

**UN ETAT DES LIEUX DES DIFFERENTES TECHNIQUES D'ANALYSE  
STATISTIQUES  
POUR LA CONSTRUCTION DES MODELES DE PREVISION DE DEFAILLANCE**

**Khadidja SADI**  
(ENSSEA).

**RESUME**

L'accroissement actuel de la taille des entreprises en faillite, donc des montants de dettes concernés, rappelle vivement la nécessité de prévoir la défaillance. La prédiction de la faillite des entreprises fait l'objet de nombreux travaux empiriques, depuis une trentaine d'années. Elle se fonde sur l'analyse économique et financière d'entreprises défaillantes et d'entreprises non défaillantes, afin de déterminer les variables, principalement comptables, qui distinguent au mieux les deux catégories d'entreprises.

L'objectif de cet article est de mettre en avant les modèles d'évaluation du risque de crédit tels que les systèmes experts et les modèles de scoring.

Nous proposons un tour d'horizon des techniques les plus utilisées afin de rendre compte de l'efficacité relative des différentes méthodes de classification utilisées.

**Introduction**

La prévision du risque de crédit est une des préoccupations majeures du monde bancaire. La mesure et la gestion du risque de crédit ont pris de plus en plus d'importance dans l'industrie bancaire, suscitant le développement de nouveaux outils et moyens permettant de minimiser les pertes.

Il est vrai qu'en octroyant un crédit, un seul objectif est recherché par les établissements financiers : générer des profits en utilisant les ressources. Cependant, ces établissements doivent anticiper la détection de défaillance des entreprises sollicitant un crédit. Les techniques d'analyses de données sont ainsi souvent utilisées afin de réaliser un premier tri dans la multitude de dossiers déposés.

Le développement récent de nouvelles techniques fondées sur l'intelligence artificielle et l'importance de leurs applications justifient qu'un nouvel état des lieux soit réalisé.

L'objectif de cet article est de mettre en avant les modèles d'évaluation du risque de crédit tels que les systèmes experts et les modèles de scoring.

**1. Utilité des modèles de prévision de défaillance**

Les modèles de prévision de défaillance sont des outils permettant de déterminer la probabilité qu'une entreprise soit défaillante dans un horizon plus ou moins lointain. Ces outils fonctionnent à partir des informations financières des entreprises sous analyse.

Dans la plupart des cas, les informations financières sont extraites de la comptabilité de ces entreprises.

L'utilisation de modèles dans le cadre de la décision de crédit est justifiée par quatre considérations :

- Premièrement, les modèles d'évaluation du risque de crédit permettent de réduire le temps requis pour l'évaluation de la demande de crédit;
- Deuxièmement, les modèles améliorent le processus de décision de crédit en spécifiant les interrelations entre les différents éléments du processus de crédit et facilitent une visualisation complète et mieux structurée de la situation;
- Troisièmement, les modèles contribuent à assurer une certaine cohérence dans les décisions de crédit relevant d'un même officier ou entre celles prises par différents officiers de crédit;
- Quatrièmement, les modèles permettent d'atténuer le biais des décideurs à accorder des poids élevés aux facteurs, indicateurs ou signaux allant en sens contraire de celui souhaité.

## **2. Les systèmes experts**

Ils sont plus fréquemment utilisés pour la mesure du risque des grands clients (corporate). L'approche des systèmes experts est de nature qualitative. Elle cherche à reproduire de façon cohérente les règles de décision des experts en matière de crédit ou leur système d'évaluation du risque (Dietsch & Petey 2003, p43). Ces règles sont déterminées de manière totalement empirique, en interrogeant les experts de crédit sur leur pratique. Cet ensemble de règles servira à constituer la base de connaissances du système expert.

(Dietsch & Petey 2003) a résumé les avantages et les importantes limites des systèmes experts comme suit :

- **Les avantages des systèmes experts**
  - Ils sont construits à partir de l'expérience et font l'objet d'une validation à posteriori.
  - Ils sont de nature qualitative, mais intègre toujours des normes quantitatives.
  - Ils reproduisent le mode de raisonnement des experts en matière de crédit.
  - Ils sont plus précis car ils intègrent des effets de l'environnement.
  - Ils ne réclament pas de disposer de longs historiques de données, car ils s'appuient sur l'expérience des experts. Ceci lui confère un avantage certain sur les méthodes de scoring.
- **Les limites des systèmes experts**
  - Ils font part d'une grande part de subjectivité des experts. Ainsi, la qualité de management d'un dirigeant célèbre n'est pas appréciée de la même façon que celle d'un dirigeant inconnu.
  - Ils donnent une part importante aux jugements des experts les plus influents dans l'institution financière.
  - Une entreprise bien notée par un système expert aura tendance à mieux se comporter qu'une entreprise moins bien notée.

