# 格付モデルの構築と検証

株式会社金融工学研究所
開発アドバイザリー部
取締役部長　森内一朗
副部長　木村和央

2009年6月

# Contents

「信用リスクの計量化手法と信用リスクマネジメントの基本的な枠組み」の構成要素となる「格付」についてご説明。

＊ 格付モデルの概要、特徴
- 統計モデル（デフォルトアプローチ）
- 構造モデル（オプションアプローチ）

＊ モデル検証のポイント

1. 信用リスクマネジメントと格付
2. スコアリングモデルの分類
3. 統計モデル構築の流れと幾つかのモデルの紹介
4. 構造モデルによるデフォルトアラームの活用
5. 格付およびモデルの検証方法


＜弊社概要＞

株式会社金融工学研究所（**FTRI**）

株式会社格付投資情報センター（**R&I**）のグループ企業。信用リスク評価モデルの開発、信用リスクコンサルティングが中心。

主要業務内容としては、

- 統計モデルである「RADAR（信用格付け推計モデル）」、「**CrediScore**（地銀協モデル）」、「CRDスコアリングモデル」の開発。
- 株価を利用した上場企業信用リスクモデルである「DEFENSE」の開発。
- スコアリングモデル、格付精度等の検証ツール「モデルチェッカーEX」の開発。
- 日経テレコン21の企業リスク評価情報の提供「risklick」。
- **R&I**中堅企業格付けモデルの開発。
- 金融機関等へのモデル導入、内部格付体系構築および検証・内部監査支援コンサルティングサービスの提供。


＜ご注意＞

本稿の内容および意見は、発表者個人に属するものであり、発表者の属する組織の公式見解ではありません。

# 1. 信用リスクマネジメントと格付

信用リスクマネジメントのキーワードは「格付」と、それを前提とした「ポートフォリオ管理」であるといっても過言ではない。ここでは、「格付」に焦点を絞り、「格付」とはいかなる存在なのかを金融機関あるいは一般事業会社の財務担当者以外の方にもイメージがつかめるように説明したい。それを踏まえて、次章以降にて、「格付」を付与するための道具としての存在である「スコアリングモデル」について語られる。

# ①格付

- 大辞林による「格付」
  - 内容・価値・能力などによって人や物の段階・等級を決めること。
  - 債券などの元本償還や利払いの確実性の度合について序列をつけること。
    アルファベットなどの簡単な記号で表示され、投資家の判断材料とされる。
    債権格付け。レーティング。

外部格付

- 一般的なイメージは、格付会社が社債の発行体に付与する「格付」
  - 『格付けとは、発行体が負う金融債務についての総合的な債務履行能力や個々の債務の支払いの確実性（信用力）に対するR&Iの意見を、一定の符号によって投資家に情報として提供するものです。』

※R&Iのホームページより引用。

※わが国で活動している格付会社としては、R&I（格付投資情報センター）のほか、JCR（日本格付研究所）、ムーディーズ・インベスターズ・サービス・インク、スタンダード＆プアーズ、フィッチ・レーティングスといったところが有名。

| 企業名 | トヨタ | ホンダ | 日産 | マツダ | 三菱 |
|---|---|---|---|---|---|
| 格付 | AAA | AA | A- | BBB+ | BB |

主要自動車メーカーに対するR&I発行体格付けの例（2009年3月31日現在）

※一般に、BBB－以上が「投資適格」、BB+以下は「投機的」とされる。社債発行体でないが銀行融資からみると優良な先である中堅・中小企業は、仮にだが、同様の枠組みで格付を付与するとBBクラスになってしまう可能性が高い。このため、R&Iでは、「投資適格」ではく、「融資適格」を捉えるべく「R&I中堅企業格付け」と呼ばれる格付（小文字aaa－ccc）を開発した。

# ②信用力・回収可能性

- 信用力
  - 契約が約定どおり履行される／されない可能性を評価。
    「債務不履行の可能性」⇒「デフォルト確率（PD）」
    - 銀行の貸付金が長期に渡って延滞
    - 売掛先が法的破綻して回収できない
    - 社債の金利や元本が支払われない
    - 債務保証している先が法的破綻

- 回収可能性
  - デフォルト時残債の回収がなされる／なされない可能性を評価。
    「デフォルト時損失率（LGD）＝１－回収率
    ＝１－回収金額÷デフォルト時残債」
    - 担保処分により回収
    - 保証人からの回収
    - 債務者からの回収（清算、訴訟、・・・）

※PD:Probability of Default
LGD:Loss Given Default

# ③内部格付制度

- 内部格付制度とは
  - 自社が保有する個々の債務者・取引に対して、信用力・回収可能性を考慮して、独自の「格付」を付与し、管理に役立てていく仕組み。

いわば「モノサシ」の役割

信用リスク計量化
取引方針等の策定

デフォルト確率（PD）

- 信用格付と案件格付
  - 信用格付・・・個々の債務者の信用力を反映。
  - 案件格付・・・個々の取引の回収可能性を反映。

※PD:Probability of Default
LGD:Loss Given Default
EL:Expected Loss

※デフォルト率が高くても、担保等によって、回収可能性が高まっているのであれば、期待損失率は低くなり、優良な取引ということになる。

期待損失率（EL率）
＝デフォルト確率（PD）×デフォルト時損失率（LGD）

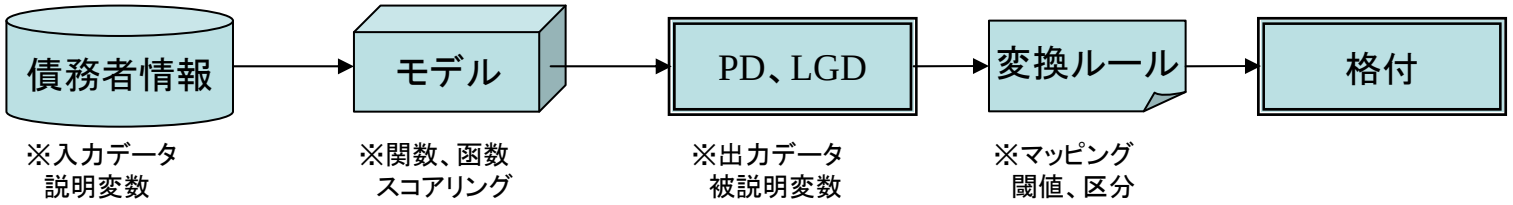
※一口にデフォルト確率といっても、「デフォルト」の定義をどうするかでイメージが異なる。金融機関が従うべきBIS規制上の定義では、後述の債務者区分が要管理以下となった場合がデフォルトとされるが、一方で、内部管理上は破綻懸念以下をデフォルトとするケースが多い。一般事業会社では実質破綻以下相当、つまり法的破綻（いわば倒産）の直前と考えるケースが多いのであろうか。なお、「デフォルト」の定義の仕方によって、デフォルト時損失率の水準は異なるので注意が必要である。

# ④評価手法

- 主観的な評価手法
  - あの会社は有名、評判もよい。
  - 社長が有名人である。
  - 保証人である社長の友人は信頼できる。
  - まじめな人だから大丈夫。
  - 大会社の社員だから安心。
  - 意外と有効な評価方法かもしれないが、このままでは客観性に欠ける。

- 客観的な評価手法
  - 統一的に把握可能な債務者の情報を利用して、デフォルト確率、デフォルト時損失率を導くためのモデルを構築する。出力数値から格付へと変換する。




  - 債務者情報としては、財務情報（決算書）、株価、社債のスプレッド等の定量情報の他、客観的な定性情報、業種別のマクロ経済変数等もその候補となりうる。

※次章で後述するモデル形態によって、求められる債務者情報が異なってくる。

# ⑤内部格付制度の設計例（信用格付）

金融機関における内部格付（信用格付）の設計例・・・債務者区分、外部格付とのリンクを含む

| 債務者区分 | | 内部格付 | 外部格付 | PD | 事象基準 | 財務基準 | 大企業 | 中小 | 個人 |
|---|---|---|---|---|---|---|---|---|---|
| 正常 | 正常先 | 1 | AA以上 | 0.02 | | | ○ | | |
| | | 2 | A | 0.08 | | | ○ | | |
| | | 3 | BBB | 0.25 | | | ○ | ○ | |
| | | 4 | | 0.75 | | | ○ | ○ | |
| | | 5 | | 1.25 | | | ○ | ○ | |
| | | 6 | | 1.75 | | | ○ | ○ | |
| | | 7 | BB | 2.50 | | | ○ | ○ | ○ |
| 要注 | 要注意先 | 8 | | 3.50 | | 赤字／債務超過 | ○ | ○ | ○ |
| | | 9 | B | 5.50 | | 決算書未提出 | ○ | ○ | ○ |
| | | 10 | | 15.00 | 1M延滞 | | ○ | ○ | ○ |
| 要管 | 要管理先 | 11 | | DF | 条件緩和 | | ○ | ○ | ○ |
| 破懸 | 破綻懸念先 | 12 | CCC以下 | DF | 3M延滞 | 赤字＆債務超過 | ○ | ○ | ○ |
| 実破 | 実質破綻先 | 13 | | DF | 6M延滞 | | ○ | ○ | ○ |
| 破綻 | 破綻先 | 14 | | DF | 法的破綻 | | ○ | ○ | ○ |

モデル結果

ノッチ調整後

※実際は、もっと複雑な債務者区分の判定（救済含む）が必要なのだが、あくまでイメージ。

※事象あるいは財務に該当した場合は、債務者区分・格付は、それ以下となり、モデル結果は関係がなくなる。
このほか金融機関独自ルールでの格付調整（ノッチアップ・ダウン）が存在する。

