

データのモニタリング手法について

岡田 絵理

 RDB 日本リスク・データ・バンク株式会社

調査企画部

Data Monitoring Method

Eri Okada
Senior Manager,
Research & Planning Department,
The Risk Databank of Japan, Limited

要 旨

信用スコアリングモデルの構築はRDBの主なサービスの一つであるが、モデル構築の事前作業として財務指標に関する分析を実施してデフォルト先の判別に有効な要因を把握し、その知見をモデル構築に活用している

キーワード: スコアリングモデル、AR、説明変数の有効性

資料の構成

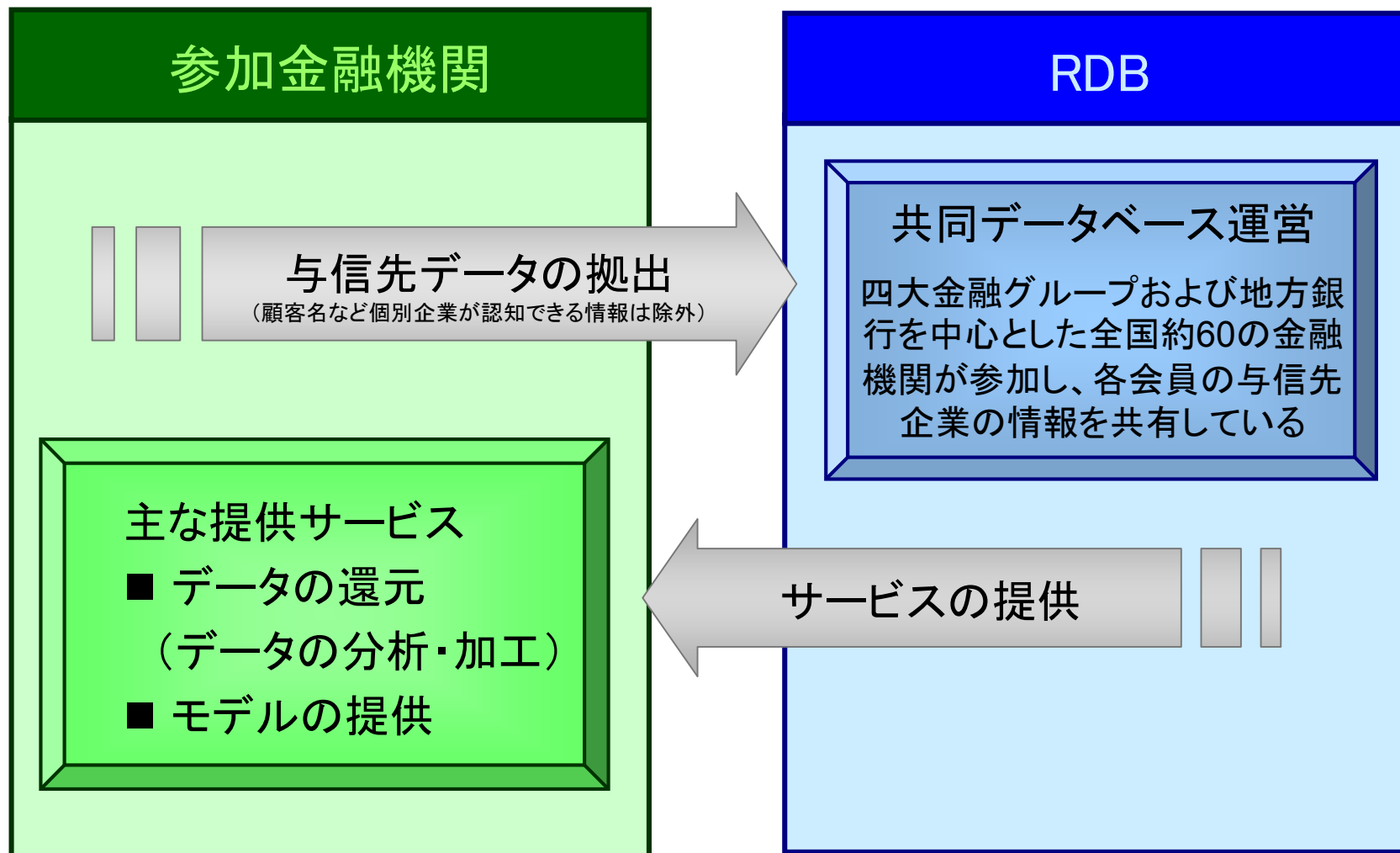
1. 日本リスク・データ・バンク(RDB)について

- 1-1. RDBの主な事業内容
- 1-2. RDBデータベースの特徴
- 1-3. RDBスコアリングモデルの特徴

2. 財務指標に関する分析

- 2-1. 分析対象データについて
- 2-2. 指標間の相関の確認
- 2-3. 単一指標分析によるデフォルト判別力の確認
- 2-4. 指標間の相対比較によるデフォルト判別力の確認

