From an exploratory point of view, we use an algorithm which selects the most relevant financial ratios to build the *D* discriminant linear function step by step and discriminates among the two farm holding groups defined by the respective values of the financial problem indicator: *g0*, the group of farm holdings considered as financially "healthy" and *g1*, the group of farm holdings considered as financially "failing". The selection of the financial ratios is based on an estimate of their discriminating capacity. The selection criterion which we use to evaluate the discriminating capacity is the multivariate Wilks lambda, computed on the whole set of financial ratios that make up the *D* function by linear combination. The Wilks lambda can be computed for a single variable (univariate Wilks lambda) making it possible to consider the discriminating capacity of each financial ratio. Its value varies between *0* (infinite discriminating capacity) and *1* (no discriminating capacity): low values indicate large differences between groups and conversely high values indicate small differences between groups.

Table 2. Equality test between the two sub-groups for each financial ratio.

Predicting Ratio	Wilks Lambda	F	dof 1	dof 2	P-level
r1	0.580	909.310	1	1258	0.000
r2	0.601	836.696	1	1258	0.000
r3	0.930	94.186	1	1258	0.000
r4	0.625	753.274	1	1258	0.000
r5	0.821	274.251	1	1258	0.000
r6	0.696	548.662	1	1258	0.000
r7	0.853	216.115	1	1258	0.000
r8	0.659	651.260	1	1258	0.000
r11	0.692	560.774	1	1258	0.000
r12	0.679	595.992	1	1258	0.000
r14	0.601	834.464	1	1258	0.000
r17	0.891	153.789	1	1258	0.000
r18	0.665	634.839	1	1258	0.000
r19	0.779	356.637	1	1258	0.000
r21	0.779	357.187	1	1258	0.000
r22	0.828	261.308	1	1258	0.000
r24	0.956	58.160	1	1258	0.000
r28	0.848	226.257	1	1258	0.000
r30	0.681	588.130	1	1258	0.000
r32	0.669	622.098	1	1258	0.000
r36	0.990	12.243	1	1258	0.000
r37	0.997	3.491	1	1258	0.062

[1] The "re-substitution method", classifying the observations from the training sample used for parameter estimation, leads to an over-estimation bias in the correctly-classified percentage. In case of relatively small samples, cross-validation or bootstrap methods can be used in order to overcome the difficulty in drawing validation samples independently from the learning sample.

[2] The classification procedure used to obtain these results assumes the equality of the local variance-covariance matrices for each group. In this case, the classification rule is linear and the obtained partitions between the groups are hyper-planes (here for two groups, a line into a plane).

Using these results, let us recall that for two groups, the *F* statistic, ratio of the between-groups variance on the intra-groups variance, is equivalent to the square of Student's *T* statistic for groups with equal variance.

We can check on Table 2 that the univariate Wilks lambda, measuring the discriminating capacity of each ratio, is an inverse function of the *F* statistic, that is to say, for example for r_1 :

$$\Lambda = \frac{1}{[1 + F/(n-2)]} = \frac{1}{[1 + 909.310/(1258)]} \approx 0.580$$

Thus, this table states that the ratio with the strongest discriminating capacity is definitely the [*r*1] ratio, the total debt rate commonly used by the professional community of financial management. However, other ratios also have a strong discriminating capacity such as [*r2*] the owned percentage of capital stocks, which is also a debt weight ratio, and [*r14*], the current liability related to the circulating asset, which is a working capital ratio.

In the multivariate case, the total sum of squares and co-product matrix *T* is broken up into the between-groups sum of squares and co-product matrix *B* and into the within-groups sum of squares and co-product matrix *W* as in the univariate case: $T = B + W$.

With two groups, we can obtain only one discriminant function. The lambda of the multivariate Wilks test associated with this discriminating linear function *D1*, is given by the ratio of the determinants of the within-groups variance-covariance matrix *W** and the between-groups variance-covariance matrix *B**:

$$\Lambda = \frac{\det(W^*)}{\det(B^*)}$$

Table 3. Value of the multivariate Wilks lambda associated with the discriminant linear function.

Discriminant Function	Wilks Lambda	Chi-square	dof	P-level
D1	0.412	1104.954	22	0.000

The statistical function associated with the Wilks lambda follows a χ^2 distribution with $p(K-1) = 9$ degrees of freedom under the null hypothesis of mean equality for K=*2* groups with $p = 9$ variables introduced into the model (see Table 3). The p-value, less than one in a thousand, leads to the conclusion that the average scores of the two groups differ significantly according to the discriminant linear function (Table 3).

Another measure of the discriminating capacity of the discriminant linear function is the canonical correlation coefficient between the subspace generated by the indicating variables of the two groups and the subspace generated by the linear combinations of financial ratios (see Table 4).

Table 4. Eigenvalue and canonical correlation coefficient associated with the discriminant linear function.

Discriminant Function	Eigen Value	Variance %	Cumulative %	Canonical Correlation
D1	1.426	100	100	0.767

As in an analysis of variance, the variance homogeneity assumption is critical to carrying out the estimates: if the groups have similar local variance-covariance matrices, the discriminant functions will be estimated on the basis of a common intra-group variance-covariance matrix and the discriminant functions will be linear; on the other hand, if the groups do not have similar variance-covariance matrices, the discriminant functions will be estimated on the basis of *K* variance-covariance matrices, i.e. for each group, and they will be quadratic.

Table 5. Natural logarithm of the within-groups variance-covariance matrix determinants and result of the multivariate Box's test.

Payment Incident Indicator	Rank	Determinant Logarithm
healthy	22	-137.658
failing	22	-99.610
Pooled Within-Groups	22	-106.509
Box's M		16,128.311
Approximate F		62.605
dof 1		253
dof 2		4,769,909
p-value		0.000

Box's *M* statistic, based on the natural logarithm of the determinant of each local variance-covariance matrix, makes it possible to build a multivariate test for the comparison of variance-covariance matrices: the probability level associated with the Fisher-Snedecor distribution of a *F* ratio value computed on the basis of Box's M statistic, less than one in a thousand, leads to the rejection of the null hypothesis of equality of the variance-covariance matrices for the two groups. This statistical test confirms the conclusions drawn from the graphic projection of the individuals from the two groups in the first factorial plane. However, like its univariate counterpart, Box's test is very sensitive to the normality assumption on the distributions. This normality not having being satisfied because of the asymmetrical character of the various ratio distributions, it is appropriate to regard the result of the test as an additional index but not as a proof of the ratio distribution heteroscedasticity related to these two groups of farm holdings.