※BIS規制では、最低、非デフォルト7区分、デフォルト1区分の格付が必要とされている。また、LGDを考慮した案件格付については割愛する。

# ⑥内部格付の活用と留意点

## ■営業・審査部署での活用

- 格付別与信決裁権限の設定（上位格付先は下位権限者に決裁権限を委譲）
- 格付別取引方針の策定（格付をベースに、取引推進・維持・縮小を決定）
- 格付に応じた与信先の期中管理（問題先、重点的に管理すべき先の選定）
- いわゆるビジネスローンへの活用（無担保、第三者保証不要の小口融資）

## ■リスク管理部署での活用

- 格付別基準金利の設定
  基準金利＝調達金利＋経費率＋信用コスト＋資本コスト＋利ざや
  信用コスト＝PD×LGD、資本コストは後述の計量化によって計算
- ポートフォリオの信用リスク計量化

※以上、新見氏（日銀金融高度化センター）の2007年3月講演資料を参考に加筆修正。

## ■内部格付活用上の留意点

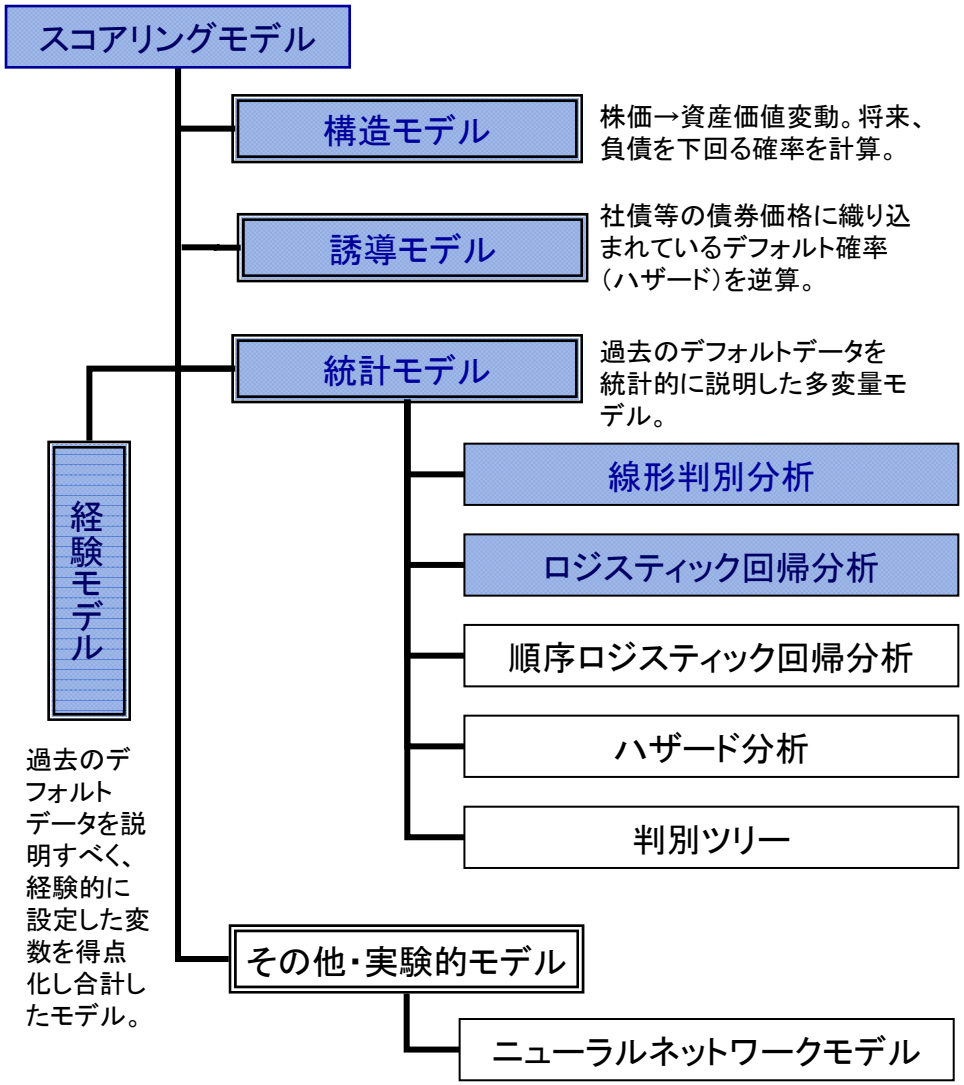
- 格付精度の構築自体が目的化してしまう、あるいは格付にしたがっていればデフォルト時の責任は問われないといった風潮にならないよう留意。
- 現場の生の主観情報のなかにも、重要なシグナルがあることを理解し、バランスのよい自社に適した審査・格付運営をたえず考えていくことが重要。

# 2.スコアリングモデルの分類

デフォルト確率を出力するためのモデルとして、決算書データを説明変数とした統計モデル、株価データから資産価値モデルを構成した構造モデル、社債のスプレッドに内在するデフォルト確率を割り出すための誘導モデルの3タイプに分類される。本章では、これらの基礎となる考え方を説明し、それぞれの特徴、利用上の留意点等について述べる。

# ①スコアリングモデルの系統と本稿の守備範囲




- 線形判別分析
  - 古典的分析手法で、デフォルトと非デフォルトの境界線を探すイメージの分析。
  - 説明しやすいが、あてはまりはあまりよいとは言えない。
- ロジスティック回帰分析
  - デフォルトと非デフォルトの間のまさに確率を表現させるイメージの分析。
  - モデル構造が線形回帰分析と類似であり、予測結果も概ね良好。
- ハザード分析
  - デフォルト率の期間構造を表現させるイメージの分析。
  - 長期間のデータが必要で、景気変動への対応が必要。
- 判別ツリー
  - 上記のモデルは、通常、説明変数とデフォルト確率の間に単調性を仮定しているが、そこに交差効果を入れて、効果的な多段クロス集計を作るイメージの分析。
  - 多数のデータが揃う個人のリテール分野では一般的だが、事業法人ではあまり見ない。
- ニューラルネットワークモデル
  - とにかく当てるためのモデル。
  - なぜ。当たったかは問わないため、説明のしにくさが難点。

※山下先生（統計数理研究所）講演資料を参考に加筆修正。

# ②経験モデル

## ■ 経験モデルの例

- 入力データは、決算書データから加工された財務指標のほか、客観的な定性指標も利用可能。
- 右図のような経験に基づいての指標選択と配点表を準備。入力データに基づいた個々の指標の得点を合計して、その得点により格付を付与。

| 観点 | 指標 | カテゴリ区分（上段）と得点（下段） | | | |
|---|---|---|---|---|---|
| 成長性 | $x_1$ | $x_1 < y_{11}$<br>$a_{11}$ | $y_{11} \leqq x_1 < y_{12}$<br>$a_{12}$ | $y_{12} \leqq x_1 < y_{13}$<br>$a_{13}$ | $y_{13} \leqq x_1$<br>$a_{14}$ |
| 増収率など | $x_2$ | $x_2 < y_{21}$<br>$a_{21}$ | $y_{21} \leqq x_2 < y_{22}$<br>$a_{22}$ | $y_{22} \leqq x_2 < y_{23}$<br>$a_{23}$ | $y_{23} \leqq x_2$<br>$a_{24}$ |
| 収益性 | $x_3$ | $x_3 < y_{31}$<br>$a_{31}$ | $y_{31} \leqq x_3 < y_{32}$<br>$a_{32}$ | $y_{32} \leqq x_3 < y_{33}$<br>$a_{33}$ | $y_{13} \leqq x_3$<br>$a_{34}$ |
| 利益率など | $x_4$ | $x_4 < y_{41}$<br>$a_{41}$ | $y_{41} \leqq x_4 < y_{42}$<br>$a_{42}$ | $y_{42} \leqq x_4 < y_{43}$<br>$a_{43}$ | $y_{43} \leqq x_4$<br>$a_{44}$ |
| 効率性 | $x_5$ | $x_5 < y_{51}$<br>$a_{51}$ | $y_{51} \leqq x_5 < y_{52}$<br>$a_{52}$ | $y_{52} \leqq x_5 < y_{53}$<br>$a_{53}$ | $y_{53} \leqq x_5$<br>$a_{54}$ |
| 回転率など | $x_6$ | $x_6 < y_{61}$<br>$a_{61}$ | $y_{61} \leqq x_6 < y_{62}$<br>$a_{62}$ | $y_{62} \leqq x_6 < y_{63}$<br>$a_{63}$ | $y_{63} \leqq x_6$<br>$a_{64}$ |
| 安全性 | $x_7$ | $x_7 < y_{71}$<br>$a_{71}$ | $y_{71} \leqq x_7 < y_{72}$<br>$a_{72}$ | $y_{72} \leqq x_7 < y_{73}$<br>$a_{73}$ | $y_{73} \leqq x_7$<br>$a_{74}$ |
| 自己資本比率など | $x_8$ | $x_8 < y_{81}$<br>$a_{81}$ | $y_{81} \leqq x_8 < y_{82}$<br>$a_{82}$ | $y_{82} \leqq x_8 < y_{83}$<br>$a_{83}$ | $y_{83} \leqq x_8$<br>$a_{84}$ |

