“A Review on Statistical and Probabilistic Models for the Control of Insurance Companies”

AUTHORS	Fabio Baione Paolo De Angelis

ARTICLE INFO	Fabio Baione and Paolo De Angelis (2006). A Review on Statistical and Probabilistic Models for the Control of Insurance Companies. *Investment Management and Financial Innovations*, *3*(4)
RELEASED ON	Friday, 15 December 2006
JOURNAL	"Investment Management and Financial Innovations"
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”

NUMBER OF REFERENCES **0**

NUMBER OF FIGURES **0**

NUMBER OF TABLES **0**

© The author(s) 2026. This publication is an open access article.

A REVIEW ON STATISTICAL AND PROBABILISTIC MODELS FOR THE CONTROL OF INSURANCE COMPANIES

Fabio Baione*, Paolo De Angelis**

Abstract

The problem of evaluating the solvency of insurance companies is tackled by means of a non-parametric statistical model, constructed using decision-tree techniques. The model is tested on a sample of Italian non-life insurance companies and its performance over the test period compared with those of linear and quadratic parametric models. In the last part a probabilistic model is proposed focused on classical Risk-Theory and implemented using simulation techniques.

Key words: Decision Tree; Discriminant Analysis; Bayesian Approach; Ruin Probability.
JEL classification: C11, C14, C15, G22.

I. Introduction

The problem of assessing the solvency of firms was tackled first in the field of corporate economics and has recently acquired a significant status in the theory of decisions in conditions of uncertainty.

The literature offers several alternative methods for the construction of corporate risk evaluation models with important practical application at the company level.

A great deal of work has been done in the industrial field using multivariate analysis methods (Altman, 1968 and 1977; Deakin, 1972 and 1977; Edmister, 1972; Blum, 1974; Eisenbeis, 1977; Forestieri et al., 1986; Barontini, 1992); less rich in the credit and insurance sectors (Pinches and Trieschmann, 1974 and 1977; Meador and Thornton, 1978; Sinkey, 1979; Buoro, 1980; Hershbarger and Miller, 1986; Ambrose and Seward, 1988; De Angelis, De Felice, Ottaviani, 1988; Barniv and Hershbarger, 1990; De Angelis, Gismondi, Ottaviani, 1992).

As regards the approach based on gambler's ruin models, a particular application of risk theory to corporate evaluation, the literature is sparser, but the following works deserve mention: Wilcox (1971, 1973 and 1979), Vinso (1979), Scott (1981), Beard, Pentikâinen, Pesonen (1984), Sandberg, Lewellen, Stanley (1987), Pentikâinen, Bonsdorff, Pesonen, Rantala, Ruohonen (1989); Savelli (2002).

In the first part of the paper a comparison of results of linear and quadratic parametric models and non-parametric models to the Italian insurance industry[1] is reported and in the last part a probabilistic model is shown, carried out by means of stochastic simulation procedures.

II. Description of the non-parametric statistical model

Non-parametric models are based on the sequential decision-making procedures typical of the techniques for the construction of decision trees; therefore it is easy to formalize a non-parametric model for the evaluation of insurance companies.

Let:

- $\mathbf{L} = \{l_1, l_2, \ldots, l_n\}$ the set of n insurance companies;

* University of Rome, "La Sapienza", Italy.
** University of Rome, "La Sapienza", Italy.

[1] See De Angelis (1988), De Angelis et al. (1988 and 1992) and Gismondi (1990 and 1992).

© Fabio Baione, Paolo De Angelis, 2006.

- $\mathbf{x}^j = \{x_1^j, x_2^i, \ldots, x_k^i\}$ the vector of numerical determinations of the multiple variable **X** with independent components (balance sheet indicators), associated with each company l_j [1].

The decision-making problem that arises consists in attributing a generic company l_j to one of the subset $\mathbf{L}_\mathbf{i}^{'}$ of **L** ($\mathbf{L}_1^{'}$ = safe companies; $\mathbf{L}_\mathbf{2}^{'}$ = unsafe companies) on the basis of the information contained in the vector $\mathbf{x}^j$.

The non-parametric model proposed[2] uses a recursive partitioning algorithm to divide the original sample (initial node) into subsamples (terminal nodes). At each pass the decomposition of a node into two subnodes is made on the basis of the comparison between each observed value of the balance sheet indicator X_i and a threshold value; iteration of the procedure for all the possible threshold values that can be defined with reference to the multiple variable **X** $= \{X_1, X_2, \ldots, X_k\}$ generates a finite set of admissible trees, among which the one corresponding to min I(a) is chosen, with:

$$I(a) = \sum_{t \in \tilde{\mathbf{T}}_\mathbf{a}} \sum_{\mathbf{L}_\mathbf{i}^{'} \subset \mathbf{L}^{'}} R_{\mathbf{L}_\mathbf{i}^{'}}(t) P(\mathbf{L}_\mathbf{i}^{'} / t), \quad a \in \tilde{\mathbf{A}} \tag{1}$$

$\tilde{\mathbf{A}}$ – set of admissible trees,

$\tilde{\mathbf{T}}_\mathbf{a}$ – set of terminal nodes of tree *a*,

$I(a)$ – impurity associated with tree *a*,

$R_{\mathbf{L}_\mathbf{i}^{'}}(t)$ – risk (average loss) of node *t* assigned to the subset $\mathbf{L}_\mathbf{i}^{'}$,

$P(\mathbf{L}_\mathbf{i}^{'} / t)$ – conditional probability of the node *t* given that the company belongs to the class $\mathbf{L}_\mathbf{i}^{'}$.