As in multiple regression, the stepwise selection process takes into account the correlations between discriminating variables (Table 6).

Table 6. Stepwise selection of the predictors discriminating the two groups of farm holdings according to the criterion minimizing Wilks' lambda.

Step	Predictors Included	Wilks Lambda		Exact F			
		Statistic	dof 1	Statistic	dof 1	dof 2	P level
1	r1	0.580	1	909.310	1	1258	0.000
2	r32	0.502	2	624.223	2	1257	0.000
3	r14	0.467	3	478.407	3	1256	0.000
4	r17	0.453	4	378.131	4	1255	0.000
5	r2	0.445	5	312.664	5	1254	0.000
6	r3	0.437	6	268.833	6	1253	0.000
7	r36	0.429	7	237.793	7	1252	0.000
8	r21	0.423	8	212.917	8	1251	0.000
9	r7	0.422	9	190.439	9	1250	0.000
10	r18	0.419	10	173.322	10	1249	0.000
11		0.420	9	192.062	9	1250	0.000

Notes: For Wilks' Lambda, dof2 = 1, dof3 = 1258. In step 11, r1 was excluded.

For instance, the total debt rate [*r1*] ratio, introduced in the first step because acting as the most discriminating variable, is eliminated in the last step of the selection process because it can be sufficiently reconstituted in an approximate way as a linear combination of the other ratios introduced into the preceding steps.

From this follows the linear discriminant function *D1* (Fisher's score), which is expressed with "standardized coefficients" as a linear combination of the standardized discriminating variables (see Table 7).

Table 7. Standardized coefficients of the discriminant function allowing computing each farm holding score.

Predictors	Discriminant Function D1
stockholders' equity / invested capital [r2]	-0.568
short-term debt / total debt [r3]	-0.413
long and medium-term debt / gross product [r7]	-0.311
short-term debt / circulating asset [r14]	-0.200
financial expenses / total debt [r17]	0.111
financial expenses / gross product [r18]	0.304
financial expenses / EBITDA [r21]	0.455
(EBITDA - financial expenses) / gross product [r32]	0.473
immobilized assets / gross product [r36]	0.499

That is to say:

$$D_1(i_0) = -0.568Z_{r2} - 0.413Z_{r3} - 0.311Z_{r7} - 0.2Z_{r14} + 0.111Z_{r17} + 0.304Z_{r18} + 0.455Z_{r21} + 0.473Z_{32} + 0.499Z_{r36}$$

where Z_{r_j} is the standardized value of the j^{th} ratio, r_j.

The score value $D_1(i_0)$ derived from this equation can be located with respect to the average scores of the groups in the following way: the farm holding i_0 is assigned to one group on the basis of the geometrical rule using the median point of the segment linking the barycenters (weighted means, see Table 8) of the two groups, the corresponding D_1-coordinate of the median point giving the threshold value $c = 0.086$.

- if the score $D_1(i_0) < c$, then the farm holding is classified as "healthy";
- if the score $D_1(i_0) > c$, then the farm holding is classified as "failing";
- if the score $D_1(i_0) = c$, then the farm holding is assigned to the group by random drawing.

Table 8. Weighted mean value of the linear discriminant function for each group of farm holdings.

Payment incident	Function D1
healthy	-1.133
failing	1.219

Note: The unstandardized Canonical Discriminant Functions are estimated at group barycenters.

To carry out the estimated classification of a farm holding which does not belong to the sample-test, we will use for convenience reasons the linear combination issued from the original variables, with unstandardized coefficients (Table 9):

Table 9. Unstandardized coefficients issued from the original variables, to be used for the score computation.

Predictors	Function D1
stockholders' equity / invested capital [r2]	-2.542
short-term debt / total debt [r3]	2.398
long and medium-term debt / gross product [r7]	1.099
short-term debt / circulating asset [r14]	0.604
financial expenses / total debt [r17]	18.702
financial expenses / gross product [r18]	-6.798
financial expenses / EBITDA [r21]	-0.622
(EBITDA - financial expenses) / gross product [r32]	-4.163
immobilized assets / gross product [r36]	0.192
(Constant)	-0.444

That is to say:

$$D_1(i_0) = -2.542r2 + 2.398r3 + 1.099r7 + 0.604r14 + 18.702r17$$
$$-6.798r18 - 0.622r21 - 4.163r32 + 0.192r36 - 0.444$$

One can then consult the distribution of values of the discriminant scores according to the groups (see Figure 7):

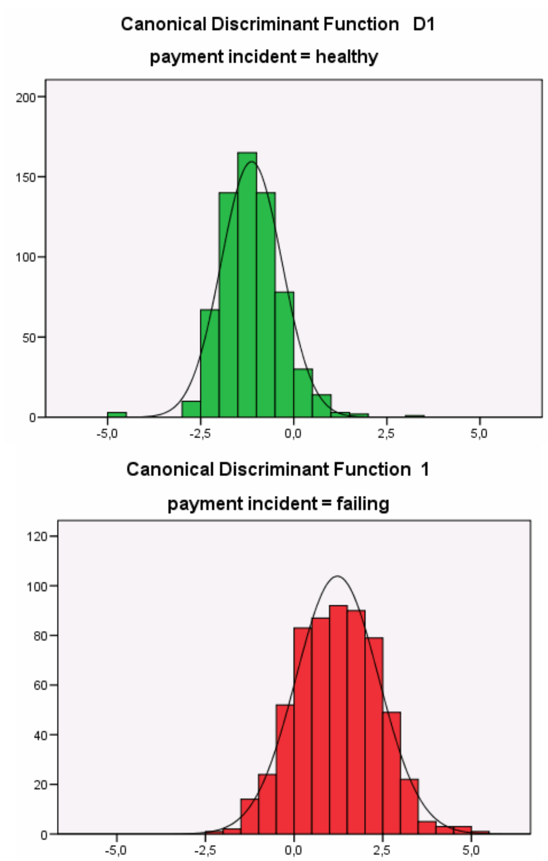

Figure 7. Distribution of the discriminant scores for "healthy" (left) and "failing" (right) farm holdings.