合計得点：$a = a_{1.} + a_{2.} + a_{3.} + a_{4.} + a_{5.} + a_{6.} + a_{7.} + a_8$

↓

格付 → 過去の格付別実績デフォルト率＝デフォルト確率

## ■ 経験モデルの利点

- 審査現場との親和性
  →審査担当者の経験を踏まえ構築。
- 理解のしやすさ
  →簡易な計算式、審査目線と共通。

## ■ 経験モデルの欠点

- 指標選択基準、配点表の根拠不足。
- 精度が統計モデルと比較して劣後。

※近年は、統計モデルに基づいて指標を選択し、得られた係数をもとに配点表を構成するといった工夫もみられる。また、遺伝的アルゴリズム（GA）を適用し、最適な配点表の閾値と得点を求めるといった先行研究もある（MTEC、東工大シンポジウム 2004）。

# ③統計モデル（1）概要

## ■経験モデルとの相違点

- 経験モデル→経験に基づく指標選択、配点表の決定。
- 統計モデル→統計学に基づく最適な指標選択、指標ウエイトの決定。

※ただし、判別ツリーは統計学ではなくて、大量データに基づき行われるデータマイニングだとする意見もある。ちなみに、データマイニングおよびデータベースからの知識発見のことをKDD（Knowledge Discovery and Data Mining）と称するが、それとの対比で「勘と経験と度胸」のことをKKDなどと言うことがある。混沌とした将来を見通す上では、過去の情報をベースとしたKDDだけでなく、KKDも重要なことは同意する。

## ■統計モデル発展の経緯

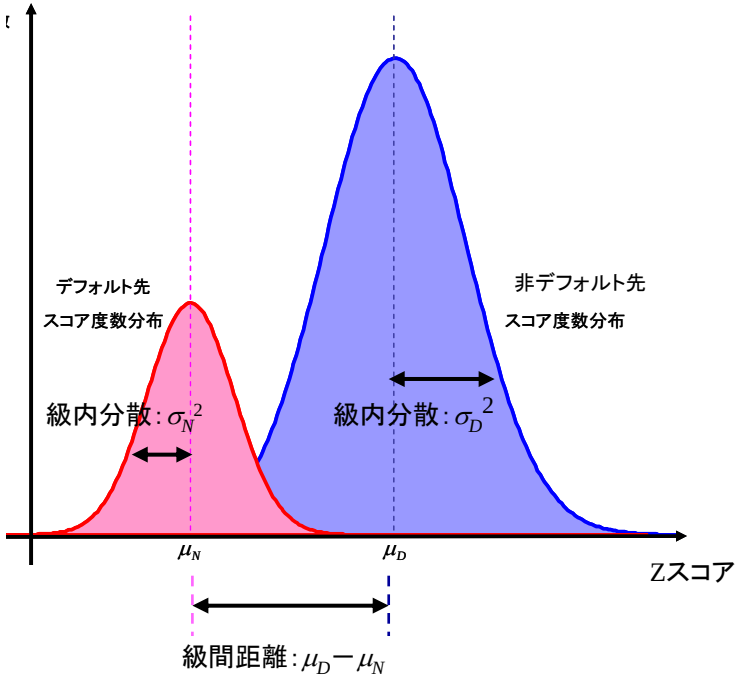
- 1960年代　　判別分析　→　AltmanによるZスコアが有名
- 1970年代　　ロジスティック回帰分析　→　標準的に利用されるモデル
- 1980年代以降　ハザード分析、判別ツリー、（ニューラルネットワーク）

※判別分析、ロジスティック回帰分析、ハザード分析は、生存・死亡の判別、死亡率の推定、時系列の生存率曲線の推定といった具合に、医薬統計の分野で発展を遂げてきたものである。金融リスク管理の分野は、その成果を追随して取込んでいるため、直感的には上記の年代より10年から20年程度遅れてブームが到来しているような感じがする。医薬統計と異なり、説明変数候補が多数あるため、最適解を探索するにも時間がかかり、実務で利用可能なモデルを構築する上では、パソコンCPUの高速化が必要であったということも理由ではないかとも思われる。

※判別ツリー、ニューラルネットワークによるモデル構築のためには、大量のデータが必要である。よって、企業（事業法人、個人事業主）を対象としたスコアリングモデル（デフォルト判別モデル）では用いられることは少ない。しかし、個人を対象とした消費者ローン、住宅ローンの審査分野では、変数の交差効果（非線形効果）を表現できるこれらのモデルが活用されている。

# ③統計モデル（2）判別分析

■ 判別分析のイメージ




$$Z = \beta_1 x_1 + \beta_2 x_2 + \Lambda + \beta_m x_m$$

– $x$は選択された指標。級間距離ができるだけ大きく、級内分散ができるだけ小さくなるように$\beta$を決定。

■ AltmanのZスコアモデル

– 1940～60年代の米国の製造業を対象。

$$Z = 1.2x_1 + 1.4x_2 + 3.3x_3 + 0.6x_4 + 1.0x_5$$

- $x_1$=運転資本　／総資産
- $x_2$=利益剰余金／総資産
- $x_3$=営業利益　／総資産
- $x_4$=時価総額　／負債簿価
- $x_5$=売上高　　／総資産
- 運転資本＝受取手形＋売掛金＋棚卸資産－支払手形－買掛金

– Z≦2.675だと倒産可能性大と判定。

– このモデルは時価総額が必要であるため、株価がないと計算ができないという欠点あり。

※日本企業に適用した例が、時折、週刊誌等で見かける。簡単なモデルながら、選択されている指標はごく一般的であるため、倒産の閾値は別として、順位づけとしては機能しているのではなかろうか。

※ところで、Altman氏であるが、現在もニューヨーク大学の教授である。2008年5月に来日され、当社とR&Iの共催特別セミナーで、ご講演いただいた。近年は回収率にもご興味をもたれて研究なさっている。

# ③統計モデル（3）ロジスティック回帰分析の必要性

■ 1つの指標による数値例にて説明

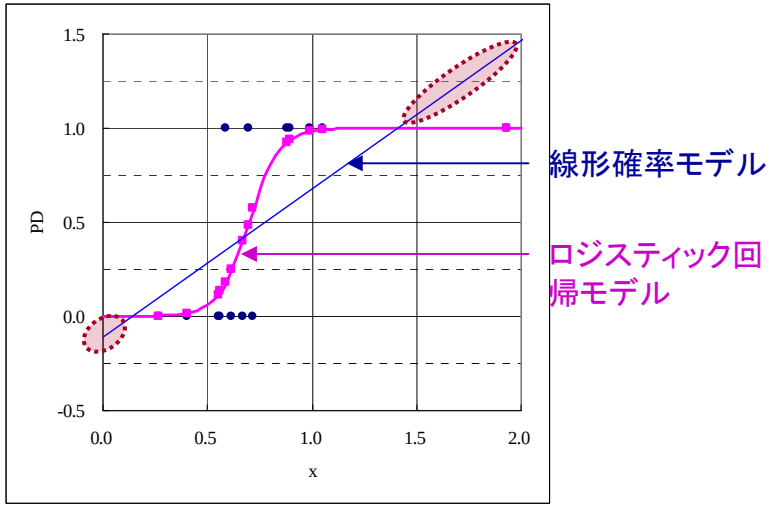
| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\delta_i$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $x_{1i}$ | 1.932 | 0.993 | 1.053 | 0.881 | 0.896 | 0.693 | 0.588 | 0.403 | 0.617 | 0.563 | 0.264 | 0.550 | 0.720 | 0.670 |

– $\delta$=1：デフォルト、0：非デフォルト、$x$：指標値

■ 線形確率モデルとその問題点

– $\delta$を被説明変数、$x$を説明変数として回帰分析。

$$PD_i = \beta_0 + \beta_1 x_{1i} = -0.108 + 0.787 x_{1i}$$




– PDは確率なのに、0から1の範囲に収まらない。

– $x_1$の増分とPDの増分の関係が$\beta_1$倍で一定。

■ ロジスティック回帰モデルの導入

– 直線ではなく、ロジスティック曲線のあてはめ。

$$PD_i = \frac{1}{1+e^{-Z_i}}, \ln\left(\frac{PD_i}{1-PD_i}\right) = Z_i = \beta_0 + \beta_1 x_{1i}$$

– 最尤法を用いて$\beta$を推計。対数尤度$\ln L$の最大化。

$$\ln L = \sum_{i=1}^{N} \left[\delta_i \ln PD_i + (1-\delta_i)\ln(1-PD_i)\right] \to \max$$

※この式が意味するところは、デフォルト先はPD、非デフォルト先は1-PDを出し合って合計し、それを最大化するように$\beta$を決めなさいということである。

$$Z_i = -9.643 + 13.818 x_{1i}$$

※判別分析のアナロジーでZと表記され、Zスコアと呼ばれてしまうことも多い。潜在変数と呼んだりする人もいる。本稿における定義では、判別分析とは異なり、Zが大なるほうがPDが高いことに注意が必要である。このため、最初からZではなく、-Zにて定義される場合もある。

# ③統計モデル（4）Excelによるロジスティック回帰分析

Microsoft Excel - Book1

ファイル(F) 編集(E) 表示(V) 挿入(I) 書式(O) ツール(T)
データ(D) ウィンドウ(W) ヘルプ(H) Adobe PDF(B)