The assignment of a generic company l_j to $\mathbf{L}_\mathbf{1}^{'}$ or $\mathbf{L}_\mathbf{2}^{'}$ corresponds to the classification of "safe" (S) or "unsafe" $(\bar{S})$ attributed to the terminal node to which it belongs; the terminal node *t* is classified as "unsafe" if:

$$R_{\mathbf{L}_\mathbf{1}^{'}}(t) > R_{\mathbf{L}_\mathbf{2}^{'}}(t), \tag{2}$$

where:

$$R_{\mathbf{L}_\mathbf{1}^{'}}(t) = c(\mathbf{L}_\mathbf{1}^{'} / \mathbf{L}_\mathbf{2}^{'}) P(\mathbf{L}_\mathbf{2}^{'}, t) + c(\mathbf{L}_\mathbf{1}^{'} / \mathbf{L}_\mathbf{1}^{'}) P(\mathbf{L}_\mathbf{1}^{'}, t), \tag{3}$$

$$R_{\mathbf{L}_\mathbf{2}^{'}}(t) = c(\mathbf{L}_\mathbf{2}^{'} / \mathbf{L}_\mathbf{1}^{'}) P(\mathbf{L}_\mathbf{1}^{'}, t) + c(\mathbf{L}_\mathbf{2}^{'} / \mathbf{L}_\mathbf{2}^{'}) P(\mathbf{L}_\mathbf{2}^{'}, t) \tag{4}$$

and

$c(\mathbf{L}_\mathbf{p}^{'} / \mathbf{L}_\mathbf{i}^{\mathbf{i}})$, $p = i =$ 1,2, the costs of mislassification[3],

[1] The condition of independence is a "strict" hypothesis and a "regular" approximation is possible in applications by using techniques of multivariate analysis; tolerance limit can be fixed at the coefficient of correlation among the components of the multiple variable *X* (see De Angelis et al. (1988, pp. 22-24)).

[2] The model is that proposed by Frydman et al. (1985), and formalized for insurance companies by Gismondi (1990). Details on the impurity functions are to be found in Marais et al. (1984).

[3] In our application it is assumed that $c(\mathbf{L}_\mathbf{1}^{'} / \mathbf{L}_\mathbf{1}^{'}) = c(\mathbf{L}_\mathbf{2}^{'} / \mathbf{L}_\mathbf{2}^{'}) = 0$.

$P(\mathbf{L}_i', t) = P(\mathbf{L}_i')P(t/\mathbf{L}_i'), i =$ 1,2, the probability that a company belonging to $\mathbf{L}_i'$ will be in node *t,*

$P(\mathbf{L}_i')$, $i = 1,2$, prior probability that a company belongs to $\mathbf{L}_i'$,

$P(t/\mathbf{L}_i')$, $i =$ 1,2, conditional probability that a company belonging to $\mathbf{L}_i'$ will be in node *t*.

III. Application of the non-parametric (NPA) model to the Italian insurance market

The data used for the analysis consist of the annual accounts of 48 non-life insurance companies operating in Italy and having a portfolio in 1981 not more than 50 billion lire. The observations cover the period of 1981-1989.

The sample, which represents 14% of the Italian market, is a significant subset of the original sample studied in De Angelis et al. (1988); the 48 companies have been divided, using cluster analysis, into two homogeneous groups consisting of 25 safe companies and 23 unsafe ones.

The set of 8 balance sheet ratios, selected by adopting a factor analysis1, has been computed for the sample, specifically:

LQ02= (securities + cash + deposits-short-term debt) / technical reserves

RD02= net financial income / average of technical reserves at the beginning and at the end of the accounting period

CR09= direct insurance: loss reserve of the year / losses incurred and paid in the year

IE01= profit or loss for the year / shareholders' equity

RI13= direct insurance: losses borne by reinsures / losses paid

TA02= motor vehicle insurance: premium reserve / premiums

RRA1= motor vehicle insurance: net financial income / average stock of technical reserves

VR01= direct operations: premium reserve / premiums (percentage change over time)

The NPA model has been estimated on the 1981 accounts of the sample companies with reference to two hypotheses:

Hypothesis A:

$$P(\mathbf{L}_1') = P(\mathbf{L}_2') \text{ and } c(\mathbf{L}_1'/\mathbf{L}_2') = c(\mathbf{L}_2'/\mathbf{L}_1')$$

Hypothesis B:

$$P(\mathbf{L}_1') = 0.9, \quad P(\mathbf{L}_2') = 0.1 \text{ and } c(\mathbf{L}_1'/\mathbf{L}_2') = 5c(\mathbf{L}_2'/\mathbf{L}_1')$$

The decision trees shown in Figures 1a and 1b are those with the least impurity among the 768 admissible trees obtained using the selection procedure. The two indicators considered to be particularly significant for distinguishing between safe (S) and unsafe $(\bar{S})$ companies, RD02 and LQ02, are present in both hypotheses.

[1] See De Angelis et al. (1988, pp. 22-24).