1-1. RDBの主な事業内容



1-2. RDB共同データベースの特徴

会員の地域構成

現時点の参加会員数:61会員

【中国地方】
5 行

【近畿地方】
7 行

【関東・
甲信越地方】
9 行

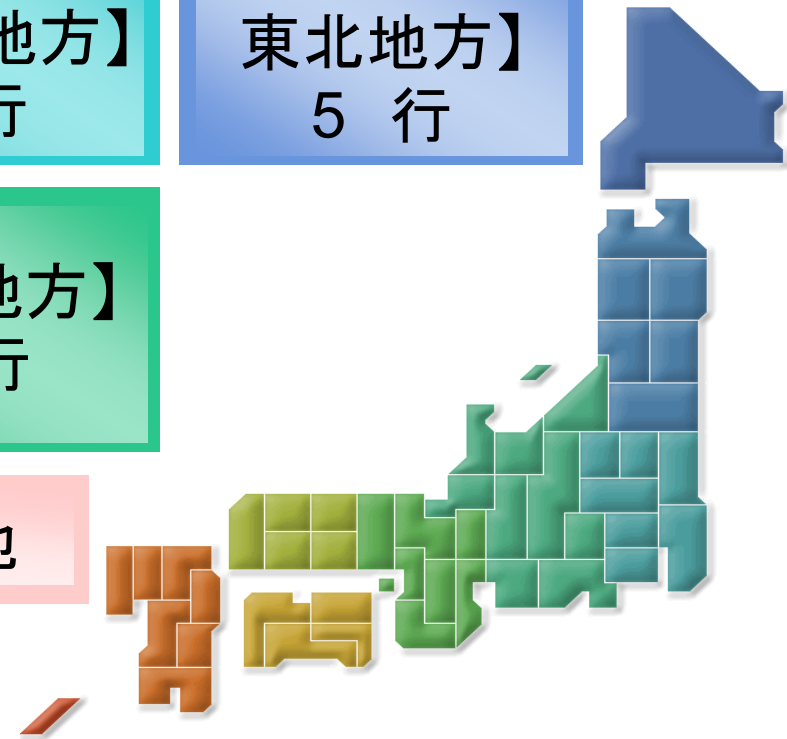
【北海道・
東北地方】
5 行

【九州地方】
9 行

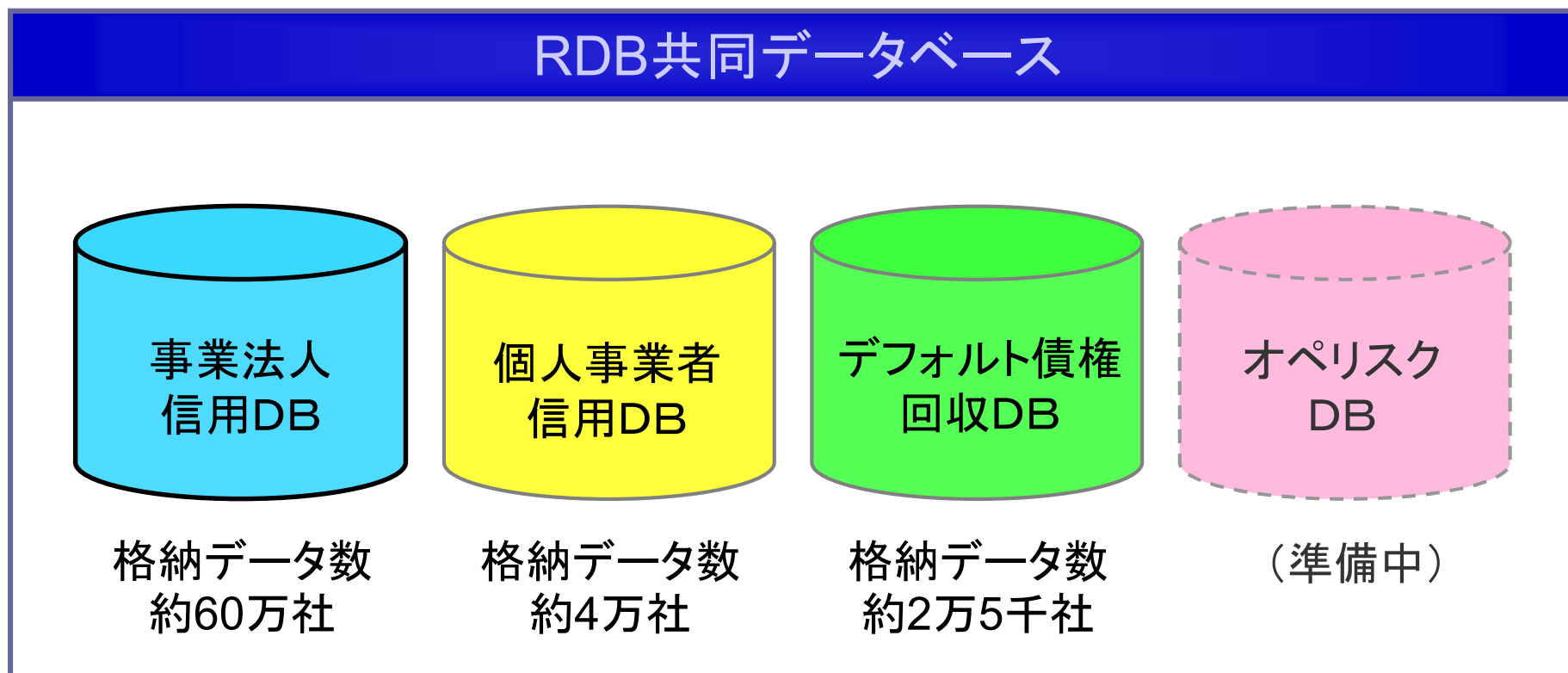
【四国地方】
4 行

【中部地方】
9 行