It can be noted that the discriminant scores are distributed according to a roughly Gaussian distribution around the average score of each group. Compared to the "healthy" group (standard deviation is equal to 0.82), the "failing" group presents a larger dispersion (standard deviation is equal to 1.17).

Box-plots (Figure 8) make it possible to compare the two distributions of discriminating scores noting their symmetry, which was not the case for the original variables.

Using a quadratic discriminant function (classification on the basis of the local variance-covariance matrices for each group), one obtains almost identical results.

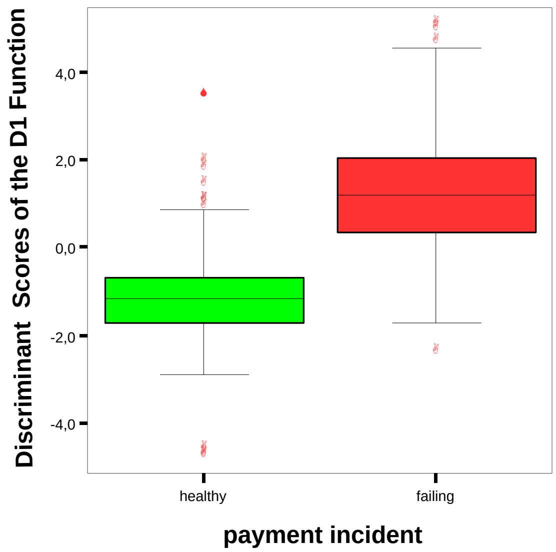

Figure 8. Comparison of the two distributions of scores, "healthy" versus "failing".

However with this option which takes into account the variance discrepancy between the two groups, a slight improvement of the total correctly-classified percentage can be observed. The "failing group" improves its correctly-classified percentage by 2 % while the "healthy" group is slightly less correctly-classified (-0.6 %), see Table 10:

Table 10. Classification results based on the quadratic discrimination rule.

	Payment incident	Predicted Group Membership healthy	failing	total
Count	healthy	601	52	653
	failing	87	520	607
%	healthy	92.0	8.0	100
	failing	14.3	85.7	100

Note: 89.0% of the original observations are correctly classified.

Logistic Modelling of the Financial Risk

As we have seen, discriminant analysis makes it possible to classify holdings into two groups "healthy" and "failing". However, this method relies on an assumption multivariate normality for the discriminating variables and on an assumption of equality for the variance-covariance matrices of the two groups.

These two hypotheses not being satisfied in many empirical situations, let us introduce the model of binomial logistic regression, technique allowing to estimate the probability that a particular event occurs. Indeed, this specific regression model requires assumptions which are much weaker than discriminant analysis.

In logistic regression, we seek to directly estimate the probability that an event occurs. In a univariate context with only one explanatory variable X, a logistic regression models the probability of an event D (financial failure) as follows:

$Prob(D) = \dfrac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}$ or in an equivalent way:

$$Prob(D) = \frac{1}{1+e^{-(\beta_0+\beta_1 X)}}$$

where β_0 and β_1 are the logistic regression parameters estimated from the data, typically via maximum likelihood.

In a multivariate context with several descriptors, the logistic model is formulated in a similar way, that is to say: $Prob(D) = \dfrac{1}{1+e^{-Z}}$

where **Z** is a linear combination formed on the basis of the ***p*** explanatory variables X_j

$$Z = \beta_0 + \beta_1 X_1 + \cdots + \beta_j X_j + \cdots + \beta_p X_p .$$

The probability that the event D does not occur (event $\overline{D}$) is derived using the following probability axiom: $Prob(\overline{D}) = 1 - Prob(D)$.

The Log-ratio or logit[3] is equal to the logarithm of the ratio of the probability $p(i)$ that a farm holding *i* is financially failing over the probability $1-p(i)$ that it is financially healthy:

$$Z(i) = \log\left(\frac{p(i)}{1-p(i)}\right) = \log\left(\frac{Prob(i \in D)}{Prob(i \in \overline{D})}\right)$$
$$= \beta_0 + \beta_1 X_1(i) + \cdots + \beta_j X_j(i) + \cdots \beta_p X_p(i).$$

Each logistic regression coefficient β_j is then interpreted as the marginal variation of the log-ratio, i.e. the relative variation in Z caused by the increment by one unit for the respective predictor X_j (when other variables are held constant).

It follows that the exponential of the parameter β_j measures the marginal variation of the Odds Ratio (OR), i.e. when the value of X_j increases from x_j^i for farm holding *i* to $x_j^{i'} = x_j^i + 1$ for farm holding *i'*:

$$OR_j(i'/i) = \frac{p(i')/[1-p(i')]}{p(i)/[1-p(i)]} = e^{\beta_j}.$$

Recall that logistic regression models are typically estimated via maximum likelihood. The likelihood of a model is defined as the probability of the observations, taking into account the values estimated for the parameters of the model. According to this criterion, the best model is that which will have the highest such probability. To facilitate numerical computations in the search for a maximum, we use the concept of deviance, a decreasing function of probability equal to – 2 times the logarithm of likelihood (LL): if the model fits the data perfectly, the likelihood equals 1 and the deviance of the model is then null. The deviance minimum of a model corresponds to its maximum likelihood.

Table 11. Deviance of the minimal logistic model: History of the iterations

Iteration		-2log-likelihood	Coefficients/constant
Step 0	1	1745.051	-0. 073
	2	1745.051	-0. 073

Note: The estimate was stopped with step 2 because the estimates of parameters changed less than 0.001.

The minimal logistic regression model comprises one parameter, the constant β_0 : $Prob(D) = \dfrac{1}{1+e^{-\beta_0}}$.

Thus, the maximum likelihood estimator leads to the following estimate of the constant: $\hat{\beta}_0 = \ln\left(\dfrac{\hat{\pi}_0}{1-\hat{\pi}_0}\right) \approx \ln\left(\dfrac{0.482}{0.518}\right) \approx -0.073$ (see Table 11). The probability π_0 to be a financially failing farm holding is estimated by the percentage of failing farm holding in the sample: $\hat{\pi}_0 = \dfrac{n_D}{n} = \dfrac{607}{1260} \approx 0.482$.