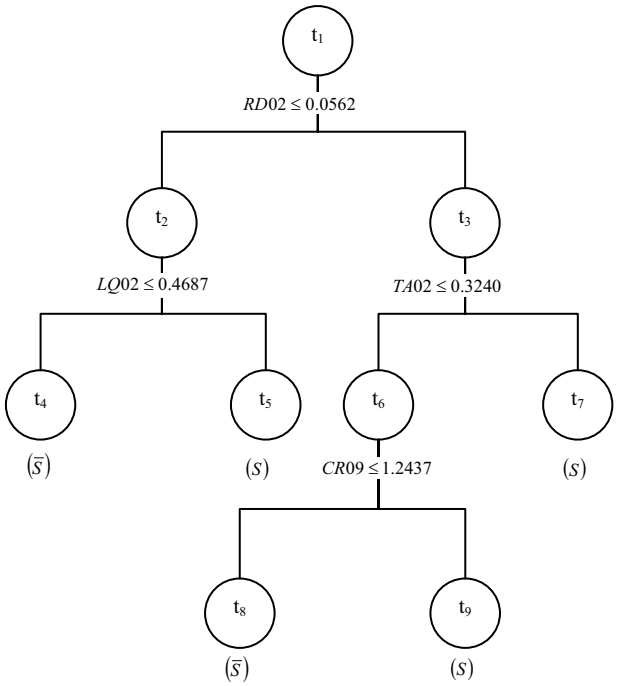

Fig. 1a. Decision trees. HYPOTHESIS A

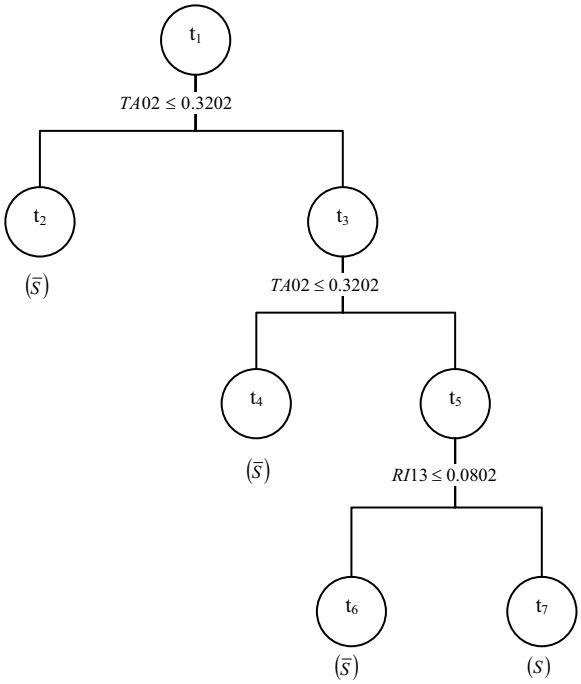

Fig. 1b. Decision trees. HYPOTHESIS B

Table 1 shows the values of $P\left(\mathrm{L}_1^{'},t\right), P\left(\mathrm{L}_2^{'},t\right)$ and $R_{\mathrm{L}_1^{'}}(t), R_{\mathrm{L}_2^{'}}(t)$

Table 1

Levels of $R_{\mathbf{L}_1'}(t)$ and $R_{\mathbf{L}_2'}(t)$

NODE	$P(\mathbf{L}_1', t)$	$P(\mathbf{L}_2', t)$	$R_{\mathbf{L}_1'}(t)$	$R_{\mathbf{L}_2'}(t)$
HYPOTHESIS A				
t1	0.5000	0.5000	0.5000	0.5000
t2	0.0218	0.4400	0.4400	0.0218
t3	0.4783	0.0600	0.0600	0.4783
t4	0.0000	0.4400	0.4400	0.0000
t5	0.0218	0.0000	0.0000	0.0218
t6	0.0218	0.0600	0.0600	0.0218
t7	0.4566	0.0000	0.0000	0.4566
t8	0.0000	0.0600	0.0600	0.0000
t9	0.0218	0.0000	0.0000	0.0218
HYPOTHESIS B				
t1	0.9000	0.1000	0.5000	0.9000
t2	0.0000	0.0800	0.4000	0.0000
t3	0.9000	0.0200	0.1000	0.9000
t4	0.0000	0.0160	0.0800	0.9000
t5	0.9000	0.0040	0.0200	0.9000
t6	0.0000	0.0040	0.0200	0.0000
t7	0.9000	0.0000	0.0000	0.0000

The discriminatory power of the model has been measured by reclassifying the sample companies on the basis of their accounts for the years from 1981 to 1989.

In accordance with the Bayesian approach adopted, the performance of the model has been measured not with reference to the usual indicators of total and conditional efficiency[1], but by introducing the indicator[2]:

$$E[C(T)] = P(\mathbf{L}_1')\frac{n_{2/1}(T)}{n_1(T)}c(\mathbf{L}_2'/\mathbf{L}_1') + P(\mathbf{L}_2')\frac{n_{1/2}(T)}{n_2(T)}c(\mathbf{L}_1'/\mathbf{L}_2'), \quad (5)$$

where: $n_i(T)$ – companies of $\mathbf{L}_i'(i=1,2)$ observed in year **T**,

$n_{1/2}(T)$, $n_{2/1}(T)$ – companies misclassified in $\mathbf{L}_1'$ and $\mathbf{L}_2'$.

Table 2 shows the values of $E[C(T)]$ observed in the "test" period of 1981-1989 for the two hypotheses. For the test period as a whole the average cost of misclassification was respectively 0.0756 and 0.1521 cost units. There was a marked deterioration in the reliability of the results in the last three years of the period, six years out from the estimation year. It needs to be stressed that the levels of reliability observed depend on the assumption that the division between safe and unsafe companies based on the accounts for 1981 is significant for the whole period.

[1] For further details, see Joy and Tollefson (1985, pp. 728-729). An application of the model to the Italian insurance industry can be found in De Angelis et al. (1988, pp. 17-30).

[2] The indicator is an estimate of the expected cost of a misclassification by the model (see Frydman et al. (1985, pp. 280)).

Table 2

Values of $E[C(T)]$ observed for the NPA model

YEAR	HYPOTHESIS A	HYPOTHESIS B
1981	0.0217	0.0200
1982	0.0435	0.2000
1983	0.0705	0.1818
1984	0.0705	0.1068
1985	0.0455	0.1443
1986	0.0788	0.1561
1987	0.1121	0.0818
1988	0.1242	0.2379
1989	0.1136	0.2403

IV. The non-parametric model compared with linear and quadratic parametric models

4.1. The discriminatory power of the NPA model is confirmed by comparison with efficiency of linear (LDA) and quadratic (QDA) models estimated on the 1981 accounts of the sample companies.