【大手行等】メガバンク、商社、リース他



1-2. RDB共同データベースの特徴

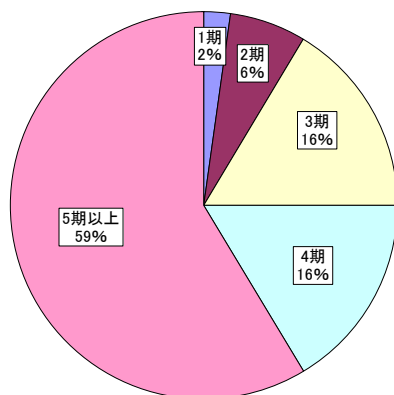


1-2. RDB共同データベースの特徴

事業法人信用 データベース

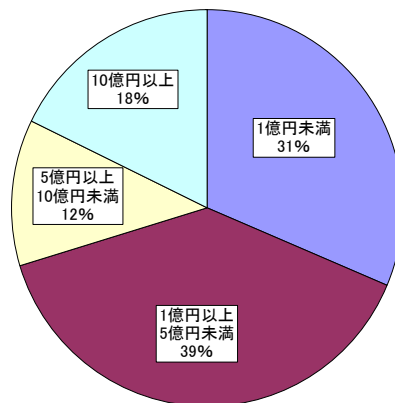
- 日本全国のデフォルト事象の約7割をカバーしている
- 時系列の財務情報が取得できる
- メガバンク、地銀、第二地銀、商社・ノンバンク等の保有データが継続的に集積され、ベンチマーク性が高い