The computation of the minimal model deviance, i.e. the total deviance, uses the estimate of the constant in the computation of the log-likelihood:

$$\Delta_0 = -2 * LL_0 = -2 * n\left[\hat{\pi}_0 \ln\left(\frac{\hat{\pi}_0}{1-\hat{\pi}_0}\right) + \ln(1-\hat{\pi}_0)\right] \approx 1745.051.$$

In a way similar to the sums of squares in linear regression, the total deviance is equal to the sum of the deviance of the model and the residual deviance:

$$\Delta_0 = \Delta_M + \Delta_R.$$

[3] The "logit" term was forged by the biometrician Joseph Berkson (1899-1982), introducing logistic modelling, by analogy with the modelling "probit" developed from 1934 by Chester Bliss (1899-1979). Logistic regression was introduced for the first time in econometrics by McFadden (1974).

Table 12. Stepwise estimate of the logistic model parameters.

		B	S.E.	Wald	dof	Signif.
Step 1	r1	8.590	.466	339.772	1	.000
	Constant	-4.890	.269	331.614	1	.000
Step 2	r1	7.380	.492	224.788	1	.000
	r32	-10.172	.998	103.923	1	.000
	Constant	-1.586	.390	16.505	1	.000
Step 3	r1	5.284	.550	92.433	1	.000
	r14	2.437	.304	64.338	1	.000
	r32	-8.031	1.035	60.264	1	.000
	Constant	-2.382	.424	31.617	1	.000
Step 4	r1	5.495	.575	91.447	1	.000
	r14	2.389	.309	59.848	1	.000
	r17	22.166	4.217	27.624	1	.000
	r32	-7.475	1.054	50.323	1	.000
	Constant	-3.996	.546	53.589	1	.000
Step 5	r1	6.027	.605	99.301	1	.000
	r14	2.295	.313	53.657	1	.000
	r17	22.591	4.293	27.698	1	.000
	r32	-6.943	1.081	41.260	1	.000
	r36	.685	.187	13.342	1	.000
	Constant	-5.223	.659	62.775	1	.000
Step 6	r1	5.950	.611	94.713	1	.000
	r12	.949	.318	8.911	1	.003
	r14	3.358	.488	47.273	1	.000
	r17	24.233	4.380	30.610	1	.000
	r32	-7.306	1.104	43.770	1	.000
	r36	.610	.188	10.498	1	.001
	Constant	-6.168	.746	68.334	1	.000

Test of the logistic regression coefficients

In order to test the H0 hypothesis of nullity of the coefficient β_j, we can use for large samples the Wald statistic: $w_j = \left[\frac{\hat{\beta}_j}{\hat{\sigma}_{\beta_j}}\right]^2$ which is equal to the squared ratio of the estimate over the standard error of the estimate. Indeed, under the null assumption H0, the Wald statistic asymptotically follows a chi-squared distribution with one degree of freedom. Thus for the ratio *r1*, reporting the coefficient estimate displayed in the column labeled B (Table 12) to its standard error (S. E. column) and squaring the result, we obtain in the first step the corresponding value of the Wald statistic, i.e.:

$$w_1 = \left[\frac{\hat{\beta}_1}{\hat{\sigma}_{\beta_1}}\right]^2 = \left[\frac{8.59}{0.466}\right]^2 \approx 339.772.$$

This value of the Wald statistic having a very low probability of occurrence in a chi-squared distribution (lower than 1 out of ten thousand according to the Signif. column), we are led to reject the null hypothesis of a zero value for this coefficient.

However, when the value of the estimated coefficient is large, its standard error of estimate becomes very large. Then, the Wald statistic becomes arbitrarily small, resulting in a systematic way that the null hypothesis is not rejected. In this case, we will prefer a testing methodology based on the log-likelihood criterion for a comparison before and after the introduction of each variable.

Selection of regressors: the forward stepwise procedure

The stepwise procedure of forward selection for the regressors is similar to that used in discriminating analysis: with each step *k*, the variable X_k is selected with the smallest p-value, provided that this p-value is not higher than a threshold value a priori fixed (here 1%). The principle of the test used to select an additional regressor is based on the deflation of deviance induced by the likelihood ratio (LR). This ratio is obtained by dividing the probability of the incomplete model (at the step *k-1*) by the probability of the complete model (step *k*):

$$G_k = -2\log(LR) = -2\log\left[\frac{\textit{Likelihood without } X_k}{\textit{Likelihood with } X_k}\right].$$

Under the null hypothesis *H0* of nullity of the logistic regression parameter β_k, the quantity $-2*\log(LR)$ follows a distribution of Chi-square $\chi^2(\nu)$ with ν degrees of freedom, provided that the sample size is sufficiently large. The number of degrees of freedom ν is the difference between the number of parameters of the complete model and the number of parameters of the incomplete model (here $\nu = 1$ at each step). For each included variable, the chi-squared test leads to a rejection of the null hypothesis with a p-value lower than 1 out of ten thousand for the first 5 steps and of 3 out of thousand in the last step (Table 13).

Table 13. Results of the forward stepwise selection according to the likelihood ratio criterion: Stepwise summary (a)

Step	Improvement			Model			% correctly classified	Variable
	Chi-square	dof	Signif.	Chi-square	dof	Signif.		
1	787.930	1	.000	787.930	1	.000	82.2	IN: r1
2	129.653	1	.000	917.583	2	.000	85.2	IN: r32
3	86.621	1	.000	1004.204	3	.000	88.1	IN: r14
4	29.213	1	.000	1033.417	4	.000	88.6	IN: r17
5	13.693	1	.000	1047.110	5	.000	88.7	IN: r36
6	9.083	1	.003	1056.193	6	.000	89.1	IN: r12

Note: No other variable can be suppressed or included in the current model.

Thus, we can note that the forward procedure selects successively the following ratios:

- *r1* total debt / total asset;
- *r32* (EBITDA - financial expenses) / gross product;
- *r14* short term debt / circulating asset;
- *r17* financial expenses / total debt;
- *r36* immobilized assets/ gross product;
- *r12* working capital / (real inputs - financial expenses).

Among the ratios positively correlated with the *F1* axis (*F1*>*0*: presence of financial problems), we find the general debt rate (*r1*) and the needs for liquidity (*r14*). Among the ratios negatively correlated with the *F1* axis (*F1*<*0*: absence of financial problems), we find a result ratio (*r32*) and a liquidity ratio (*r12*).

Let us note that the first four steps select the same regressors which the stepwise procedure of the discriminant analysis found by minimizing Wilks' lambda (Table 6).