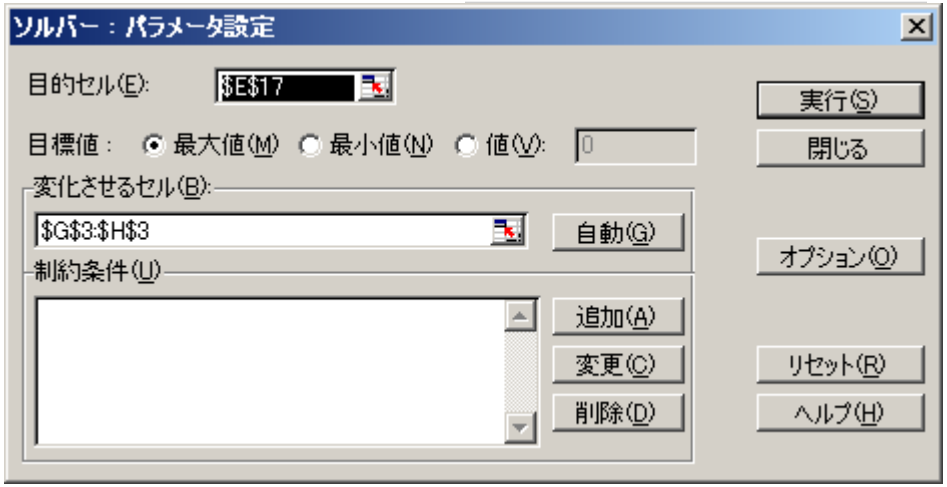
J14

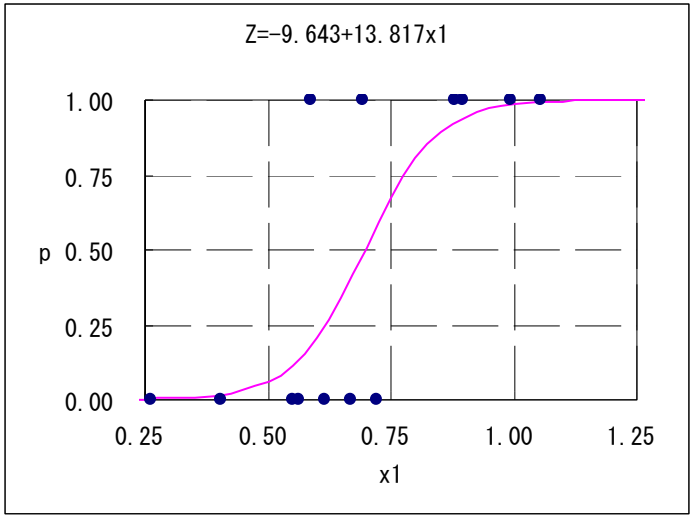
| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | i | δ i | x1 i | pi | logL | | β 0 | β 1 |
| 3 | 1 | 1 | 1.932 | 1.000 | 0.000 | | −9.643 | 13.817 |
| 4 | 2 | 1 | 0.993 | 0.983 | −0.017 | | | |
| 5 | 3 | 1 | 1.053 | 0.993 | −0.007 | | | |
| 6 | 4 | 1 | 0.881 | 0.926 | −0.077 | | | |
| 7 | 5 | 1 | 0.896 | 0.939 | −0.063 | | | |
| 8 | 6 | 1 | 0.693 | 0.483 | −0.727 | | | |
| 9 | 7 | 1 | 0.588 | 0.180 | −1.716 | | | |
| 10 | 8 | 0 | 0.403 | 0.017 | −0.017 | | | |
| 11 | 9 | 0 | 0.617 | 0.247 | −0.283 | | | |
| 12 | 10 | 0 | 0.563 | 0.134 | −0.144 | | | |
| 13 | 11 | 0 | 0.264 | 0.002 | −0.002 | | | |
| 14 | 12 | 0 | 0.550 | 0.115 | −0.122 | | | |
| 15 | 13 | 0 | 0.720 | 0.576 | −0.858 | | | |
| 16 | 14 | 0 | 0.670 | 0.405 | −0.519 | | | |
| 17 | | | | | −4.552 | | | |

例

コマンド NUM

※G3とH3セルの初期値は適当に、0としておけばよい。これは、すべての先のPDを50%としたことに相当する。右図のようにソルバーを設定して、実行ボタンを押せば、自動的に計算がなされ解が得られる。



# ③統計モデル（５）ロジスティック回帰分析の流行理由

- 判別分析同様、Zスコアは個々の指標の一次結合（係数×説明変数の和）で表すことができ、透明性が高く、対外的な説明がしやすい。
  - 指標とZスコア、PDとの関係が明確である。
  - 新BIS規制の内部格付手法のPDモデルとして、多数の金融機関が採用し、モデルベンダーも推奨
- PDを0から1の間で評価できる。
  - 確率であるから、区間[0，1]に収まることが自然
  - 線形確率モデルでは、区間[0，1]の外になる場合の対応が必要
- 個々の指標の分布に対して仮定が緩い。
  - 線形判別分析では、デフォルト・非デフォルトサンプル群ごとに、説明変数の等分散性（分散・共分散行列が等しいこと）を仮定
- 質的な変化を意味するダミー変数を利用することができる。
  - たとえば、赤字ダミー（赤字のとき1、そうでないとき0なる変数）も投入可能
- 2値から多値の順序ロジスティック回帰分析への拡張が可能である。
  - 非デフォ・デフォの2状態から、3状態以上の確率を求めるモデルへ発展可能

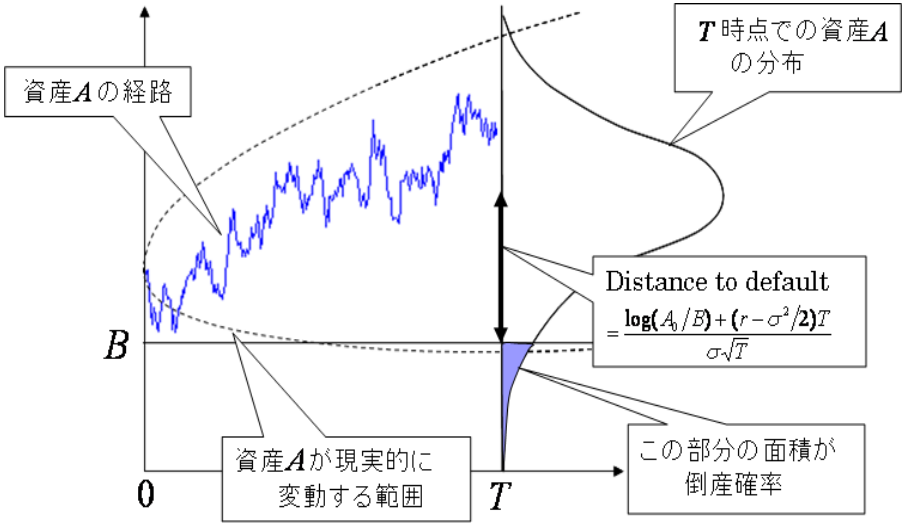
 

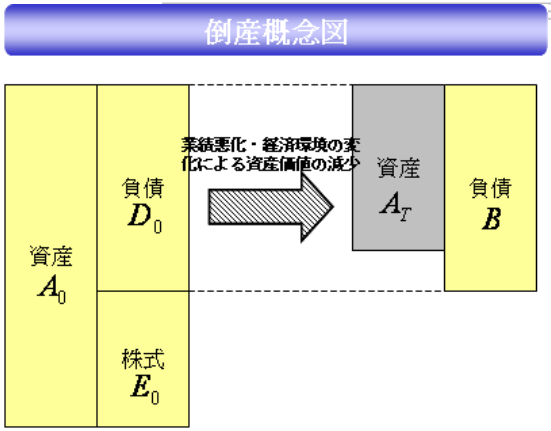
# ④構造モデル（1）基本概念

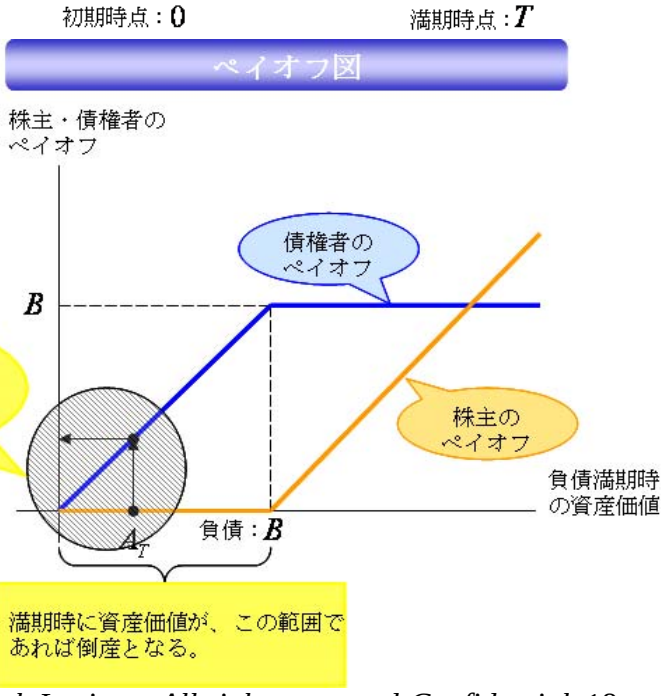
■ Mertonモデル

– 満期時点$T$において、企業の資産価値が負債価値を下回ったらデフォルト。
債務超過状態⇒これを「倒産」と呼称。

– 資産価値の変動を株式の時価総額の変動で代用してモデル化。




※資産価値が負債価値を1円でも下回ればデフォルト（倒産）であり、デフォルト確率を求めるには、その範囲に陥る確率を計算すればよい。一方、資産価値の値そのものも得られるわけであるから、負債価値と資産価値の比率からデフォルト時損失率を求めることも可能である。

# ④構造モデル（2）利点と欠点

## ■ 構造モデルの利点

- 株価などのマーケットデータを利用するため、公開企業であれば評価可能。
- 将来の企業業績への期待が反映されるならば、予見性のある評価が可能。
- 統計モデルと比較して、評価精度が時間の経過とともに低下しにくい。
- 決算書データを利用した統計モデルに比べ、タイムリーな評価が可能。