Specifically, the LDA model considered is given by[1].

$$Z = b_1 LQ02 + b_2 RD02 + b_3 CR09 + b_4 IE01 + b_5 RI13 + b_6 TA02 + b_7 RRA1 + b_8 VR01, \quad (7)$$

where: $\boldsymbol{b}$ = (0.21, 0.230, -0.001, -0.004, 0.006, 0.065, 0.019, -0.111)
and the QDA model by

$$Q = -\left(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{L}_1'}\right)' \sum_{\mathbf{L}_1'}^{-1} \left(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{L}_1'}\right) + \left(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{L}_2'}\right)' \sum_{\mathbf{L}_2'}^{-1} \left(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{L}_2'}\right), \quad (8)$$

where[2]: $\boldsymbol{\mu}_{\mathbf{L}_1'}$ = (0.574, 0.105, 1.888, 0.094, 0.430, 0.395, 0.086, 0.020)

$\boldsymbol{\mu}_{\mathbf{L}_2'}$ =(0.022, 0.030, 1.399, -0.061, 0.375, 0.281, 0.032, -0.002)

and

[1] This is the linear model used in De Angelis et al. (1988) and constructed following the approach to discriminatory analysis due to Fisher and Mahalanobis (1936).

[2] The quadratic model is derived from linear model when it is not possible to sustain the hypothesis of equal covariance in the two groups of companies on which the estimation is based. Specifically, $\boldsymbol{\mu}_{\mathbf{L}_i'}$ and $\Sigma_{\mathbf{L}_i'}(i = 1,2)$ are the vector of the means and the variance-covariance matrix in $\mathbf{L}_1'$ and $\mathbf{L}_2'$. Interesting comments can be found in Eisenbeis (1977, pp. 876-881) and some interesting comments and applications of the model to the Italian insurance industry are offered in De Angelis et al. (1988).

$$\sum\nolimits_{\mathbf{L}_1^{'}}^{-1} = \begin{bmatrix} -48.5 & 869.6 & -15.8 & -45.9 & 31.7 & -414.6 & -527.7 & 15.8 \\ 869.9 & -7213 & 35.5 & 358.7 & -331.2 & 8185 & 3831 & -465.9 \\ -15.8 & 35.5 & 2.6 & -0.3 & 1.5 & -64.7 & -3.3 & 11.2 \\ -45.9 & 358.7 & -0.3 & 99.5 & 0.8 & -373 & -123.7 & 23.0 \\ 31.7 & -331.2 & 1.5 & 0.8 & 22.3 & 74.4 & 161.9 & 11.0 \\ -414.6 & 5185 & -64.7 & -373 & 74.4 & -368.7 & -2675 & 189.7 \\ -526.7 & 3831 & -3.3 & 123.7 & 161.9 & -2675 & -810.8 & 154.7 \\ 15.8 & -465.9 & 11.2 & 23.0 & 11.0 & 189.7 & 154.7 & 564.2 \end{bmatrix}$$

$$\sum\nolimits_{\mathbf{L}_2^{'}}^{-1} = \begin{bmatrix} 69.8 & -173.4 & 21.3 & -77.2 & 14.3 & -181.2 & 64.3 & -18.5 \\ -173.4 & 3245 & -61.9 & 1.1 & -77.4 & 993.6 & -2744 & -49.9 \\ 21.3 & -61.9 & 16.8 & -14.8 & -2.6 & -93.9 & 83 & -2.1 \\ -77.2 & 1.1 & -14.8 & 121 & -27.4 & 115.9 & 66 & 28.4 \\ 14.3 & -77.4 & -2.6 & -27.4 & 35.7 & -35.1 & 229.8 & -8.8 \\ -181.2 & 993.6 & -93.9 & 115.9 & -35.1 & 992.1 & -1454 & -60.4 \\ 64.3 & -2744 & 83 & 66 & 229.8 & -1454 & 7994 & 209.6 \\ -18.5 & -49.9 & -2.1 & 28.4 & -8.8 & -60.4 & 209.6 & 88.8 \end{bmatrix}$$

According to the Bayesian approach adopted, the automatic classification procedure for the LDA and QDA models has been strengthened by introducing a correction factor in the "cut-off value" of the intervals of safe and unsafe companies.

For the LDA model the cut-off value is defined by[1]:

$$h^{LDA}\left(\bar{z}_{1,t}\bar{z}_{2,t}, s_t^2, t\right) = \frac{\left(\bar{z}_{1,t} + \bar{z}_{2,t}\right)}{2} + \frac{s_t^2}{\left(\bar{z}_{1,t} - \bar{z}_{2,t}\right)}k, \tag{9}$$

where: $\bar{z}_{1,t}\bar{z}_{2,t}$ are the z-scores mean values of the companies in $\mathbf{L}_1^{'}$ e $\mathbf{L}_2^{'}$ valued in year t,

s_t^2 is the z-scores variance, valued in year t,

$k = \log \frac{c\left(\mathrm{L}_1^{'}/\mathrm{L}_2^{'}\right)P\left(\mathrm{L}_2^{'}\right)}{c\left(\mathrm{L}_2^{'}/\mathrm{L}_1^{'}\right)P\left(\mathrm{L}_1^{'}\right)}$ is the correction factor;

and for the QDA model by[2]

[1] The Bayesian formulation of the cut-off of the linear discriminatory model is due to Graham and Johnson (1977, pp. 317-319).

[2] The cut-off of the quadratic model is obtained from the maximum likelihood ratio under the hypothesis: $\frac{c\left(\mathbf{L}_1^{'}/\mathbf{L}_2^{'}\right)P\left(\mathbf{L}_2^{'}\right)}{c\left(\mathbf{L}_2^{'}/\mathbf{L}_1^{'}\right)P\left(\mathbf{L}_1^{'}\right)} = 1.$

$$h^{QDA}\left(\Sigma_{\mathbf{L}'_{1,t}},\Sigma_{\mathbf{L}'_{2,t}}k\right)=\log\frac{\left|\Sigma_{\mathbf{L}'_{1,t}}\right|}{\left|\Sigma_{\mathbf{L}'_{2,t}}\right|}+2k, \tag{10}$$

where: $\Sigma_{\mathbf{L}'_{1,t}},\Sigma_{\mathbf{L}'_{2,t}}$ are the variance-covariance matrix in $\mathbf{L}'_1$ and $\mathbf{L}'_2$, estimated in year *t*.