【1社あたり決算期数】



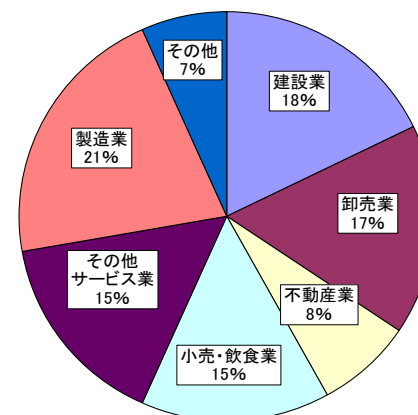
(平均: 5.1期)

【売上規模構成】

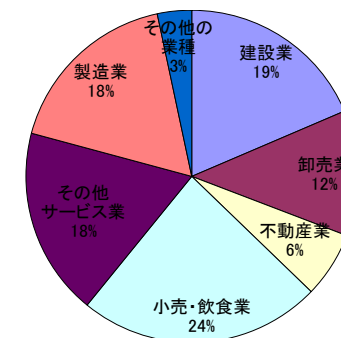


(中央値: 206百万円)

【業種構成】



(参考) 中小企業白書



1-2. RDBデータベースの特徴

事業法人データベースに関する主なサービス

信用情報、財務情報等のデータ還元

モデル構築、検証等に利用可能なデータの提供

RDB企業デフォルト率の計測

過去の実績デフォルト件数から算出した企業デフォルト率を月次にて提供

中小企業財務統計資料の発行

代表的な財務指標に関する集計結果を提供

スコアリングモデルの提供(後述)

1-3. RDBモデル(RDBによる信用スコアリングモデル)の特徴

RDBモデルの特徴

- 汎用型、業種別(6業種)のモデル
- 説明変数として財務指標を利用
- スコア、推定デフォルト率を算出
- アウトサンプルデータによる検証を毎年実施

財務指標に関する分析

- ◆ 指標間の相関の確認
- ◆ 単一指標分析
- ◆ 指標間の判別力の相対比較

定期的にデータのモニタリングを実施し、その情報をモデル構築に活用する。

2. 財務指標に関する分析

単一指標分析によるデフォルト判別力の確認

年度別データにより指標ごとのデフォルト判別力を計測し、判別力の水準の変動等を確認する。

(注)判別力の評価にはAR(AUCと等価)を用いた。ARの定義は参考資料を参照。

指標間の相対比較によるデフォルト判別力の確認

年度別データによりモデルを試作し、採用される説明変数を比較する。説明変数は基本的にはデフォルト判別に有効な財務指標から採用されるので、当該指標の相対的な判別力を時系列にて確認する。

分析結果からデフォルト判別にはどのような財務指標が有効であるか、時系列でのデフォルト判別力に変化があるかを把握できる

2-1. 分析対象データについて

分析対象データ

- ◆ 2003年度から2008年度までの6年間のステータス（信用状態）に関するデータ
- ◆ 財務情報はステータス認定時点の前年の決算期のものを利用

財務指標

- ◆ 財務分析等で使用されることが多い財務指標のうち、欠損値や異常値の少ない約100指標を対象とする
- ◆ 分析に利用しやすいように上下限値の設定や対数化を行う

2-2. 指標間の相関の確認

分析対象とした指標間の相関を確認するため、因子分析を行う。その結果を利用し、この後の分析結果の解釈を容易にするため指標をいくつかのカテゴリに分類する。

因子分析にはSASのFACTORプロシジャを利用し、主因子法により因子を抽出した。その結果、11の因子が得られた。財務指標は因子負荷量の大きさにより分類し、各カテゴリに含まれる財務指標の特徴から因子の解釈を行った。