Table 14. Percentage of correctly-classified holdings: **Classification Table (a).**

		Observed	Estimated Payment incident: Healthy	Estimated Payment incident: Failing	Corrected percentage
Step 6	Payment incident	Healthy	594	59	91.0
		Failing	78	529	87.1
	Global percentage				89.1

Note: The cutting point is 0.500.

Classification is performed by comparing the estimated probability of failure to a cut-off value (.50 by default in many statistical software packages). The total percentage of correctly-classified holdings obtained in the final step is *89.1*%, with *87.1*% of correctly-classified for the financial failures (Table 14). On the one hand, the false-failing rate is 9.5% (59 out of 653 "healthy" farm holdings are classified "failing"); on the other hand the false-healthy rate is 12.9% (*78* farm holdings out of 607 "failing" farm holdings are classified "healthy"). Using all the information provided by the ratio battery doesn't significantly enhance the percentage of correctly-classified: 75 out of 607 failing farm holdings are classified as "healthy", resulting in a 12.3% false-healthy rate. A Monte-Carlo cross-validation procedure adapted from Picard & Cook (1984) gives us an overall 88.8 % percentage of correctly classified with a standard error of 0.1%. Hence, it appears to be slightly but significantly better than the 88.3% rate previously obtained with the corresponding two-group linear discriminant analysis.

The predicted probabilities histogram (Figure 9) shows that the logit model produces a satisfactory estimate of the membership probability for each of the two groups because the density of each group is important in each end of the interval and the overlap remains thin among the support of the two empirical distributions. The *U*-shaped distribution indicates that the probability predictions issued from the logistic regression are well-differentiated.

Regarding the *78* failing farm holdings wrongly classified as "healthy" by the model, the histogram of the estimated probabilities indicates that about twenty are classified "healthy" with a strong probability (more than 75% for the farm holdings labeled by the *f* letter in the interval [0; 0.25]) and about sixty are classified with a poorer probability (less than 75% for the farm holdings represented by the three *f* letters in the interval [0.25; 0.50].

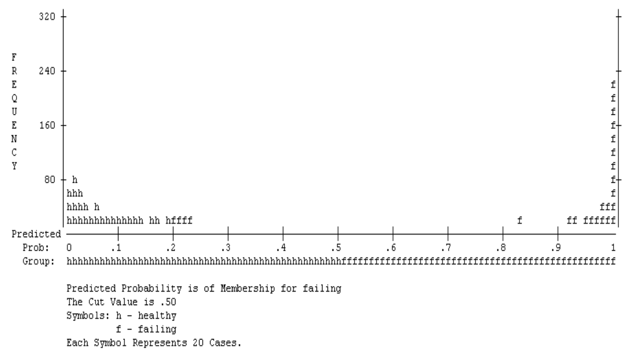

Figure 9. Histogram of the predicted probabilities.

In the context of logistic regression, as in that of linear regression, the quality assessment of the model fit involves the study of residuals, i.e. deviations between the observed probabilities and the predicted probabilities $e(i) = p(i) - \hat{p}(i)$, e.g. $e(75) = 1 - 0.084 = 0.916$, but also the study of standardized residuals $r(i) = \dfrac{e(i)}{\sqrt{\hat{p}(i)[1-\hat{p}(i)]}}$,

e.g. $r(75) = \dfrac{0.916}{\sqrt{0.084 \times 0.916}} \approx 3.304$.

Among poorly classified farm holdings (Table 15), obtained by selecting farm holdings whose standardized residual is higher than 2 in absolute value, let us consider for example i = 397, a very atypical farm holding, since *Zresid*, its standardized residual, is equal to -187.922.

From the estimates of the logistic regression coefficients provided by step 6 in Table 12 for the variables in the equation, we obtain the following model to estimate the probability of membership to the class of failing farm holdings:

$$Z(i) = -6.168 + 5.95 r_1(i) + 0.949 r_{12}(i) + 3.358 r_{14}(i) + 24.233 r_{17}(i) - 7.306 r_{32}(i) + 0.61 r_{36}(i).$$

Replacing by the actual values of the predictor ratio for the i = 397 farm holding, we obtain the value of the Z log-ratio, as follows:

$$Z(397) = -6.168 + 5.95(1.183) + 0.949(-0.891) + 3.358(2.383) + 24.233(0.057) - 7.306(0.002) + 0.61(1.762) \approx 10.472$$

For the #397 farm holding, we can derive an estimate from the 'failing' probability according to the following formula: $\hat{p}(397) = \frac{1}{1+e^{-10.472}} = 0.99997$.

Table 15. Poorly classified farm holdings with residual greater than two standard deviations: **Observation list(b).**

Observation	Selected state (a)	Observed Payment incident	Predicted	Predicted Group	Temporary Variable	
					Resid	ZResid
75	S	f**	.084	s	.916	3.304
102	S	f**	.078	s	.922	3.431
107	S	f**	.095	s	.905	3.091
185	S	f**	.067	s	.933	3.738
188	S	f**	.079	s	.921	3.414
191	S	f**	.095	s	.905	3.093
218	S	f**	.081	s	.919	3.373
301	S	f**	.040	s	.960	4.882
304	S	f**	.073	s	.927	3.573
359	S	h**	.979	d	-.979	-6.907
388	S	h**	.934	d	-.934	-3.759
393	S	h**	.867	d	-.867	-2.556
397	S	h**	1.000	d	-1.000	-187.922
405	S	f**	.099	s	.901	3.015
412	S	f**	.007	s	.993	12.199
454	S	h**	.988	d	-.988	-9.130
484	S	h**	.918	d	-.918	-3.345
493	S	h**	.923	d	-.923	-3.462
499	S	f**	.128	s	.872	2.606
507	S	f**	.014	s	.986	8.316
548	S	h**	.984	d	-.984	-7.830
580	S	h**	.904	d	-.904	-3.076
587	S	h**	.895	d	-.895	-2.918
599	S	f**	.032	s	.968	5.499
602	S	f**	.021	s	.979	6.779
917	S	f**	.061	s	.939	3.917
1028	S	f**	.055	s	.945	4.129
1051	S	f**	.094	s	.906	3.105
1069	S	h**	.866	d	-.866	-2.545
1127	S	f**	.127	s	.873	2.618
1148	S	f**	.085	s	.915	3.290

Notes: a. S = selected observations, U = non selected observations and ** = poorly classified observations.
b Observations with studentized residual greater than 2.000 are displayed.