## ■ 構造モデルの欠点

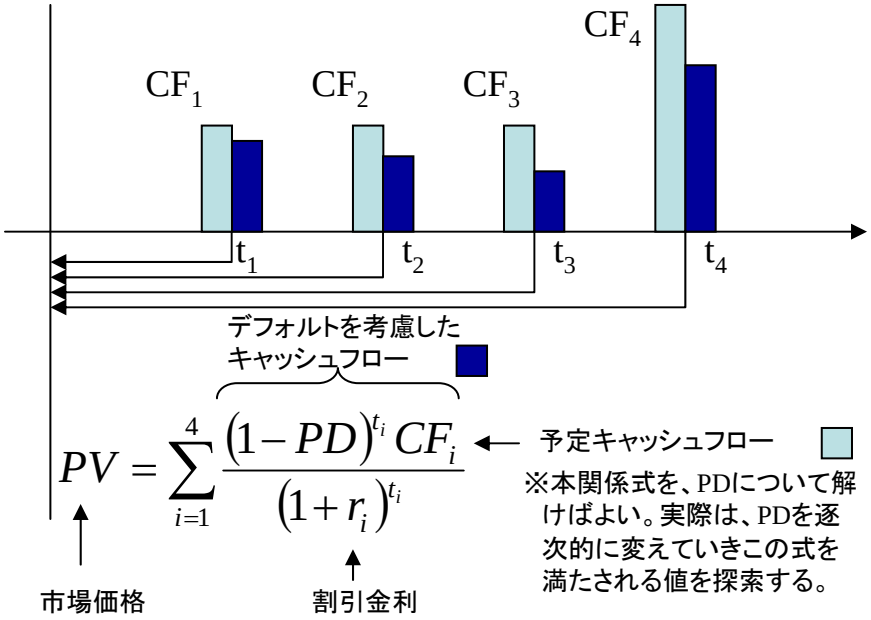
- そもそも株価の変動は、真の企業資産価値の変動でない要素もあるはず。
  ノイズトレーダーなど株価が過剰反応している場合の評価結果は？
- 債務超過となってもデフォルトするとは限らない。
  メインバンクの支援などの考慮が必要か？
- 当然だが、未公開企業はマーケットで評価された株価はなく対象外。

格付用モデルではなく、上場企業のモニタリングツールとしての利用が一般的

# ⑤誘導モデル

■ 基本概念
- 社債のスプレッドは信用リスクを反映している。
- このことを前提として、デフォルト確率を誘導する。

■ イメージ
- PD（年率表示）が一定なら逆算可能。




$$PV = \sum_{i=1}^{4} \frac{(1-PD)^{t_i} CF_i}{(1+r_i)^{t_i}}$$

■ 利点
- 社債の価格データを用いるため、株価を利用した構造モデル同様に、タイムリーな評価が可能。

■ 欠点
- そもそも社債が取引なされていなければ計算不可能。
- マーケット参加者の予想を反映しているだけと言えなくもない。

学術研究の世界では、論文も多く、広く扱われている課題である。

※実際の誘導モデルは、Duffie-Singletonなどが有名だが、連続変数により、積分の形で書かれるため、何とも理解しにくいが、エッセンスは左図のようなものと考えてよい。なお、社債のスプレッドには、デフォルト確率のほか回収率も考慮されていると考えるべきである。ただし、構造モデルとは異なり、一方を別の方法で推定しておいて、もう一方を誘導モデルで定めるという手続きになることに注意が必要である。

# 3.統計モデル構築の流れと幾つかのモデルの紹介

前半は、与信先を保有する企業が、各種データを収集し、自企業内で独自にモデルを構築する場合の手法について説明し、独自にモデルを作らない場合でも、基礎知識として習得すべきポイントを確認する。

後半は、外部モデルの活用方法と、幾つかのモデルについて、概要および特徴を紹介する。

# ①統計モデル構築の流れ

1.分析用データ収集・抽出

モデル構築に必要なデータを収集し、分析用データセットとして格納。
紙の決算書や申込書を数千先、手入力していた時代もあった。
※過去のセミナー資料では、デフォルト先は全件、非デフォルト先は1万件を目標に入力を促していた。

2.データクレンジング

申込書や手入力データには、誤入力や欠損値も存在。
同一カラム(列)に、異なる情報が入っているケースも。
※意図的なW-meaningのケース、ある時期を境にカラム定義を変更してしまったケース等がある。

3.説明変数の作成・加工

財務指標を計算し、必要な欠損値処理、異常値処理、変数変換処理を実施。
モデル投入変数を、離散型とするか連続型にするかで処理内容は変化。
※このあたりが、各コンサルティング会社の商売の源泉。本稿では、公開資料ベースで説明する。

4.単変数ごとの基礎分析

単変数回帰、分割表分析等を実施し、各変数の振る舞いをチェック。
多重共線性問題の考察のために、変数間相関のチェックもしておきたい。
※実は、次のステップで、相関が強い変数は2つと選ばれないようになるので、さほど気にしない。

5.モデル採用変数選択

符号条件に注意しながら、ステップワイズ(変数増減法)で変数を選択。
変数選択順序に結果が依存するため、さまざまな方法で候補モデルを追加。
※医薬統計では、総当り法を試すようだが、サンプル数と変数の数が多い金融の世界では厳しい。
最適なものを探索する時間よりも、まぁまぁのものを使えるように仕上げる時間が大事と心得よ。

6.パフォーマンス検証

モデルのロバストネス(頑健性)を検証用データにてチェック。
モデル構築用データと検証用データに分割しておくという方法が一般的。
※各種検証方法は、後述とする。

※勿論、結果によっては前のステップに戻り、再度処理することがある。

# ②モデル構築の原点回帰

■ モデル構築の3大要素

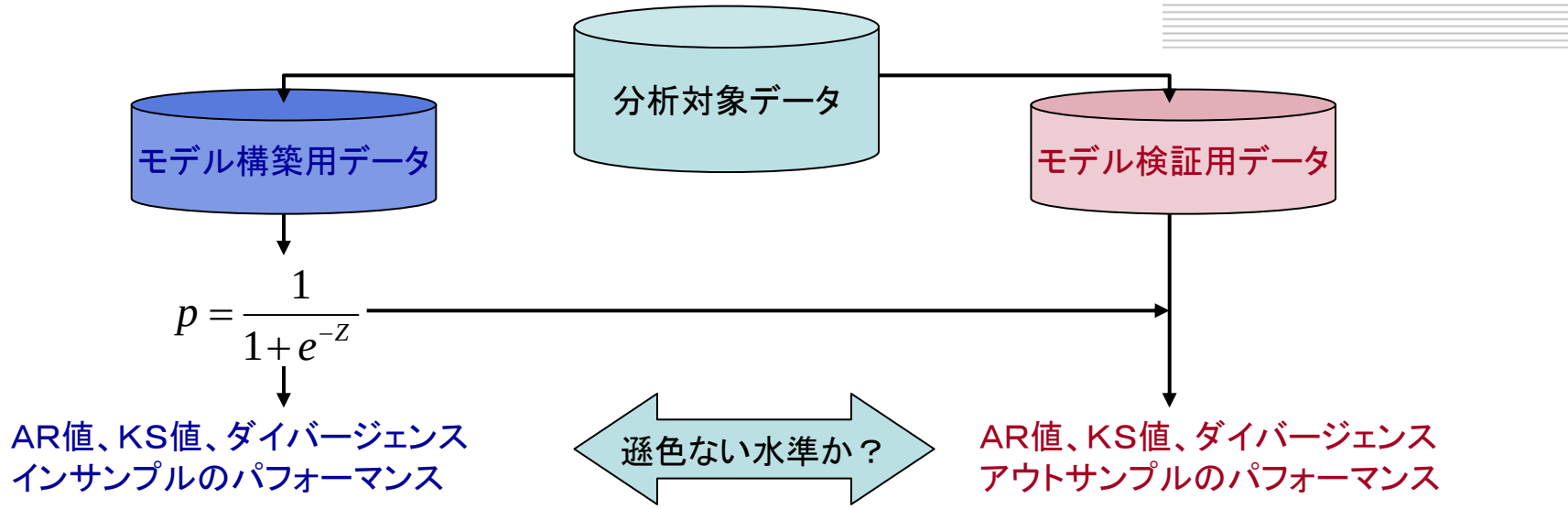

■ モデリング手法は、結果に与える影響は限定的。
- 手法を変えても同じデータを使っている以上、同じような結果が出るだけ。
- 重要なのはスピード感。構築から適用まで急がないとモデルは陳腐化。

※といっても、構築にはある程度時間がかかるのであり、システムセットのスケジュールから急がされるといいことはない。
作ったモデルをそのままセットできないシステムを使っている限り、これでは未来永劫改善しないであろう。

■ 土台となる基礎データが成功の鍵。
- 過去から一定の精度でデータ収集しているか？
- その他、有用なデータを蓄積する努力をしているか？

■ モデル構築時に、業務知識・ノウハウを反映することが競争力の源泉。
- 暗黙知をモデル構築を通じて形式知化させることが、真の差別化要因。
- PDCAプロセスを通じて、見直しを行うことで、ノウハウを拡大再生産。

# ③パフォーマンス検証（1）クロスバリデーション概略




※事実上AR値がデファクトスタンダードである。KS値もダイバージェンスもAR値と大差ない統計量である。

- 検証用データにモデルを適用
  - ロバストネス（頑健性）チェック。検定統計量が検証用で急落してないか？
  - もし、急落しているとすると、モデル構築時に実施した分割表と同様のものを作成することで、問題点は判明する場合が多い（検証用データでは、説明力がない指標、単調性がない指標が存在している可能性が高い）。