(注)11の因子のいずれに対しても負荷量が小さい指標についてはその他として分類し、計12のカテゴリとした。



2-3. 単一指標分析によるデフォルト判別力の確認

年度別データに対するAR上位5指標

指標単独でのデフォルト判別力の有効性を確認するため、年度別データにより指標ごとにARを計測した。ARの算出にはSASのFREQプロシジャを利用した。

ARの上位5指標をみると、いずれの年度においても金利負担、債務負担、自己資本に関する指標が判別力が高いという結果であった。

年度別データに対するARの上位5指標が属するカテゴリ

債務負担

金利負担

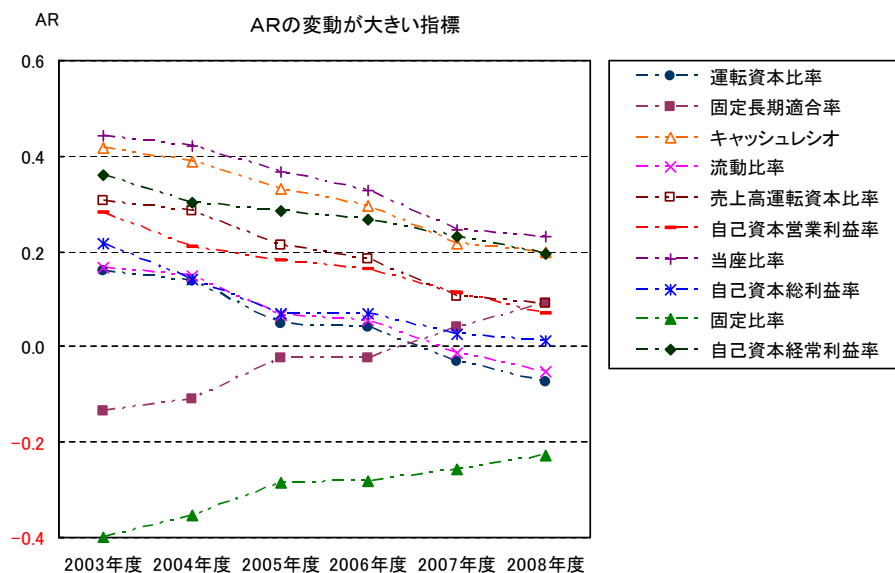
自己資本

2-3. 単一指標分析によるデフォルト判別力の確認

時系列でのARの変動

時系列でみた場合に指標ごとのデフォルト判別力に変動があるかをみるため、ARの変動の大きさを確認する。

【ARの変動が大きい指標】



資金効率

運転資本比率
固定長期適合率
流動比率
売上高運転資本比率

自己資本

自己資本営業利益率
自己資本総利益率
固定比率
自己資本経常利益率

流動性

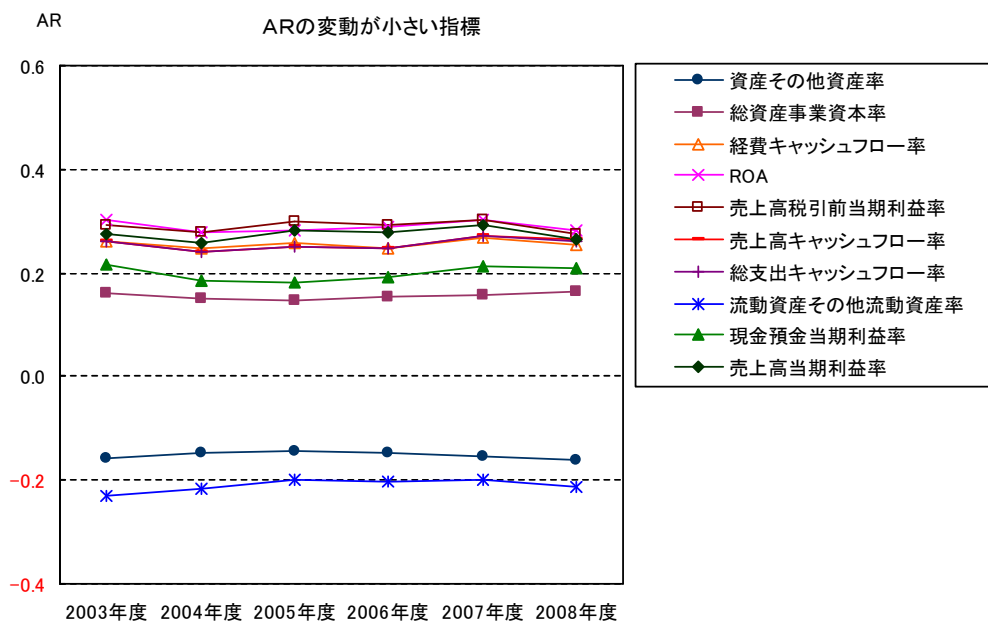
キャッシュレシオ
当座比率

(注) 変動の大きさは年度間の最大値と最小値の差により判断した。ARの正負の意味は、正の場合は指標値が大きいほど債務者の信用力が高いことを示す。