A priori being catalogued as "healthy" (no financial failure), the residual of the #*397* farm holding estimate is equal to: $e(397) = p(397) - \hat{p}(397) = 0 - 0.99997 \approx -1.000.$ Hence, the residual divided by the estimate of its standard deviation, namely the standardized residual is equal to: $r(i) = \frac{e(i)}{\sqrt{\hat{p}(i)\times[1-\hat{p}(i)]}} \approx \frac{-0.99997}{\sqrt{0.99997\times 0.00003}} \approx -187.922$.

The standardized residual constitutes the individual contribution of each farm holding to the residual deviance of the model, used as an overall quality index for the fitting of the logistic model to the data $\Delta_R = \sum_{i\in I} \frac{e_i^2}{\hat{p}(i)\times[1-\hat{p}(i)]}$.

According to the equation of the deviance analysis, the residual deviance of the model is equal to the difference between the total deviance and the model deviance. Thus the residual deviance of the model with 6 regressors is equal to the total deviance minus the model deviance with 6 regressors:

$$\Delta_{R(6)} = \Delta_0 - \Delta_{M(6)} = 1745.051 - 1056.19 = 688.86$$

Under the H0 hypothesis that all the logistic regression coefficients are null (except for the constant), the model deviance follows a Chi-square distribution $\chi^2_{(p-1)}$ where p is equal to the number of model regressors. This property of the deviance allows a test for this model similar to the F test with the linear regression model. Under the null hypothesis H0, the deviance of this model with 6 regressors has a probability lower than 10^{-4} to be exceeded. Thus rejecting this null hypothesis as very unlikely, we conclude to the global relevance of the logistic regression model.

Logistic Regression versus Linear Discriminant Analysis: an appraisal

Although logistic regression was originally designed as a modeling tool in order to study the influence of factors in classification processes, it has been extensively used since as an individual prediction technique in decision making processes such as credit scoring. In a binary logistic regression, the probability of belonging to the "failing" group G2 of farm holdings given the x regressor values $P(G2/X = x)$ can be used as a predictive score which ranges between 0 and 1. Hence, the Fisher linear discriminant score can be compared to the estimated probability produced by a logistic regression, viewed as a score.

From the empirical point of view, the results obtained with a linear discriminant analysis and a logistic regression are very consistent (except in the extreme tails) as demonstrated in Figure 10. In this financial case study, the scores are highly correlated as the following diagram shows (Figure 10): 0.923 for the Bravais-Pearson coefficient and 0.975 for Spearman's Rho.

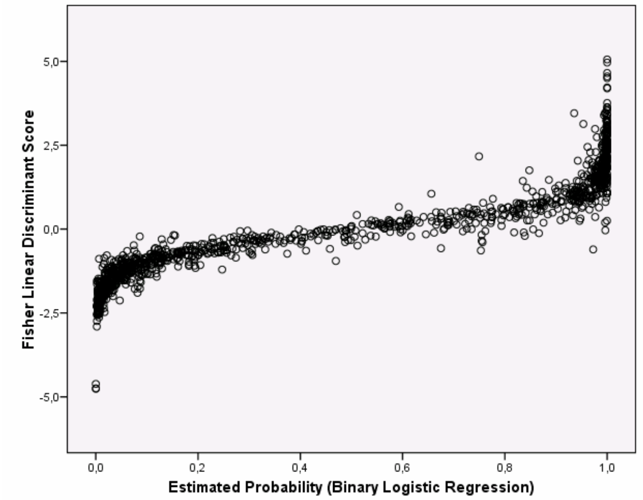

Figure 10. Linear discriminant scores against logistic regression scores.

In credit scoring among many other scoring applications, the Receiver Operating Characteristic (ROC) curve is extensively used to compare score performances because the ROC curve depends only on the ranking of the score values.

The ROC curve provides a graphical display for the trade-off between the sensitivity and specificity of a classification. Namely, the ROC curve is a display of the sensitivity of the procedure (vertical axis) against the risk of false negative ("*1-Specificity*", horizontal axis), for various choices of the cut-off probability used to classify holdings as healthy or failing. On Figure 11, each point on the curves corresponds to a particular choice of a cut-off point.

The *sensitivity* of a test procedure is defined in a medical context as the probability of having a positive test when being genuinely sick (in our case, that means being classified as a "failing" farm holding, given that the holding is indeed a "failing" holding). The *specificity* of a test procedure is the probability of having a negative test when not being sick (i.e. being classified as a "healthy" farm holding given that the holding is indeed "healthy").

The complement of the sensitivity to 1 ("1-Sensitivity") is the probability that a member of the "failing" group (G2 farm holdings) is wrongly classified as a member of the

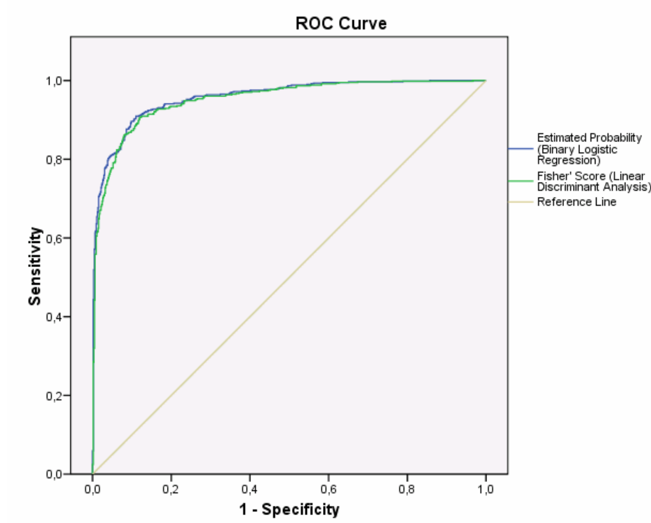

Figure 11. Comparing classification performance between logistic regression and linear discriminant analysis.

"healthy" group (G1 farm holdings). This probability is estimated by the frequency of false negatives. The complement of the specificity to 1 ("*1-Specificity*") is the probability that a member of the "healthy" group (G1 farm holdings) is wrongly classified as a member of the "failing" group (G2 farm holdings). This probability is estimated by the frequency of false positives.

Table 16. Area under the ROC curve as a measurement of classification performance.