- Enfin, dans la mesure où les systèmes experts reposent sur l'expérience des experts, il est difficile de définir des procédures scientifiques de test des résultats de ces systèmes, alors que les modèles de scoring peuvent être soumis à toute une batterie de tests d'inférence statistique.

### 3. Les différentes techniques de construction des modèles de scoring

Les statisticiens ont mis au point des techniques d'analyses descriptives qui permettent d'opérer une sélection dans les données disponibles et de ne conserver que les plus explicatives.

Ces techniques bien que déjà anciennes, n'ont pu être réellement appliquées qu'à partir des années 60, avec l'arrivée d'outils de calculs informatiques.

La prévision de la défaillance des entreprises a fait l'objet de nombreux travaux empiriques. Elle se fonde sur l'analyse économique et financière d'entreprises défaillantes et d'entreprises non défaillantes ou saines, afin de déterminer les variables qui séparent au mieux les deux classes. Nous proposons un état des lieux des différentes méthodes ou techniques de classification utilisées et l'énumération n'est en aucun cas exhaustive.

Ces modèles sont adaptés à la mesure du risque de défaut dans la clientèle de la banque de détail et celle des petites et moyennes entreprises (PME). Ces modèles de score sont des outils de mesure de risque qui utilisent des données historiques et des techniques statistiques.

Les modèles de crédit scoring peuvent être répertoriés en trois classes :

- La plus grande classe regroupe les méthodes statistiques paramétriques comme l'analyse discriminante linéaire ou quadratique, et les techniques économétriques.
- La deuxième classe regroupe les méthodes non paramétriques comme le partitionnement récursif et les estimateurs à noyau.
- La troisième classe, la plus récente, s'inspire de l'intelligence artificielle comme les réseaux de neurones artificiels et les algorithmes génétiques.

#### 3.1 Les méthodes statistiques paramétriques

Les méthodes paramétriques de classification statistique visent à établir une relation fonctionnelle entre les variables explicatives et la variable expliquée. Il existe trois familles de méthodes :

##### a) La méthode unidimensionnelle de Beaver<sup>1</sup>

Beaver (1966) a été le réel point de départ et la référence de nombreuses études empiriques relatives à la prédiction de défaillance. Son objectif est de classer les entreprises

---

<sup>1</sup> Pour plus de détails, voir Beaver 1966.

sur la base du ratio le plus discriminant. C'est une classification dichotomique unidimensionnelle, fondée sur un ratio unique. Cette méthode qui est de nature exploratoire n'est plus utilisée actuellement.

Son manque de robustesse est lié à l'unicité du ratio utilisé. Les analyses multidimensionnelles permettent une description plus riche de la situation de l'entreprise.

### **b) Les modèles fondés sur l'analyse discriminante**

Altman est le premier, en (1968), à avoir exploité simultanément plusieurs ratios à l'aide d'une analyse discriminante linéaire multidimensionnelle<sup>2</sup>. Altman utilise une règle de décision simple qui consiste à affecter l'entreprise au groupe dont elle est le plus proche. A la différence de l'analyse univariée le rattachement à l'une des deux classes ne se fait pas à partir d'un seul ratio mais sur la base d'une combinaison de plusieurs indicateurs. L'objectif est alors de déterminer la combinaison optimale qui permet de distinguer au mieux les deux classes.

C'est le principe de l'analyse discriminante de Fisher qui aboutit à la construction d'une fonction appelé score .

Si l'analyse discriminante reste la technique la plus utilisée pour la prévision de défaillance au sein des entreprises bancaires, elle présente cependant certaines limites. Elle repose notamment sur un ensemble d'hypothèses a priori sur la distribution des variables. D'autres outils sont donc parfois utilisés.