- ブートストラップによる検証
  - 検証サンプルが、この1つのデータでは心もとないということでは、モデル検証用データからブートストラップで複数回、異なるサンプルセットを作成し、その統計量のぶれをみることで、検証精度を高めることも行われている。

※このほか、過去の時点を基準日としたバックテスト、そのブートストラップなど、検証メニューは幾つも考えられる。

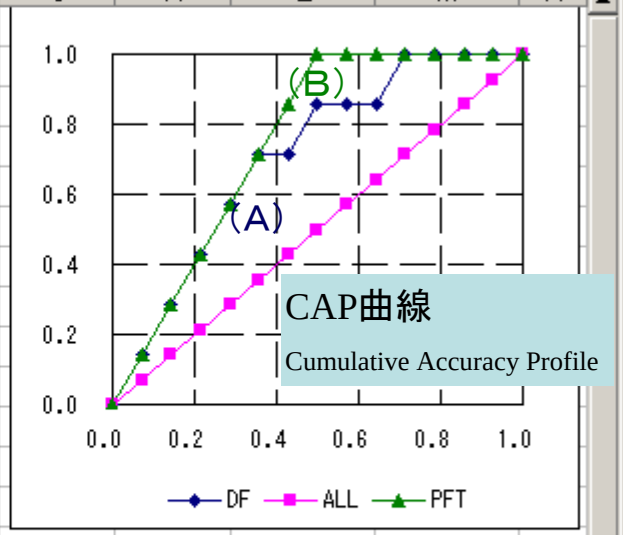
# ③パフォーマンス検証（2）CAP曲線とAR値

Microsoft Excel - Book1

ファイル(F) 編集(E) 表示(V) 挿入(I) 書式(O) ツール(T) データ(D) ウィンドウ(W) ヘルプ(H) Adobe PDF(B)

H3 =(F3-E3)*(E3-E2)

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | ALL | DF | PFT | S-DF | S-PFT |
| 2 | i | δ i | x1 i | pi | 0.000 | 0.000 | 0.000 | | |
| 3 | 1 | 1 | 1.932 | 1.000 | 0.071 | 0.143 | 0.143 | 0.005 | 0.005 |
| 4 | 3 | 1 | 1.053 | 0.993 | 0.143 | 0.286 | 0.286 | 0.010 | 0.010 |
| 5 | 2 | 1 | 0.993 | 0.983 | 0.214 | 0.429 | 0.429 | 0.015 | 0.015 |
| 6 | 5 | 1 | 0.896 | 0.939 | 0.286 | 0.571 | 0.571 | 0.020 | 0.020 |
| 7 | 4 | 1 | 0.881 | 0.926 | 0.357 | 0.714 | 0.714 | 0.026 | 0.026 |
| 8 | 13 | 0 | 0.720 | 0.576 | 0.429 | 0.714 | 0.857 | 0.020 | 0.031 |
| 9 | 6 | 1 | 0.693 | 0.483 | 0.500 | 0.857 | 1.000 | 0.026 | 0.036 |
| 10 | 14 | 0 | 0.670 | 0.405 | 0.571 | 0.857 | 1.000 | 0.020 | 0.031 |
| 11 | 9 | 0 | 0.617 | 0.247 | 0.643 | 0.857 | 1.000 | 0.015 | 0.026 |
| 12 | 7 | 1 | 0.588 | 0.180 | 0.714 | 1.000 | 1.000 | 0.020 | 0.020 |
| 13 | 10 | 0 | 0.563 | 0.134 | 0.786 | 1.000 | 1.000 | 0.015 | 0.015 |
| 14 | 12 | 0 | 0.550 | 0.115 | 0.857 | 1.000 | 1.000 | 0.010 | 0.010 |
| 15 | 8 | 0 | 0.403 | 0.017 | 0.929 | 1.000 | 1.000 | 0.005 | 0.005 |
| 16 | 11 | 0 | 0.264 | 0.002 | 1.000 | 1.000 | 1.000 | 0.000 | 0.000 |
| 17 | DF | 7 | | | | | | 0.209 | 0.250 |
| 18 | AL | 14 | | | | | | AR | 0.837 |




CAP曲線

Cumulative Accuracy Profile

AR値=A/(A+B)

Accuracy Ratio

PDの降順にソート

*x*座標：全サンプルの累積割合

*y*座標：DF ＝デフォルトサンプルの累積割合
ALL＝全サンプルの累積割合
PFT ＝完全（パーフェクト）モデルの累積割合

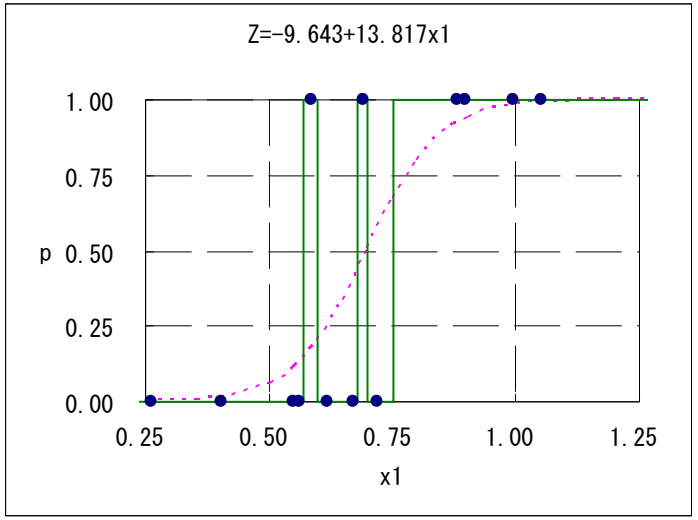
※左図の方法でAR値が計算可能なのは、同一のPDをもつデフォルト先と非デフォルト先が存在しないことが条件である。このような場合にデフォルト先を非デフォルト先より上にすると、下にしたときよりもAR値は上昇してしまう。こだわらないのであれば、保守的に、非デフォルト先が上にくるようにしておけばよい。正確に計算するには、同一のPD先をひとまとめとして、それぞれの累積割合を計算することになるため、シートに工夫が必要である。

# ③パフォーマンス検証（3）オーバーフィッティング・ロバストネス

- オーバーフィッティング（過剰最適化、以下OF）とは？
  - サンプルのデフォルト・非デフォルト状態をかなりよく説明できる状態。
  - 「尤度関数が1に近い＝対数尤度関数が0に近い」状態。

- 実際に、何が起きているのか？
  - 極端に言えば、右図のような状態。
    ※$x_1$を変数変換すれば、ロジスティック回帰モデルでも、このような状態は作り出すことができる。実は、変数変換には細心の注意が必要。
  - このモデルは、実態を表現しているのか？
    →おそらくNo！
    ※何故、Noなのかは統計学からは導くことはできない。
    $x_1$とPDは単調な関係にあるはずという感覚の判断が必要である。

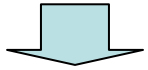



- OFの判断→ロバストネス（頑健性）チェック
  - モデル構築用データのほか検証用データを残す。
  - 検証用データでAR値を計算すれば、大幅低下。
    - たとえば、$x_1$=0.58の企業の多くはデフォルトしない。
    - 一方、$x_1$=0.74の企業の多くはデフォルトするであろう。

$$\begin{cases} p=0 \quad \Lambda \quad x_1 < 0.57 \\ p=1 \quad \Lambda \quad 0.57 \le x_1 < 0.60 \\ p=0 \quad \Lambda \quad 0.60 \le x_1 < 0.68 \\ p=1 \quad \Lambda \quad 0.68 \le x_1 < 0.70 \\ p=0 \quad \Lambda \quad 0.70 \le x_1 < 0.75 \\ p=1 \quad \Lambda \quad 0.75 \le x_1 \end{cases}$$

# ④外部モデル（1）必要性と留意点

■ 必要性
- 保有するデータ数が、内部モデル構築においては不足と考えられる場合。
- そもそもデータ自体が存在せず、これから収集していこうという場合。
- 社内に内部モデルを構築し、メンテナンスできる人材が欠如する場合。



外部モデルを可能な限り自社顧客の特徴に合わせて調整し利用。
モデルに組み込みにくい情報は、ノッチ調整時に対応していく。

■ 留意点
- 外部モデルが前提としているデータと構築の概略は情報提供を受けるべき。
  ブラックボックス化は、可能な限り排除。

  ※外部モデル提供者の立場からすると、どんなに厳格な守秘義務契約を結んでいても、すべてのノウハウを開示してしまうことには抵抗がないわけではなかろう。ビジネスモデル特許で縛ることは難しく、むしろ信頼関係をいかに築いていくかが重要と考えている。
- 少量であっても自社データがあるのであれば、外部モデルを適用し、結果を吟味しておく必要あり。

  ※たとえば、出力されたデフォルト確率の序列が正しくても、水準がまったく異なるといったことがある。この場合は、後述のPD調整を実施するのが一案である。序列も正しくないという場合は、説明変数のレベルで何か効きの悪いものがある可能性がある。