4.2. The LDA, QDA and NPA models have been compared with reference to the values of the indicator $E[C(T)]$ for each test year and for both hypotheses, A and B. Table 3 shows the values of $E[C(T)]$ for selected years.

Table 3

The values of $E[C(T)]$ in selected years for the LDA, QDA and NPA models

YEAR	1981	1984	1987	1989
HYPOTHESIS A				
LDA model	0.0000	0.0727	0.1348	0.2980
QDA model	0.2217	0.4250	0.4227	0.5000
NPA model	0.0217	0.0705	0.1121	0.1136
HYPOTHESIS B				
LDA model	0.2200	0.1409	0.1667	0.3571
QDA model	0.2591	0.4250	0.4409	0.5000
NPA model	0.0200	0.1068	0.0818	0.2403

a. With reference to the hypothesis:

$$P\left(\mathrm{L}'_1\right)=P\left(\mathrm{L}'_2\right)=0 \text{ and } c\left(\mathrm{L}'_1/\mathrm{L}'_2\right)=c\left(\mathrm{L}'_2/\mathrm{L}'_1\right)$$

- the NPA model performs better than the LDA and QDA models. The average expected cost of misclassification over the whole test period was found to be 0.0756 cost units, compared with 0.106 units for the LDA model and 0.383 units for the QDA model. On average the NPA model correctly classified 92% of the companies, as against 89% for the LDA model and 64% for the QDA model;
- the NPA and LDA models had the same discriminatory power in the first four years after the estimation year, with $E[C(T)]$ averaging 0.050 cost units.

After the sixth year, the expected cost of misclassification for the NPA model was 40% less than that for the LDA model.

b. With reference to the hypothesis:

$$P\left(\mathrm{L}'_1\right)=0.9,\quad P\left(\mathrm{L}'_2\right)=0.1 \text{ and } c\left(\mathrm{L}'_1/\mathrm{L}'_2\right)=5c\left(\mathrm{L}'_2/\mathrm{L}'_1\right)$$

- the NPA model had an average overall efficiency of 89.3%, as against 80% for the LDA model and 65% for the QDA model. The conditional efficiency for the group of unsafe companies was particularly high: the NPA model correctly classified 92.2% of the unsafe companies, as against 57% for the LDA model and 24% for the QDA model;
- the average expected cost of misclassification for the NPA model was 35% less than for the LDA model and 62% less than that for the QDA model.

The use of additional information with respect to the classification rule of each model confirms the Bayesian value of the NPA model.

The better performance of the NPA model can be explained by the fact that prior probabilities and misclassification costs are incorporated directly in its construction, both when selecting the indicators and when fixing the rule for classifying terminal nodes as safe or unsafe.

Nonetheless, in view of its operational simplicity, the LDA model's potential contribution to decision making should not be undervalued.

For large samples, and in the absence of information on $P\left(\mathbf{L}_{\mathbf{j}}^{'}\right)$ and $c\left(\mathrm{L}_{\mathrm{i}}^{'}/\mathrm{L}_{\mathrm{j}}^{'}\right)$, it performs well in the first three or four years after the estimation year.

V. A stochastic model for evaluating ruin probabilities

In reference to the traditional literature on classical Risk Theory, it is possible to introduce on a set of insurance companies a preference order on the basis of their solvency quality, adopting a ruin probability measure defined by $\mathrm{Prob}\left(\widetilde{U}_t \le 0\right)$, where $\widetilde{U}_t$ is the stochastic Risk Reserve at the end of year *t*; in particular $\left\{\widetilde{\mathbf{U}}_t; t = 1,2,...\right\}$ is the stochastic process describing the Risk Reserve's dynamic over the time, whose *t-th* component is[1]:

$$\widetilde{U}_t = \left(1 + i\left(t-1,t\right)\right)\widetilde{U}_{t-1} + \left(\Pi_t - \widetilde{S}_{\mathrm{t}} - \mathrm{E}_{\mathrm{t}}\right)\left(1 + i\left(t-1,t\right)\right)^{0,5}, \tag{11}$$

where: $i\left(t-1,t\right)$: risk free rate over the period [*t-1, t*],

Π_t : gross premium volume, including safety and expense loadings,

$\widetilde{S}_t$: stochastic aggregate claim amount,

E_t : general and acquisition expenses.

Assuming that expense loading amounts are equal to actual expenses and not considering the capitalization factor $\left(1 + i\left(t-1,t\right)\right)$, the (11) can be simplified as $\widetilde{U}_t = \widetilde{U}_{t-1} + \left(\mathrm{P}_t - \widetilde{S}_{\mathrm{t}}\right)$, where $\mathrm{P}_t = \mathrm{E}\left(\widetilde{S}_t\right)$ is the risk premium, including safety loadings. In accordance with the actuarial perspective it is usual to assume $\widetilde{S}_t = \sum_{i=1}^{\widetilde{N}_t} \widetilde{Y}_{i,t}$ to be a compound mixed Poisson process, where $\widetilde{N}_t$ (number of claims occurred in year *t*) is a mixed Poisson random variable, with parameter $n_t = n_0\left(1+g\right)^t$, increasing by the real growth rate *g* and $\widetilde{Y}_{i,t}$ (i.i.d. random variables) the random variable representing the *i-th* claim of year *t*, $\widetilde{N}_t$ and $\widetilde{Y}_{i,t}$ being reciprocally independent for each year *t*.