### **c) Les modèles de régressions avec des variables qualitatives**

Face à la contrainte de multinormalité qui est rarement respectée empiriquement dans l'analyse linéaire et quadratique, d'autres méthodes sont proposés : les techniques économétriques sur variables qualitatives comme le modèle Probit et le modèle Logit. Ils supposent une distribution différente des variables comptables.

Le modèle Logit suppose que les erreurs suivent une loi logistique. Le modèle Probit suppose que les erreurs suivent une loi normale. Dans les deux modèles, la variable endogène est une variable qualitative. Elle prend la valeur 0 ou 1 selon que l'entreprise est défaillante ou non.

Le modèle Probit a été le moins utilisé dans la prédiction de défaillances des entreprises.

Cette technique présente deux avantages :

- Elle ne nécessite pas d'hypothèses à priori sur la distribution des variables.
- Elle tient compte de la probabilité à priori d'appartenance à un groupe.

---

<sup>2</sup> Pour plus de détails, voir par exemple Bardos 2001.

Par contre elle pose des problèmes pour la prévision des entreprises qui n'appartiennent pas à l'échantillon de construction. En effet, l'échantillon de prévision doit contenir exactement les mêmes pourcentages d'entreprises des deux classes que l'échantillon ayant servi à la construction du modèle. Cette hypothèse présente un réel obstacle à l'utilisation de cette technique.

### **3.2 Les méthodes statistiques non-paramétriques**

On utilise les modèles non-paramétriques lorsque les ratios ne suivent pas de loi multi normale et que les erreurs ne suivent pas une loi logistique ou une loi normale.

Ces méthodes ne nécessitent aucune hypothèse relative à la distribution des variables. Il existe deux catégories de techniques : le partitionnement récursif et l'estimation non paramétrique des lois de distribution des ratios.

#### **a) Les modèles de partitionnement récursif**

C'est un algorithme qui consiste à construire un arbre décisionnel<sup>3</sup>. Comme le signale (Perez M. 2002, p105), à chaque nœud de l'arbre les entreprises sont séparées en deux groupes en fonction d'un ratio et de sa valeur seuil. La construction se fait pas à pas, en minimisant le risque d'erreur probable de classement, à chaque nœud. On aboutit ainsi à la construction d'un arbre de décision, avec des sous-groupes d'entreprises auxquels sont associés des taux d'erreur (figure 2.3). Un nœud est considéré comme final ou feuille quand aucune diminution d'erreur de classement n'est plus possible.

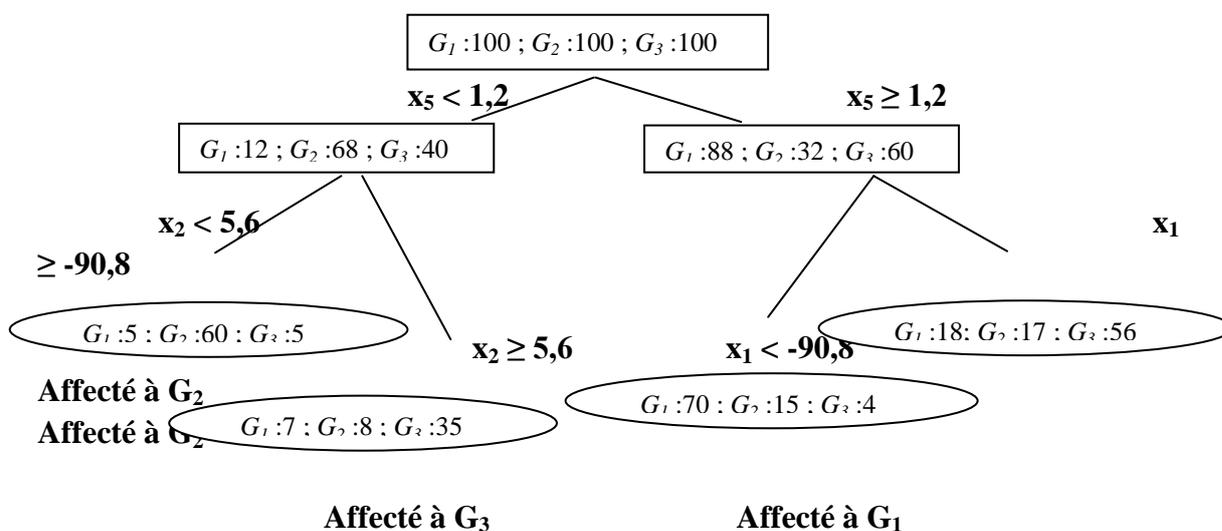
Le partitionnement récursif est néanmoins peu utilisé dans la prévision de défaillance. Contrairement à l'analyse discriminante, il présente les inconvénients suivants :