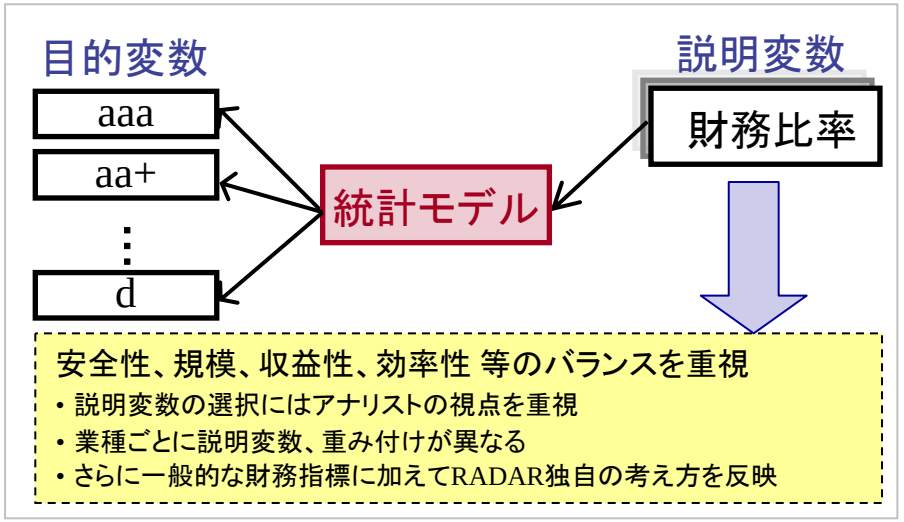
# ⑤RADAR （弊社商品を外部モデルの一例として）

■ 概要

– R&Iの発行体格付けを財務データ（過去6期分あることが前提だが、1期でも可）から推計するモデル。

– 上場企業が対象だが、財務データがあれば、それと同等の企業も評価可能。

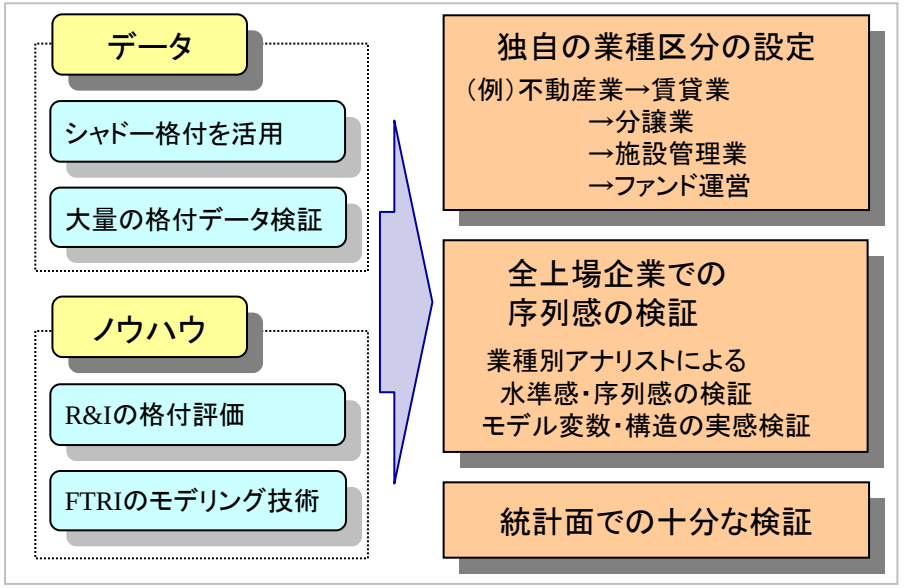
※地方自治体については、RADAR Publicというモデルがあり、この場合、格付のみならず、財務スコアも推計される。




■ モデル構築

– 業種別に財務データから算出される3～11程度の説明変数（財務比率）を用いた「統計モデル」にて格付を推計。

– データと実務ノウハウとの融合に主眼をおいたモデル構築。



RADARは、銀行、証券会社、系統金融機関、保険会社など大手金融機関を中心に多くのユーザー様に利用いただいており、事業法人への導入実績もあります。

## 財務情報から格付けを推計

RADARは、金融工学研究所とみずほフィナンシャルグループが共同開発した、財務情報からR&Iの発行体格付けを推計する定量モデルです。R&I格付けとの高い整合性があります。

## 広範囲の業種をカバー

12業種に対応した一般事業法人モデルと、5つの金融法人モデルにより、広い業種をカバーし、一般的に定量評価が難しいとされる金融機関、ノンバンク、不動産、電力、鉄道などの業態も格付けが可能です。

## 年1回以上のモデルの検証・メンテナンス

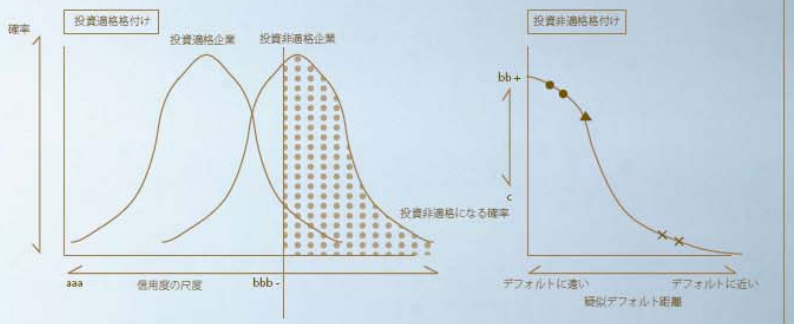
毎年最新のデータを用いてモデルのパフォーマンス検証や、水準調整、変数の重み付けの変更や必要に応じて変数の入れ替え、会計基準の変更への対応等のメンテナンスを行っています。

## 付加情報の提供

毎年、『技術資料』として、直近の格付け別デフォルト率等を提供しています。

### 理論的背景

業種モデルにより財務データから算出される3～11程度の説明変数を用いて、多変量回帰モデルによって格付けを推計しています。モデルの説明変数の選択には、デフォルトや格付けの説明力の視点に加えて、アナリストの視点を重視しています。




## 公開企業モデルと非公開企業モデル

連結決算を基準とした公開企業モデルであるRADAR1と、単体決算を基準とした非公開(中小)企業モデルのRADAR2があります。

## 開示情報

RADARはブラックボックスのツールではありません。説明変数の算出式、格付け推計のロジック等各種の技術情報を開示することが可能です。

## RADARで実現する→

### → 内部格付け制度の定量評価モデル

高い精度で信用格付けを推計できるRADARは多数の企業の信用リスクを瞬時に把握できるツールです。定量評価モデルとして内部格付けのプロセスに組みこみ、効率的に信用リスク管理を行うことができます。

### → ベンチマークに

RADARを内部格付けのベンチマークとして活用することにより、多面的な信用リスクの管理が可能となります。公開企業や金融機関など、現行の内部格付け制度を部分的に補完するような利用方法も有効です。

### → 全社のリスク管理に

RADARは金融機関や地公体等を含む幅広い業種に対して格付けを付与することが可能です。取引先全体をカバーするきめ細やかな信用リスク管理が可能となります。

### → 関係会社の評価

財務情報により信用格付けを推計するため、財務構成の変化による格付けのシミュレーションが可能です。これを利用して業態の異なる関係会社の管理に信用格付けを利用することも可能です。

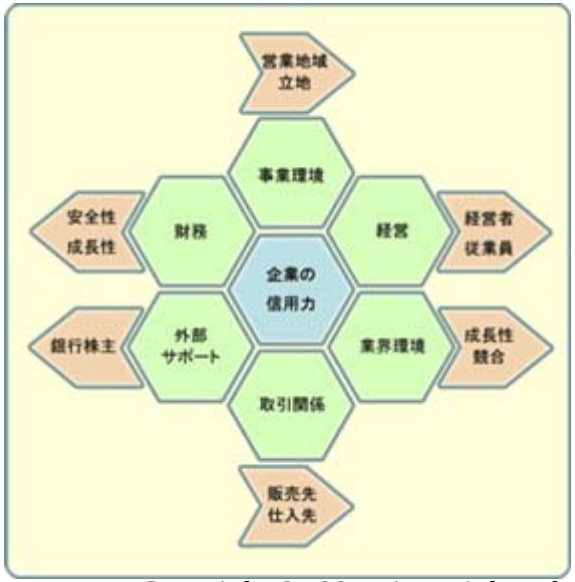
### → コンサルティングサービスとの連携

金融工学研究所では主に金融機関向けにリスク管理にかかわるコンサルティングサービスを提供しています。ご要望に応じて、各種サービスをご用意しています。

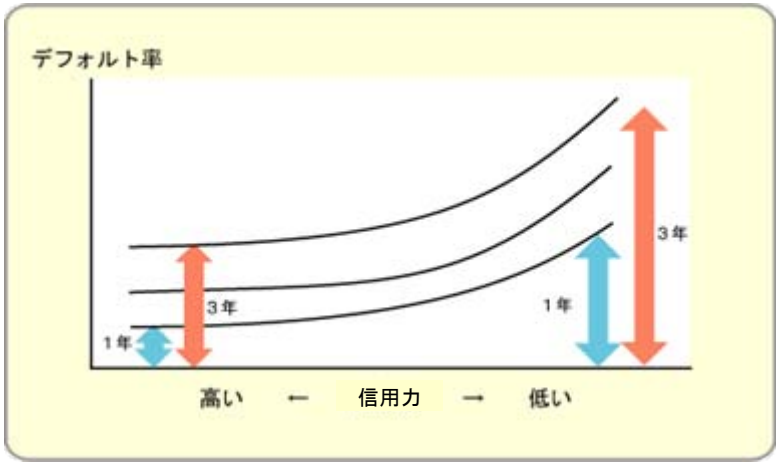
# ⑥risklick（1）概要と特徴　（弊社商品を外部モデルの一例として）

■ 概要
- TSR社が保有するDB登録企業（約112万社）に対し、倒産確率を付与するほか、信用リスク評価に必要な情報をビジュアルに表現し出力。
- 日経テレコン21によるオンライン提供のため、一般の個人ユーザーも含め幅広いご利用。なお、オフライン提供も可能ですので、お問合せください。