Under the above mentioned assumptions, if no autocorrelation is in force for all the components of the aggregate claims amount, the first three moments of the $\widetilde{S}_t$ distribution are:

$$\begin{aligned} E\left(\widetilde{S}_t\right) &= n_t\, a_{1Y,t} = \left(1+g\right)^t \left(1+k\right)^t E\left(\widetilde{S}_0\right), \\ \sigma^2\left(\widetilde{S}_t\right) &= n_t\, a_{2Y,t} = \left(1+g\right)^t \left(1+k\right)^{2t} \sigma^2\left(\widetilde{S}_0\right), \end{aligned} \tag{12}$$

[1] For further details see Pentikainen and Rantala (1995, pp. 116-122) and Savelli (2002, pp. 4-6).

$$\gamma\left(\widetilde{S}_t\right)=\frac{1}{\sqrt{n_t}}\frac{a_{3Y,t}}{\left(a_{2Y,t}\right)^{1,5}}=\frac{1}{\sqrt{(1+g)^t}}\gamma\left(\widetilde{S}_0\right),$$

where $a_{jY,t}=E\left(\widetilde{Y}_{I,t}^{j}\right)=(1+k)^{j\,t}\;E\left(\widetilde{Y}_{i,0}^{j}\right)=(1+k)^{j\,t}\;a_{jY,0}$ is the *j-th* absolute moment of the random amount of the *i-th* claim of year *t*, depending on the inflation rate *k*.

It is possible to reach the aim of building a preference order on a set of insurance companies with reference to a ruin probability measure using the Pentikainen-Rantala simulation method; this method is based on randomizing the aggregate losses of all claims for an accident year and it requires as input parameters the mean, the standard deviation and the skewness of $\widetilde{S}_t$, computed when the mean claim number and the lowest moments of the individual claims are available as estimates from company's observed data and assuming *g* is deterministic and *k* moves over the time as Wilkie's autoregressive model. Steps of a simulation procedure referred to each realisation on *t-th* year can be represented as follows:

1st step: generate number $n_t=n_0\;q_t\;(1+g)^t$, where q_t is a realisation of a structure variable $\widetilde{Q}_t$ introducing into the model the stochastic fluctuation; in particular $\widetilde{Q}_t - 1=\alpha\left(\widetilde{Q}_{t-1}-1\right)+\varepsilon_t$ is a first order autoregressive process with $\varepsilon_t \approx N(0,1)$;

2nd step: compute k_t as a realisation of the stochastic inflation rate $\widetilde{K}_t - \bar{k}=\beta\left(\widetilde{K}_{t-1}-\bar{k}\right)+\xi_t$ with $\xi_t \approx \mathrm{N}(0,1)$;

3rd step: compute the first three moments of $\widetilde{S}_t$, conditioned to generated numbers n_t and k_t;

4th step: compute the Wilson-Hilferty's [1] formula using the first three moments of $\widetilde{S}_t$ to derive s_t as a realisation of $\widetilde{S}_t$;

5th step: compute $\mathrm{P}_t=\mathrm{E}\left(\widetilde{S}_\mathrm{t}\right)$;

6th step: compute $u_t=u_{t-1}+\left(\mathrm{P}_t - s_\mathrm{t}\right)$ as a realisation of the Stochastic Reserve;

7th step: go back to the 1st step to run a new realisation.

In lack of specific data to create a preference order on the insurance companies sample described in par. 3.1., the simulation procedure has been built to analyse its properties and to give some sensitivity tests. It has been used:

- PC tools like Visual Basic for Application, Matlab and Excel,
- ISVAP's statistics on car accidents, referred to Italian insurance market and claims statistical data of an insurance company of medium size with n =100000 and y = 2,273 euro,
- u_0 = minimum solvency margin (*m.s.m.*),
- 100.000 realisations for each yearly node, over a projection period of 25 years.

Figures 2a, 2b, 2c show respectively a simulated pattern of the inflation rate, the claims number and the risk reserve over a period of 25 years, making evident default situations near the barrier at level 0.

---

[1] See Wilson and Hilferty (1931); the algorithm represents a usual actuarial tool to simulate $\widetilde{S}_t$, when the first three moments are known.