- Cette méthode est très sensible au choix des probabilités de défaut à priori et des coûts d'erreur de classement.
- Il requiert qu'à chaque nœud de l'arbre de décision soit déterminée une valeur seuil pour le ratio considéré, ce qui nuit à la stabilité temporelle de l'indicateur.
- L'arbre se construit en chaînage avant et ne permet pas de retour en arrière dans l'arborescence. Il peut arriver qu'un indicateur soit utilisé deux fois ce qui rend l'interprétation difficile.

---

<sup>3</sup> . Voir Frydman H., Altman E., et Kao D, 1985.

**Figure 2.3 : Le partitionnement récursif**



Trois groupes  $G_1$ ,  $G_2$ ,  $G_3$  sont à discriminer. A chaque nœud sont représentés sur le schéma les effectifs des groupes et la coupure.

Source : Bardos 2001, p 85

### b) Les modèles d'estimation non-paramétrique des lois de distribution des ratios

Afin de résoudre le problème de non-multinormalité des ratios, il est possible d'estimer la loi de probabilités des ratios par les méthodes non-paramétriques. Les méthodes non-paramétriques ne s'appuient pas sur une forme spécifique de loi déterminée a priori. Le principe est de faire un lissage des observations permettant d'estimer une fonction de densité par une moyenne pondérée locale des fréquences. C'est le principe de la méthode du noyau<sup>4</sup>.

Un point est pondéré d'autant plus fortement qu'il est proche du point  $x$  en lequel la fonction de densité est estimée. Le paramètre de lissage est crucial car il détermine la taille du voisinage, c'est-à-dire le nuage de points pris en considération pour estimer localement la fonction de densité; il fixe également la pondération de chaque point. Il conditionne donc la précision de l'estimation.

Le principal défaut de cette technique statistique est le nombre important de données qu'elle requiert, par rapport à l'estimation paramétrique de la loi de distribution des ratios.

### 3.3 Les méthodes issues de l'intelligence artificielles

<sup>4</sup> Voir Calia et Ganugi 1997.

Face aux nombreuses contraintes liées aux méthodes statistiques traditionnelles, d'autres méthodes issues de l'intelligence artificielle ont vu le jour. Elles relèvent plus précisément, de l'apprentissage automatique. Il existe deux catégories de méthodes : les réseaux de neurones artificiels et les algorithmes génétiques.

#### **a) Les réseaux de neurones artificiels**

La technique des réseaux de neurones a pris une ampleur considérable au cours des années 90, tant par le développement de la recherche et la multiplication des modèles auxquels ils ont donné lieu, que par l'extension des applications à des domaines nouveaux et très variés. Le principe des réseaux de neurones artificiels sera plus détaillé dans le chapitre suivant.

#### **b) Les algorithmes génétiques**

Leur fonctionnement est assez proche des réseaux de neurones. Un algorithme génétique traite les informations disponibles, les ratios comptables sélectionnés et la situation de chaque entreprise (défaillante ou saine). L'algorithme, par itérations successives, déduit des régularités, des règles générales du processus de défaillance qu'il peut ensuite appliquer à d'autres entreprises<sup>5</sup>. Le principe d'un algorithme génétique est d'imiter le processus d'évolution naturelle des espèces. D'après (Varetto 1999), Il se fonde sur les trois principes du Darwinisme :

- sélection des individus
- reproduction des individus : 'les meilleurs'
- croisements génétiques et mutations aléatoires des gènes.

L'algorithme génétique analyse donc des populations de solutions - c'est-à-dire des modes de traitement de l'information à prévoir la défaillance - et évalue leur qualité, en l'occurrence le taux de bon classement des entreprises qu'elles fournissent. Il aboutit alors à une nouvelle population de solutions, plus homogène. La procédure est réitérée et s'arrête lorsque la population de solutions atteint un certain niveau d'homogénéité, fixé préalablement.