2-3. 単一指標分析によるデフォルト判別力の確認

時系列でのARの変動

【ARの変動が小さい指標】



(注) 変動の大きさは年度間の最大値と最小値の差により判断した。ARの正負の意味は、正の場合は指標値が大きいほど債務者の信用力が高いことを示す。

不健全資産

資産その他資産率
総資産事業資本率
流動資産その他流動資産率

収益性

経費キャッシュフロー率
ROA
売上高税引前当期利益率
売上高キャッシュフロー率
総支出キャッシュフロー率
現金預金当期利益率
売上高当期利益率

2-3. 単一指標分析によるデフォルト判別力の確認

分析結果

- ◆ 債務負担、金利負担に関する指標はいずれの年度においてもARが高く、変動も比較的小さいため、デフォルト判別に有効であり時系列でもその有効性は比較的安定しているといえる
- ◆ 流動性、自己資本、資金効率に関する指標はARの変動が大きく、直近になるにつれ判別力は低下している
- ◆ 収益性、不健全資産に関する指標はARの変動が小さい

2-4. 指標間の相対比較によるデフォルト判別力の確認

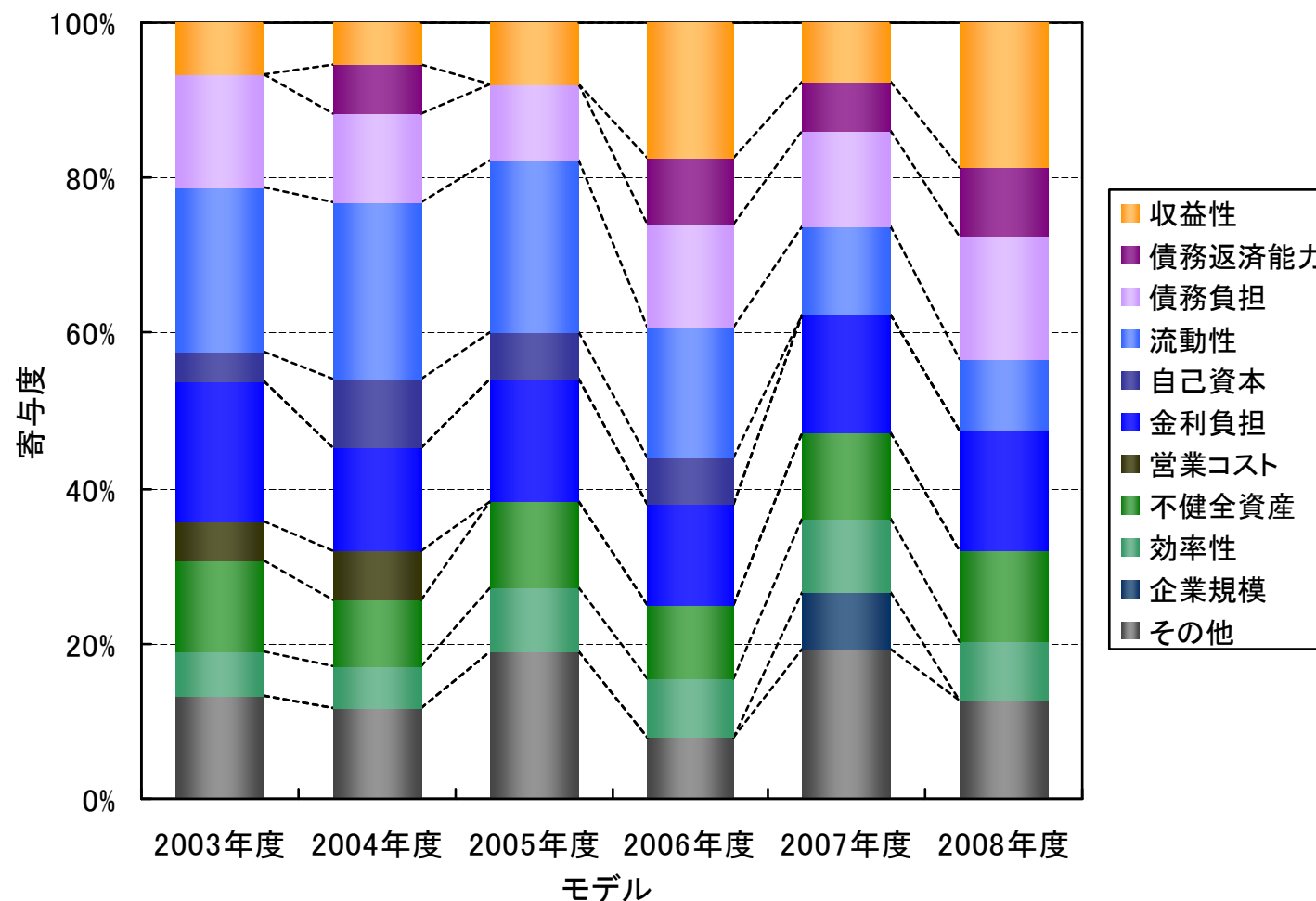
モデルの説明変数は基本的にはデフォルト判別に有効な指標から選択されるので、これによって指標の相対的な判別力を確認することができる。