Scores	Zone	Std Error	95% Asymptotic Confidence Interval Upper bound	Lower bound
Estimated Probability (Binary Logistic Regression)	.958	.005	.948	.968
Fisher' Score (Linear Discriminant Analysis)	.952	.006	.941	.963

The ideal classification rule would correspond to a top-left point on the ROC curve (where sensitivity and specificity would both equal 1). So the area under the ROC curve (AUC) measures the classification performance of a score (Table 16). The AUC of logistic regression is slightly greater than the linear discriminant analysis one. However, given the standard error of the estimates, this difference appears as not statistically significant.

The AUC empirical estimator typically used in AUC computations is unbiased, but the standard error estimator is biased, with a bias that depends upon the shape of distributions. In our case of non-Gaussian distributions, the standard error of estimates is too large and the asymptotic confidence intervals are conservative. This drawback can be avoided using an empirical resampling technique according to a stratified sampling scheme to derive unbiased asymptotic confidence intervals for the AUC (Saporta & Niang 2006). Moreover, when the numbers of regressors are not the same in competing models, the comparison of ROC curves by means of their AUC needs to be performed on test samples.

However, those methodological improvements are beyond the current scope of this paper.

Discussion and prospects

This case study describes how to combine exploratory methods such as principal component analysis with decision-oriented methods such as discriminant analysis and logistic regression in order to cope with the complexity of assessing the financial risk for farm holdings through a standard battery of financial ratios. A choice has been made to select standard methods in order to focus on the logic of discovery and assessment through a specific multidimensional set of real data. Thanks to multidimensional data analysis, this logic of discovery can be based on not only graphical but geometrical methods as well as statistical tests, alternatively or in a complementary (see Fénelon 1999).

Alternative methods have been used for classification purposes in the framework of scoring methods. With regards to our comparative appraisal between linear discriminant analysis and logistic regression, the results show no significant differences between the two classifying methods we favored. The use of such tools can be easily extended to other types of micro-economic studies, for the purpose of business group analysis, for example comparing farm holdings and other types of business units.

Considered as a more appropriate method in the hypothetic-deductive framework used for scientific research, the choice of logistic regression as a methodology is largely a matter of taste. It is sometimes contrasted to the choice of a linear discriminant analysis, which is much more regarded as a multivariate exploratory tool used for empirical analysis. Opposing those two methods, some considerations should be kept in mind (see Saporta 2006):

- From the theoretical point of view, both these methods are based on probabilistic models, which are however different in terms of specification. A logistic regression is based on the conditional distribution $Y / X = x$ while a discriminant analysis is based on the conditional distributions $X / Y=k$ in each group k. Moreover, the results of each model can be expressed as a “score”, which is a linear function of the x observed values. In this respect, the main difference is the numerical optimizing criterion used to get the estimates: maximum likelihood for logistic regression versus least squares for linear discriminant analysis.
- Logistic regression is theoretically appropriate even in the case of non-Gaussian distributions; the geometric approach to linear discriminant analysis can also be applied whatever the nature of the distribution might be.
- While logistic regression leads to unique estimates for coefficients, they are estimated up to a scale factor for Fisher’s linear functions. Logistic regression asymptotically provides standard errors for the estimates. This is not the case in linear discriminant analysis but bootstrap and other re-sampling methods can be used to provide empirical standard errors.

Many other methods can be used to assess the financial risk of farm holdings, including algorithmic approaches proposed under the “data mining” umbrella, which compete with distribution free and non linear statistical methods.

The output from classification and regression tree methods (CART methodology, Breiman & al, 1984) consists in a hierarchy of predictors displayed in a very convenient way (as a decision tree). These trees can accommodate all types of predictors, without any assumption on their distribution. However, the main flaw of such techniques is the fundamental instability of the results.

Neural networks, derived from the original Rosenblatt “perceptron”, are favored, in particular by non-statisticians for their flexibility despite the fact that they tend to suffer from over-fitting problems and lack of readability. However, a recent family of supervised learning methods, the support vector machines (Cortes & Vapnick, 1995) also known as maximum margin classifiers, seems to provide equivalent performances under a thoroughly statistical formulation.

The multiplicity of modeling techniques stresses the need for a robust model assessment methodology for practitioners.

Correspondence: dominique.desbois@agriculture.gouv.fr

A tribute to Jean-Pierre Fénelon: two orthogonal projection planes revealing the "Rouge des Prés" PDO morphology. *Source:* UPRA Maine-Anjou.

REFERENCES

Altman E.I. 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance* 23(4):589-609.

Breiman L., Friedman J., Olshen R.A., Stone C.J. 1984. *Classification and regression trees*. Wadsworth.

Bardos M. 2001. *Analyse discriminante : application au risque et scoring financier*. Dunod, Paris.

Bardos M. 1985. Le risque de défaillance d'entreprises. *Cahiers Economiques et Monétaires*, 19, Banque de France.

Beaver W. 1966. Financial Ratios as Predictors of Failures. *Empirical Research in Accounting, selected studies*, supplement to the *Journal of Accounting Research*, January 1967.

Blogowski A., Colson F., Léon Y. 1992. Les exploitations agricoles en difficulté financière dans la CEE. *Note INRA/Ministère de l'Agriculture et de la Forêt*, April.

CNCER 1996. *Guide comptable des exploitations agricoles*. CNERTA, Dijon.

Colson F., Blogowski A., Dechambre B., Chia E., Désarménien D., Dorin B. 1993. Prévenir les défaillances financières en agriculture. Application de la méthode des scores. *Cahiers d'économie et de sociologie rurales,* 29: 22-44.

Cortes C. and Vapnik V. 1995. Support-Vector Networks. *Machine Learning* 20: 273-297.

Dietsch M. 1989. La mesure des difficultés financières des exploitations agricoles à partir du RICA. *Rapport final - Convention SCEES*, Institut d'Etudes Politiques, Strasbourg.

Fénelon J-P. 1999. *Qu'est-ce que l'Analyse des Données*. LEFONEN, 2nd Edition.

Flavigny P-O. and Lebeaux M-O. 1998. *Manuel d'utilisation des procédures ADDADSAS, version SAS/Toolkit 94-98*.

Isnard M. and Sautory O. 1994. *Les macros SAS d'analyse des données*, Working paper n° *F 9405,* Direction des Statistiques Démographiques et Sociales, INSEE.