### **4. La méthodologie commune des différentes techniques de crédit scoring.**

De nombreux travaux ont tenté avec succès d'évaluer le risque de défaut des entreprises en se basant sur l'analyse financière. Le principe général est similaire, bien qu'il existe une variété de techniques. Disposant de données comptables d'entreprises et connaissant la classe d'appartenance durant la période d'observation, ces techniques parviennent à sélectionner les variables comptables les plus discriminantes puis à établir une relation statistique entre ces variables et la classe d'appartenance des entreprises. Une probabilité de défaillance est parfois estimée. La qualité de l'indicateur de risque ainsi élaboré est jugée selon les erreurs de classement : le pourcentage d'entreprises défaillantes considérées comme saines (erreur de

---

<sup>5</sup> Pour une illustration, voir Varetto F., 1998

type I) et inversement, le pourcentage d'entreprises saines considérées comme risquées (erreur de type II).

D'après (Refait 2004), il existe quatre étapes nécessaires à la construction d'un modèle de prévision de défaillance : la construction d'un échantillon, la sélection des variables explicatives de la défaillance, le choix du mode de classification des entreprises entre les deux classes, et enfin l'estimation de la qualité de prévision effectuée.

#### **4.1 La construction de l'échantillon**

L'élaboration de l'échantillon joue un rôle important dans la qualité de la prévision. Il est basé sur deux piliers importants : le premier est la définition de la défaillance et le choix de l'horizon de prévision sont importants, le second est le choix des caractéristiques des entreprises sélectionnées.

- **Définition de la défaillance et de l'horizon de prévision** : elle est primordiale dans la construction de l'échantillon puisqu'elle caractérise les deux populations qui déterminent l'indicateur de risque. A partir d'un échantillon global d'entreprises, deux sous-échantillons doivent être construits : l'échantillon des entreprises dites 'saines' et l'échantillon des entreprises dites 'défaillantes'.

La détermination de ce qu'est une entreprise défaillante est un problème délicat, susceptible de nuire à la robustesse de l'indicateur. La majorité des auteurs considère la défaillance comme l'ouverture d'une procédure judiciaire. Cependant, certains auteurs considèrent la défaillance comme un défaut de paiement. Le comité de Bâle met en effet l'accent sur l'estimation du risque de défaut et non sur celle du risque de défaillance.

Parallèlement à la définition de la définition de la défaillance, le choix de l'horizon de prévision est important. Il conditionne la date de la défaillance des entreprises et la date des données qui vont fonder la discrimination, l'écart de temps entre les deux étant l'horizon de prévision. Un arbitrage est à faire entre une échéance trop proche, dont l'intérêt est limité car elle ne permet pas d'éviter la faillite et limiter les pertes, et une échéance trop lointaine, qui ne permet pas une prévision précise. Les études diffèrent légèrement sur ce point, mais généralement, deux horizons sont choisis : un an et trois ans avant la défaillance.

- **Les caractéristiques des entreprises : représentativité et homogénéité de l'échantillon** : la construction des deux sous-échantillons pose le double problème de la représentativité et homogénéité des deux sous-échantillons. L'échantillon de base doit être représentatif : du point de vue des secteurs d'activités, de la taille des entreprises mais également du rapport entre le nombre d'entreprises et d'entreprises saines. Néanmoins, cette représentativité crée une homogénéité qui est susceptible de créer un biais statistique : des facteurs explicatifs peuvent être masqués par des effets sectoriels ou des effets de taille. Afin de concilier ces deux exigences et d'améliorer la qualité de prédiction de la défaillance, les solutions suivantes ont été préconisées :

- Une bonne représentativité est possible lorsque le nombre d'entreprises exploitées est élevé. Elle doit cependant être vérifiée par la comparaison entre les caractéristiques de l'échantillon et celle de la population globale.
- Le fait de limiter la portée de l'indicateur à une population cible, par un secteur économique limité ou un intervalle restreint de taille d'entreprises.

L'inconvénient de cette méthode est qu'elle exige que soit élaboré un indicateur par secteur et par taille. Il se pose alors le problème du choix d'agrégation des secteurs économiques et de l'ampleur des intervalles de taille.