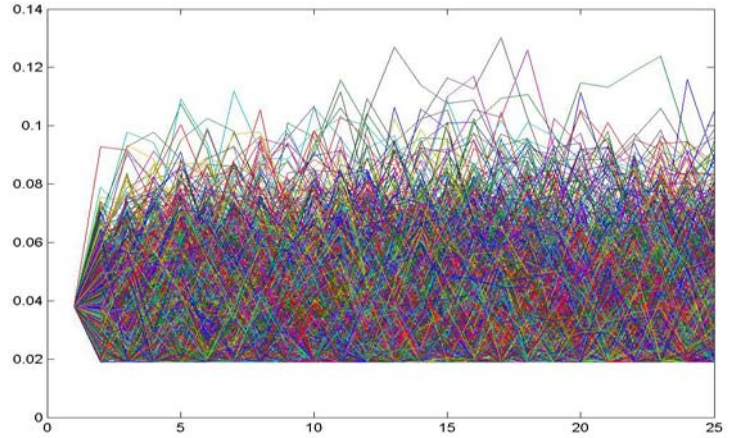

Fig. 2a. A simulated pattern of the inflation rate

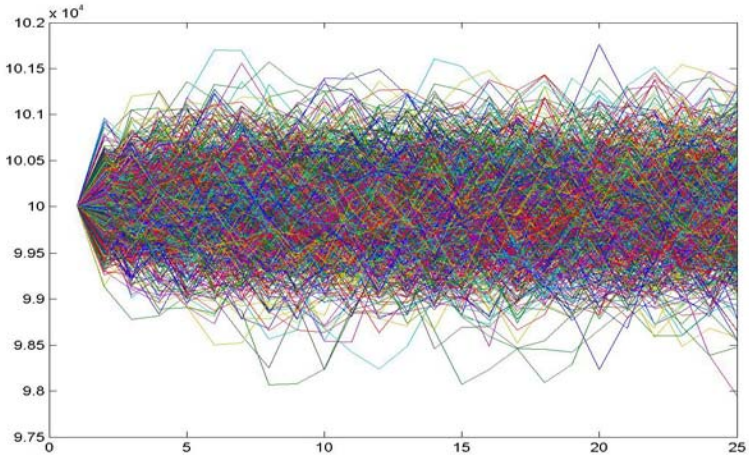

Fig. 2b. A simulated pattern of the claims number

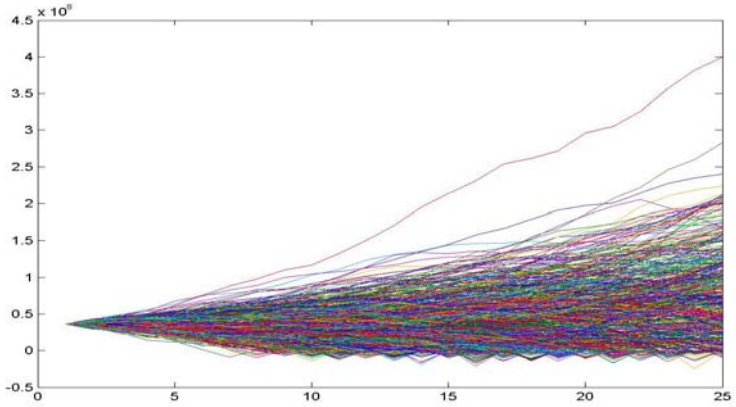

Fig. 2c. A simulated pattern of the risk reserve

First cases of default arrive not before the end of the 9th year, depending on the specific situation of the insurance portfolio analysed and moreover a similar behaviour is found out from tables 4-7, where an observed cumulative ruin probability is reported in reference to different levels of the Risk Reserve and, respectively, for three different assumptions on the average amount and the number of claims at the time t=0: in particular the cumulative ruin probability increases with the time and drops when the Risk Reserve grows over the m.s.m.

Table 4

n_0 =100000, y_0 = 2,273 euro

$u_0 =$	$\text{Prob}(\widetilde{U}_t \le 0)$				
	$t \in [0,5]$	$t \in [0,10]$	$t \in [0,15]$	$t \in [0,20]$	$t \in [0,25]$
(m.s.m.)	0.00%	5.64%	21.65%	36.00%	46.63%
1.25 * (m.s.m.)	0.00%	2.30%	14.79%	29.09%	40.59%
1.50* (m.s.m.)	0.00%	0.89%	9.76%	23.23%	35.49%
1.75* (m.s.m.)	0.00%	0.29%	6.41%	18.71%	31.09%
2.00* (m.s.m.)	0.00%	0.08%	3.98%	14.49%	26.59%

Table 5

n_0 =110000. y_0 = 2,500 euro

$u_0 =$	$\text{Prob}(\widetilde{U}_t \le 0)$				
	$t \in [0,5]$	$t \in [0,10]$	$t \in [0,15]$	$t \in [0,20]$	$t \in [0,25]$
(m.s.m.)	0.00%	5.34%	21.39%	35.70%	46.23%
1.25 * (m.s.m.)	0.00%	2.18%	14.60%	28.79%	40.38%
1.50* (m.s.m.)	0.00%	0.70%	9.49%	22.97%	35.16%
1.75* (m.s.m.)	0.00%	0.26%	6.08%	18.18%	30.54%
2.00* (m.s.m.)	0.00%	0.06%	3.85%	14.18%	26.27%

Table 6

n_0 =120000. y_0 = 2,728 euro

$u_0 =$	$\text{Prob}(\widetilde{U}_t \le 0)$				
	$t \in [0,5]$	$t \in [0,10]$	$t \in [0,15]$	$t \in [0,20]$	$t \in [0,25]$
(m.s.m.)	0.00%	5.22%	21.07%	35.40%	45.97%
1.25 * (m.s.m.)	0.00%	2.09%	14.18%	28.52%	40.06%
1.50* (m.s.m.)	0.00%	0.73%	9.42%	22.87%	35.10%
1.75* (m.s.m.)	0.00%	0.25%	6.14%	17.92%	30.24%
2.00* (m.s.m.)	0.00%	0.07%	3.82%	14.13%	26.29%

Table 7

n_0 =130000. y_0 = 2,955 euro

u_0 =	$\text{Prob}(\tilde{U}_t \le 0)$				
	$t \in [0,5]$	$t \in [0,10]$	$t \in [0,15]$	$t \in [0,20]$	$t \in [0,25]$
(m.s.m.)	0.00%	4.96%	20.83%	35.35%	45.98%
1.25 * (m.s.m.)	0.00%	1.97%	13.95%	28.23%	39.91%
1.50* (m.s.m.)	0.00%	0.73%	9.44%	22.85%	34.83%
1.75* (m.s.m.)	0.00%	0.23%	6.01%	17.85%	30.11%
2.00* (m.s.m.)	0.00%	0.07%	3.73%	13.95%	25.80%