McFadden D.L. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. *Frontiers in Econometrics*:105-142, *Academic Press*, New York.

Picard R.R. and Cook R.D. 1984. Cross-validation of regression models. *Journal of the American Statistical Association* 79:575-583.

Saporta G. 2006. *Probabilités, analyse des données et statistique,* Editions Technip, 2nd Edition.

Saporta G. and Niang N. 2006. « Model Assessment» *Knowledge Extraction by Modeling (provisional papers),* Compstat Satellite IASC Meetings, Capri, Italy, September 4-6.

Addendum. SPSS code to perform the main statistical analyses.

```
* Principal Component Analysis (PCA) of  predictor battery.
FACTOR
 /VARIABLES r1 r2 r3 r4 r5 r6 r7 r8 r11 r12 r14 r17 r18 r19 r21 r22 r24 r28
 r30 r32 r36 r37  /MISSING LISTWISE /ANALYSIS r1 r2 r3 r4 r5 r6 r7 r8 r11 r12
 r14 r17 r18 r19 r21 r22 r24 r28 r30 r32 r36 r37
 /PRINT INITIAL EXTRACTION FSCORE
 /FORMAT SORT
 /PLOT EIGEN ROTATION
 /CRITERIA MINEIGEN(1) ITERATE(25)
 /EXTRACTION PC
 /ROTATION NOROTATE
 /SAVE REG(ALL)
 /METHOD=CORRELATION .

* Classifying with PCA Factor F#1 according to the geometrical rule.
RECODE
 FAC1_1
 (Lowest thru 0.0256=0)  (0.0256 thru Highest=1)  INTO  RISK
.
VARIABLE LABELS RISK 'geometrical rule of classification'.
EXECUTE .

* Displaying the geometrical rule classification results.
IGRAPH /VIEWNAME='Diagramme en bâtons' /X1 = VAR(DIFF) TYPE = CATEGORICAL /Y
 = $count /STYLE = VAR(RISK) STACK /COORDINATE = VERTICAL /NORMALIZE
 /X1LENGTH=3.0 /YLENGTH=3.0 /X2LENGTH=3.0 /CHARTLOOK='NONE' /CATORDER
 VAR(DIFF) (ASCENDING VALUES OMITEMPTY) /CATORDER VAR(RISK) (ASCENDING VALUES
 OMITEMPTY) /BAR KEY=ON LABEL INSIDE N  SHAPE = RECTANGLE BASELINE = AUTO.

* Classifying with the total debt rate according to the third quartile rule.
RECODE
 r1
 (Lowest thru 0.48=0)  (0.48 thru Highest=1)  INTO INDEBTED .
VARIABLE LABELS INDEBTED 'Total Debt Rate third quartile'.
EXECUTE .

* Displaying the total debt rate classification results.
IGRAPH /VIEWNAME='Diagramme en bâtons' /X1 = VAR(DIFF) TYPE = CATEGORICAL /Y
 = $count /STYLE = VAR(INDEBTED) STACK /COORDINATE = VERTICAL /NORMALIZE
 /X1LENGTH=3.0 /YLENGTH=3.0 /X2LENGTH=3.0 /CHARTLOOK='NONE' /CATORDER
 VAR(DIFF) (ASCENDING VALUES OMITEMPTY) /CATORDER VAR(INDEBTED) (ASCENDING
 VALUES OMITEMPTY) /BAR KEY=ON LABEL INSIDE N SHAPE = RECTANGLE BASELINE =
 AUTO.

* Linear discriminant analysis  (LDA) with stepwise selection of predictors.
DISCRIMINANT
 /GROUPS=DIFF(0 1)
 /VARIABLES=r1 r2 r3 r4 r5 r6 r7 r8 r11 r12 r14 r17 r18 r19 r21 r22 r24 r28
 r30 r32 r36 r37
 /ANALYSIS ALL
 /SAVE=CLASS SCORES PROBS
  /METHOD=WILKS
  /FIN= 3.84
  /FOUT= 2.71
 /PRIORS  EQUAL
  /HISTORY=NONE
 /STATISTICS=MEAN STDDEV UNIVF BOXM COEFF RAW FPAIR TABLE CROSSVALID
 /PLOT=COMBINED SEPARATE MAP
 /CLASSIFY=NONMISSING POOLED .

* Quadratic discriminant analysis (QDA)  with stepwise selection of predictors.
DISCRIMINANT
 /GROUPS=DIFF(0 1)
 /VARIABLES=r1 r2 r3 r4 r5 r6 r7 r8 r11 r12 r14 r17 r18 r19 r21 r22 r24 r28
 r30 r32 r36 r37
 /ANALYSIS ALL
 /SAVE=CLASS SCORES PROBS
  /METHOD=WILKS
  /FIN= 3.84
  /FOUT= 2.71
 /PRIORS  EQUAL
  /HISTORY=NONE
 /STATISTICS=MEAN STDDEV UNIVF BOXM COEFF RAW FPAIR TABLE
 /PLOT=COMBINED SEPARATE MAP
 /CLASSIFY=NONMISSING SEPARATE .

* Forward selection procedure to compute Binary Logistic Regression .
LOGISTIC REGRESSION VARIABLES  diff
 /METHOD = FSTEP(LR) r1 r2 r3 r4 r5 r6 r7 r8 r11 r12 r14 r17 r18 r19 r21 r22
 r24 r28 r30 r32 r36 r37
 /SAVE = PRED PGROUP COOK LEVER DFBETA RESID LRESID SRESID ZRESID DEV
 /CLASSPLOT /CASEWISE OUTLIER(2)
 /PRINT = GOODFIT CORR ITER(1) SUMMARY CI(95)
 /CRITERIA = PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .

* 2D scattergram displaying LDA  and Logistic Regression scores .
GRAPH
 /SCATTERPLOT(BIVAR)=Pre_1 WITH Dis1_1
 /MISSING=LISTWISE .

* ROC Curves procedure  in order to compare LDA  and Logistic Regression .
ROC
 Pre_1 Dis1_1  BY DIFF (1)
 /PLOT = CURVE(REFERENCE)
 /PRINT = SE
 /CRITERIA = CUTOFF(INCLUDE) TESTPOS(LARGE) DISTRIBUTION(FREE) CI(95)
 /MISSING = EXCLUDE .
```