## 4.2 La sélection des variables explicatives

La prévision de la défaillance est réalisée à partir de données comptables. Les variables explicatives retenues sont généralement des ratios comptables diverses. Mais les caractéristiques des entreprises prises en compte sont similaires.

Conformément à l'enseignement de l'analyse financière, les trois éléments les plus corrélés à la défaillance sont : la rentabilité de l'entreprise (économique ou financière), la structure de son bilan et sa capacité de remboursement.

- **La rentabilité de l'entreprise** : La rentabilité économique met en considération un résultat d'exploitation. La rentabilité financière met en rapport une variable de résultat avec le capital financier. Elle est souvent exprimée par le ratio résultat global sur capitalisation boursière.

- **La structure du bilan** : elle met en rapport un élément de l'actif et un élément du passif. Elle permet de rendre compte d'éventuels déséquilibres qui peuvent causer le non-respect des engagements financiers de l'entreprise.

- **La capacité de remboursement** : elle est mesurée de différentes manières : par le ratio cash flow sur dette totale, par le ratio intérêts dus sur chiffre d'affaires, et par le ratio revenu d'exploitation sur intérêts dus.

Conjointement aux données comptables, des informations liées à l'organisation de l'entreprise ou à la nature de son financement, par exemple, peuvent être utilisées.

## 4.3 La méthode de validation

Une fois l'échantillon constitué, il convient d'établir un lien statistique entre les variables explicatives retenues et l'appartenance à l'une des deux classes. La qualité de prévision de l'indicateur de risque est mesurée par le taux de bons classements.

La divergence entre les études se situe dans la construction de l'échantillon qui sert à calculer les taux de bons classements, nommé échantillon test. Les études qui utilisent une très grande base de données mesurent le taux de bons classements à partir d'un autre échantillon différent de l'échantillon de construction tel que Bardos (1998) et Bardos M. et Zhu W.H. (1997). Dans d'autres études, l'échantillon de base ou initial est le même que celui qui est utilisé pour la validation ce qui rend l'indicateur de risque peu performant.

En général trois méthodes de validation sont utilisées : la validation croisée, le jack-knife, et le bootstrap.

- **La validation croisée** : en anglais cross-validation. Une validation croisée d'ordre  $V$  consiste à découper l'échantillon initial en  $V$  sous-échantillons de taille approximativement égale, puis à estimer l'indicateur de risque à partir de  $V-1$  groupes pour enfin calculer le taux de bons classements sur le  $V^{\text{me}}$  groupe. Le taux de bons classements théorique est estimé par la moyenne des  $V$  taux de bons classements empiriques.

- **Le jack-knife** : il consiste à construire  $n$  sous-échantillons de  $n-1$  entreprises, en négligeant successivement chacune des  $n$  entreprises de l'échantillon initial. Les  $n$  taux de bons classements sont calculés sur l'échantillon initial à partir des  $n$  indicateurs estimés des  $n$  sous-échantillons. Le taux de bons classements théorique est estimé par la moyenne des  $n$  taux de bons classements empiriques.

- **Le bootstrap** : il consiste à construire un échantillon à partir d'un tirage aléatoire avec remise des entreprises de l'échantillon initial. L'échantillon est généralement de même taille que l'échantillon initial. Un indicateur est élaboré à partir de l'échantillon construit et le taux d'erreur de classement qu'il procure est calculé sur l'échantillon initial. Un taux d'erreur apparent est calculé sur l'échantillon construit. L'écart entre les deux taux d'erreur de classement est mesuré. Le taux d'erreur de classements théorique est estimé par le taux d'erreur apparent calculé sur l'échantillon initial auquel s'ajoute la moyenne empirique des écarts.

Le niveau des taux de bons classements dépend donc, du mode de validation choisi. Les trois modes de ré-échantillonnage cités permettent de calculer une moyenne de bons classements.

Les recherches académiques de méthodes de prévision de la faillite restent actives, pendant que leur utilisation opérationnelle s'intensifie. Nous pourrions rappeler leurs avantages et leurs inconvénients (Refait 2004):

Les avantages :

- Systématisation du traitement de l'information
- Gain de temps et de coût

Les inconvénients :

- Problème de la robustesse temporelle et sectorielle,
- Problèmes statistiques (communs à l'ensemble des techniques évoquées),

Mais notre objet est plutôt de mettre en perspective les différentes méthodes.