最初に年度別データによりモデルを構築する。モデルはロジスティック回帰モデルとし、SASのLOGISTICプロシジャを利用した。変数選択は変数増加法により行い、各モデルとも15指標を選択した。

次に各モデルに採用された15指標について、2-2の指標分類結果に基づき12のカテゴリに分類し、各カテゴリのモデルに対する寄与度を集計した。

2-4. 指標間の相対比較によるデフォルト判別力の確認

【各カテゴリのモデルに対する寄与度】



2-4. 指標間の相対比較によるデフォルト判別力の確認

分析結果

- ◆ 各モデルについて寄与度が高い指標カテゴリをみると、2003年度から2005年度までは流動性、ついで金利負担の寄与が大きかった。その後2006年度は収益性、流動性、2007年度は金利負担、債務負担、2008年度は収益性、債務負担の順に寄与度が高かった
- ◆ 指標カテゴリ別に見た寄与度の時系列での変動に注目すると、流動性、自己資本については寄与度が低下傾向にある一方で、収益性、債務返済能力、効率性については寄与度は上昇傾向にあった
- ◆ 債務負担、金利負担についてはいずれの年度のモデルにおいても寄与度は比較的高く、変動は小さい

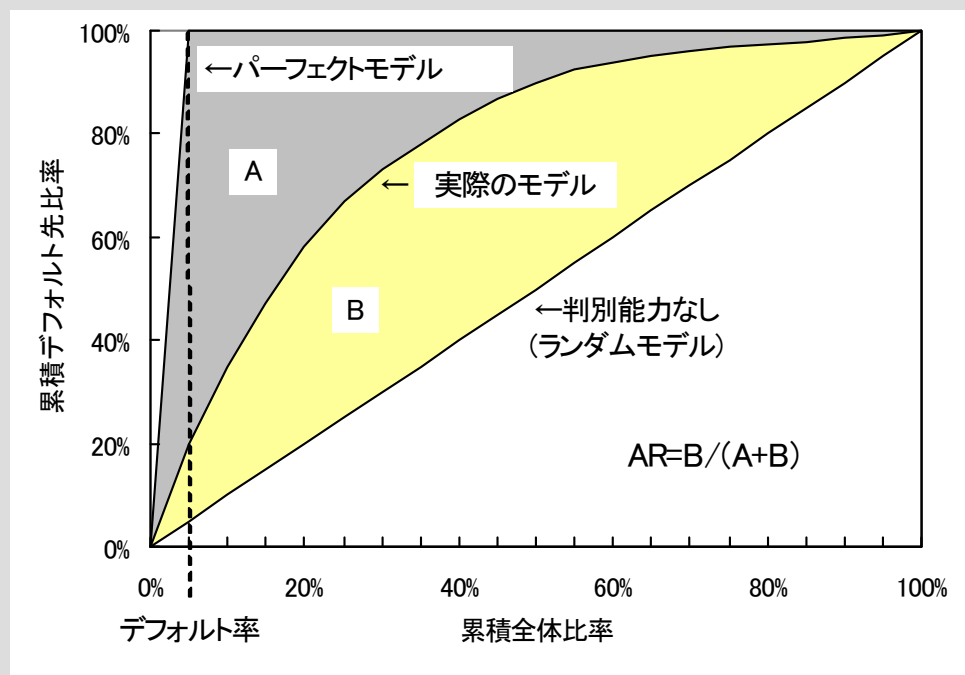
総 括

本稿では、デフォルト判別にはどのような要因が有効か、またその有効性に変動があったかということについて分析した。

その結果、債務負担、金利負担といった要因は、分析対象とした2003年度～2008年度のいずれの期間においてもデフォルト判別に有効であり、債務者が過大な債務を抱えていないか、金利負担の水準は妥当であるかといったポイントはデフォルト判別において重要であると考えられる。

今回取り上げた手法は、過去に構築したモデルの有効性を判断する際にも活用できる。一般的にはモデル改訂の基準にARが利用されることが多いが、モデル採用指標についても定期的に検証を実施し、想定どおりの有効性が維持されているか確認する必要があるだろう。

参考資料 ARについて



$$AR = 2 \times AUC - 1$$

AR (Accuracy Ratio) は、モデルの判別能力を評価する統計量である。データセットに含まれる企業をモデルスコアの昇順に並べ、デフォルト先の累積構成比率を縦軸、対象企業全体の累積構成比率を横軸としてプロットした曲線をパワーカーブという。モデルに判別能力が全くない場合、この曲線は原点と縦横軸の値がともに100%の点を結ぶ直線となる(ランダムモデル)。一方、理想的な判別能力がある場合は、この曲線は点(全体のデフォルト先件数比率, 100%)と原点および点(100%, 100%)を結ぶ2本の直線で表される(パーフェクトモデル)。実際のモデルとランダムモデルの曲線には含まれる部分と、パーフェクトモデルとランダムモデルの曲線には含まれる部分の面積比をARという。