References

1. ALTMAN E.I., Financial ratios. Discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance. 1968.
2. ALTMAN E.I., HALDEMAN R., NARAYANAN P., Zeta analysis: a new model to identify bankruptcy risk of corporations. Journal of Banking and Finance. 1977.
3. AMBROSE J.M., SEWARD A.J., Best's Rating. Financial ratios and prior probabilities in insolvency prediction. The Journal of Risk and Insurance. 55. 1988.
4. BAIONE F., La cartolarizzazione della riserva sinistri:un modello attuariale per il ramo R.C. Auto. Doctoral Dissertation in Actuarial Sciences. University of Rome "La Sapienza". 2002.
5. BARNIV R., HERSHBARGER A., Classifying financial distress in the life insurance industry. The Journal of Risk and Insurance. 1990.
6. BARONTINI R., l'efficacia dei modelli di previsione delle insolvenze: risultati di una verifica empirica. Finanza Imprese e Mercati. Anno IV. N. 1. 1992.
7. BEARD R., PENTIKAINEN T., PESONEN E., Risk Theory. Chapman and Hall. London. 1984.
8. BLUM M.P., Failing company discriminant analysis. Journal of Accounting Research. 1974.
9. BREIMAN L., FRIEDMAN J.H., OLSHENR A., STONE C.J., Classification and regression trees. Wadsworth. Belmont CA. 1984.
10. BUORO G., Sistemi di controllo preventivo della solidità delle imprese assicuratrici: una applicazione al mercato italiano. Italian Institute of Actuaries Journal. 63. 2. 1980.
11. COOLEY P.L., Bayesian and cost consideration for optimal classification with discriminant analysis. Journal of Risk and Insurance. 42. 2. 1975.
12. DEAKIN E., A discriminant analysis of predictors of business failure. Journal of Accounting Research. 1972.
13. DEAKIN E., Business failure prediction: an empirical analysis. In Financial Crises. Institutions and Markets in a Fragile Environement a cura di Altman E.I., Sametz A.W., New York. Wiley Interschence. 1977.
14. DE ANGELIS P., Modelli di valutazione delle imprese assicuratrici: alcuni risultati in condizioni di multinormalità e eteroschedasticità, Italian Institute of Actuaries Journal, Anno LI, 1-2, 1988.
15. DE ANGELIS P., DE FELICE M., OTTAVIANI R., Un modello statistico per il controllo delle compagnie di assicurazione. Quaderni Isvap 1, 1988.
16. DE ANGELIS P., GISMONDI F., OTTAVIANI R., l'approccio bayesiano nella valutazione della solvibilità delle imprese assicuratrici: una applicazione al mercato assicurativo italiano di modelli parametrici e non parametrici., Giornale dell'istituto Italiano degli Attuari, Anno LV,1-2,1992.
17. EDMISTER R.O., An empirical test of financial analysis for small business failure prediction. Journal of Finance and Quantitative Analysis. 1972.

18. EISENBEIS R.A., Pitfalls in the application of discriminant analysis in business, finance and economics. Journal of Finance. 32. 1977.
19. FORESTIERI G., La previsione delle insolvenze aziendali. Profili teorici ed analisi empiriche. Milano. Giuffrè 1986.
20. FRYDMAN H., ALTMAN E.I., DUEN-LI KAO, Introducing recursive partitioning for financial classification: the case of financial distress. Journal of Finance. 1. 1985.
21. GISMONDI F., Sulla costruzione di un modello non parametrico per la valutazione del rischio d'impresa in campo assicurativo. Department of Actuarial Science and Mathematics for Economic and Financial Decisions. University of Rome "La Sapienza". Note n. 16. 1990.
22. GISMONDI F., Modelli per la valutazione del rischio d'impresa in campo assicurativo: problemi e prospettive. Doctoral Dissertation in Actuarial Sciences. University of Rome "La Sapienza". 1992.
23. GRAHAM R.E., JONHSON L.W., Unified bayesian approach to optimal classification with discriminant analysis. Journal of Risk and Insurance. 46. 2. 1977.
24. HERSHBARGER R.A., MILLER R.K., The NAIC information system and the use of economic indicators in predicting insolvencies. The Journal of Insurance Issues and Practices. 9. 1986.
25. JOY O.M., TOLLEFSON J.O., On the financial application of discriminant analysis. Journal of Financial and Quantitative Analysis. 20. 4. 1985.
26. MARAIS M.L., PATELL J.M., WOLFON M.A., The experimental design of classification models: an application of recursive partitioning algorithm and bootstrapping to commercial bank loan classification. Journal of Accounting Research. 22. 1984.
27. MEADOR J.W., THORNTON J.H., The NAIC Early Warning System: a one-state review. Best's Review. July. 1978.
28. PENTIKAINEN T., BONSDORFF H., PESONEN M., RANTALA J., RUOHONEN M., Insurance solvency and financial strength. Helsinki: Finnish Insurance Training and Publishing Company. 1989.
29. PENTIKAINEN T., RANTALA J., A Simulation Procedure for Comparing Different Claims Reserving Methods, ASTIN Bulletin, Vol. 22, n. 2, 1992.
30. PINCHES G.E., TRIESCHMANN J.J., The efficiency of alternative models for solvency surveillance in the insurance industry. The Journal of Risk and Insurance. 41. 19 74.
31. PINCHES G.E., TRIESCHMANN J.J., Discriminant analysis. Classification results and financial distressed P-L insurers. The Journal of Risk and Insurance. 44. 1977.
32. SANDBERG C.M., LEWELLEN W.G., STAKEY K.L., Financial strategy: planning and managing the corporate leverage position. Strategic Management Journal. 8. 1987.
33. SAVELLI N., Solvency and Traditional Reinsurance for Non-Life Insurance, Proceedings of the 9th Congress on Risk Theory, University of Molise (Campobasso), 2002.
34. SINKEY JR. J.F., Problem and failed institutions in the commercial banking industry. Greenwich-Connecticut: JAI-Pres Inc. 1979.
35. SCOTT J.T., The probability of bankruptcy: a comparison of empirical predictions and theoretical models. Journal of Banking and Finance. 5. 1981.
36. VINSO J., A determination of the risk of ruin. Journal of Finance and Quantitative Analysis. 3. 1979.
37. WILCOX J.W., A gambler's ruin prediction of business failure using accounting data. Sloan Management Review. 3. 1971.
38. WILCOX J.W., A prediction of business failure using accounting data. Empirical Research in Accounting: Selected studies. Journal of Accounting Research. 1973.
39. WILCOX J.W., The gambler's ruin approach to business risk. Sloan Management Review. 1979.
40. WILKIE A.D., A Stochastic Investment Model for Actuarial Use, TFA, 39, 1986.
41. WILSON E.B., HILFERTY M., The Distribution of Chi-Square, Proceedings of the National Academy of Science, USA, 17, 1931.