## **Conclusion**

L'analyse discriminante linéaire, après une phase exploratoire lors des années soixante-dix, est la méthode la plus utilisée du point de vue opérationnel. Elle fournit en effet pour le moment les prévisions les plus robustes. De plus, la fonction score qu'elle permet d'établir est riche d'applications utiles pour les praticiens, comme le calcul de probabilité a posteriori ou la construction de classes de risques (voir Bardos 2001).

D'autres méthodologies ont été mises au point comme l'analyse discriminante quadratique, les régressions sur variables qualitatives ou encore les techniques non-paramétriques comme l'estimateur fondé sur la méthode du noyau. Leur objectif était principalement d'éviter les contraintes statistiques imposées par la discrimination linéaire. Ces techniques ne sont néanmoins pas aussi utilisées dans la pratique que la discrimination linéaire, principalement à cause de problèmes de robustesse (pour l'analyse discriminante quadratique) ou de disponibilité des données (pour les techniques non-paramétriques).

Les récentes techniques empruntées à l'intelligence artificielle, tels les réseaux de neurones et les algorithmes génétiques, connaissent un très grand engouement académique. Elles permettent en effet de bonnes prévisions tout en présentant l'avantage de ne pas exiger de restrictions statistiques. De plus, les algorithmes génétiques sont particulièrement robustes, notamment parce qu'ils ne sont pas soumis à des contraintes mathématiques telles que la dérivabilité ou la continuité des fonctions utilisées. Leur utilisation, principalement celle des algorithmes génétiques, reste exploratoire. Leur haut degré de technicité ne leur permettait pas de fournir d'éléments explicatifs clairs de la défaillance, inconvénient qui semble pouvoir être réduit.

Une critique souvent formulée porte sur l'utilisation exclusive de données comptables : une amélioration significative de la détection de la faillite serait sans doute plus apportée par l'exploitation d'informations autres que comptables et financières, par exemple, la stratégie globale des entreprises.

Le développement des systèmes-experts s'inscrit dans la même logique. Ils sont construits par l'application des arbres de décision puis des réseaux neuronaux et des algorithmes génétiques à des critères qualitatifs utilisés par les experts, tels que la qualité de la gestion des ressources humaines ou le degré de concentration des clients. Les systèmes-experts permettent non

seulement une compréhension plus riche du processus de défaillance mais également une bonne prévision de la faillite.

Rappelons néanmoins pour finir que les modèles de prédiction du risque de faillite ne constituent qu'une aide à la décision parmi d'autres et se doivent d'être enrichis par d'autres formes d'information.

## **BIBLIOGRAPHIE**

1. Bardos M. et Zhu W.H. (1997). Comparaison de l'analyse discriminante linéaire et des réseaux de neurones – application à la détection de défaillance d'entreprises, *Revue statistique appliquée*, tome 45, n°4.
2. Bardos Mireille (1998), Detecting the risk of company failure at the bank de France, *Journal of banking and finance*, vol. 22.
3. Bardos Mireille (2001). *Analyse discriminante, application au risqué et scoring financier*. Dunod.
4. Calia P. et Ganugi P. (1997), Kernel and nearest neighbour discriminant analysis : business failure classification in industrial district, *Applied stochastic models and data analysis*, Colloque Capri.
5. Dietsch Michel, Petey Joël (2003). *Mesure et gestion du risque de crédit dans les institutions financières*. Edition d'Organisation.
6. Frydman H., Altman E., et Kao D. (1985), Introducing recursive partitioning for financial classification: the case of financial distress. *Journal of finance*, vol. 40, n°1.
7. Perez Muriel (2002). *De l'analyse de la performance à la prévision de défaillance*. Thèse de doctorat en Sciences de gestion, Lyon III.
8. Perez Muriel (2006). *Artificial neural networks and bankruptcy forecasting: a state of the art*. *Neural Comput & Applic*.
9. Refait Catherine (2004), *La prévision de la faillite fondée sur l'analyse financière de l'entreprise : un état des lieux*. *Economie et prévision*, n°162.
10. Varetto F. (1998), Genetic algorithms applications in the analysis of insolvency risk, *Journal of banking and finance*, vol. 22, n°10-11, October.