参考文献

- [1] Larry Hatcher, 「A Step-by-Step approach to Using SAS for Factor Analysis and Structural Equation Modeling」, SAS
- [2] 永田靖, 棟近雅彦, 「多変量解析入門」, 株式会社サイエンス社
- [3] 田中豊, 森川敏彦, 山中竹春, 富田誠, 「一般化線形モデル入門」, 共立出版株式会社
- [4] 丹後俊郎, 山岡和枝, 高木晴良, 「ロジスティック回帰分析」, 朝倉書店
- [5] 山下智志, 川口昇, 敦賀智裕, 「信用リスクモデルの評価方法に関する考察と比較」, 金融庁金融研究センター(2003)
- [6] 白田佳子, 「倒産予知の実務」, 日本経済新聞社
- [7] 藤井真理子, 竹本遼太, 「中小企業のデフォルトリスクとその期間構造: 大規模財務データによる実証分析」, FSAリサーチレビュー第6号(2010)
- [8] Giulio Bottazzi, Marco Grazzi, Angelo Secchi, Federico Tamagni, 「Financial and Economic Determinants of Firm Default」, Laboratory of Economics and Management Working Paper Series(2009)
- [9] 大久保豊, 尾藤剛, 「プライムレート革命」, 社団法人金融財政事情研究会
- [10] 柳澤健太郎, 下田啓, 岡田絵理, 清水信宏, 野口雅之, 「RDBデータベースにおける信用リスクモデルの説明力の年度間推移に関する分析」, 第27回JAFEE大会予稿集(2007)

本件に関するお問い合わせは、下記までお願い致します。

日本リスク・データ・バンク株式会社
〒105-0011 東京都港区芝公園1-8-21 芝公園リッジビル9F
TEL:03-5425-2188 FAX:03-5425-2332
E-mail:jimukyoku@riskdatabank.co.jp
<http://www.riskdatabank.co.jp>

.....
● 本資料の作成につきましては弊社保有のデータベースその他信頼に足ると判断した情報に基づき正確性を
● 期しておりますが、その完全性を保証するものではありません。また内容は作成時点における筆者の個人的
● な見解を示しており、所属会社の公式見解を示すものではありません。
.....