■ 特徴①
- 多面的な評価。

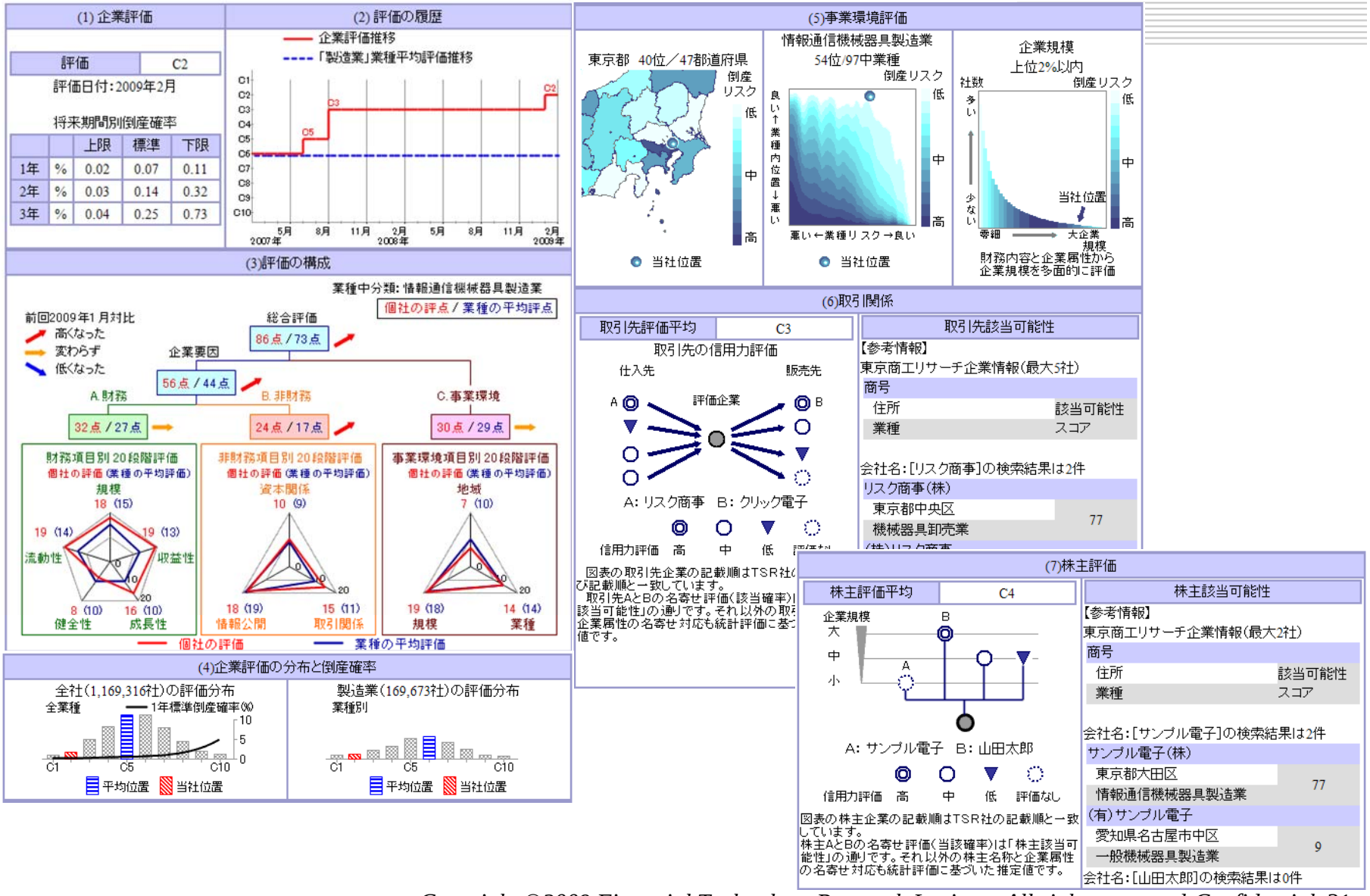

■ 特徴②
- 3年累積倒産確率まで出力。




※一般には、信用力が高い層では1年から3年にかけて、PDの増加率は逓増し、信用力が低い層ではPDの増加率は逓減するという関係になる。

# ⑥risklick（2）出力イメージ



# 4.構造モデルによるデフォルトアラームの活用

統計モデルは基本的に決算書を対象に構成されるため、四半期決算が入手・反映可能としても、年に4回しか結果が更新されない。これに対し、構造モデルは株価を利用するため、それより短い時間スケールでの変化を捉えることが可能である。

ここでは、弊社製品であるDEFENSEの出力結果を例として、デフォルトアラーム検出機能としての役割を果たすことを確認する。

# ①DEFENSE（１）概要と特長　（弊社商品を構造モデルの一例として）

■ 概要
- 日本で取引されている上場および店頭公開企業を対象としたモデル。
- 株価のみならず、決算書情報も加えた総合的判断。
- 企業の信用リスクを「格付け値」、「格付け」、「ステータス」で表現。
  - 格付け値　：DCRI　　DEFENSE Credit Rating Index
  - 格付け　　：DCR　　DEFENSE Credit Rating
  - ステータス：6段階　　Green-1, Green-2, Yellow-1, Yellow-2, Red-1, Red-2

■ 特長
- 日本的経営の要素である株式持合いによる信用力の支えやメインバンクによる信用力バックアップもモデルの要素として組み入れ。
- DEFENSEは通常の理論に加え、より実践に即したモデルにチューニングすることで、高い倒産捕捉力と整合的な序列感を実現。
- DEFENSEのシステム上で、RADARの格付け、ユーザー独自の格付けを（データ連携なされていれば、外部格付会社の格付けも含め）取り込み可能。

# ①DEFENSE（2）主な利用方法

## ■ 期中モニタリング

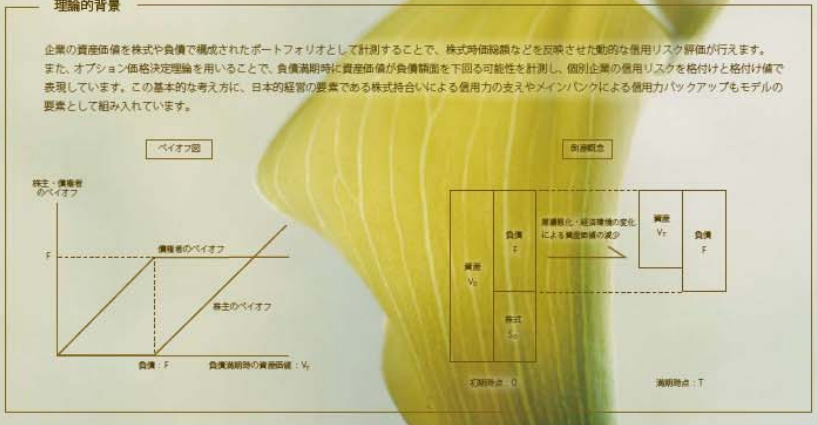
- 次回格付け見直しまでの期中モニタリング情報として活用。
- フロント（証券部等）、ミドル（審査部等）においても幅広くご契約。
- 「Yellow-2以下」 or 「Red-1以下」を注意喚起対象とするのが一般的。
- 「本部→営業部」の連絡経路の確保、情報の共有化に有益。

## ■ 営業審査判断ツール

- 情報の少ない取引先の与信判断情報として利用。
- 決算書等による統計モデルの結果と併せて、対象企業の信用力推移を捕捉。

DEFENSEは、政府系金融機関，都市銀行・地方銀行，資産運用会社等の金融機関・団体を中心に多くのユーザー様にご利用いただいております。

## 高精度な信用リスク評価／データ分析／銘柄管理／レポート機能




新開発の評価結果（DCR：DEFENSE Credit Rating）を利用することで、高度なリスク管理体制が構築できます。

自社ポートフォリオを登録することで、容易なモニタリングが可能です。オンライン版では複数のポートフォリオを登録可能です。部署別、商品別に別名柄を管理する事が可能です。

### ポートフォリオ登録による銘柄管理

条件選定によるレポート出力機能を新たに追加しました。印刷も可能とすることで、自社内での情報共有にお役立ていただけます。

### レポート出力機能

ウェブブラウザーによるオンライン対応となりました。外出先からでも評価結果が随時確認可能です。また、DCRに加えて、市場実勢に即したスプレッドの表示も可能です。

### オンラインに対応

非常に高い評価を頂いております、従来モデルでの基本ロジックはそのままに、モデルのチューニングを実施することで、より高い倒産捕捉力を実現しました。

### 高い倒産捕捉力

ECSからDCR（DEFENSE Credit Rating）およびDCRI（DEFENSE Credit Rating Index）へと信用リスク指標を改訂しました。従来のスプレッド表示に替わって、格付け（DCR）と格付け値（DCRI）を信用リスク指標として算出します。オンライン版では、市場実勢に即したスプレッドも新たに算出します。

### 信用リスク指標を改訂

### 公開企業全てに評価結果を算出

全ての上場会社、店頭公開会社にDCRI（新規株式公開後1年未満を除く）を算出します。DEFENSEの評価対象銘柄は、事業法人、金融法人（銀行・証券・保険）、REIT、の各種法人です。

### 信用リスクのステータス表示

GREEN、YELLOW、REDによるステータス評価を踏襲しました。オンライン版ではステータスが切り替わる水準を任意に設定可能です。また、予め設定した条件に合致するとワーニングメールを配信します。

### 様々な格付と比較可能

DCRは様々な格付けと比較可能です。弊社製品RADAR、および行内格付けや格付け会社の格付けとの比較が可能です。

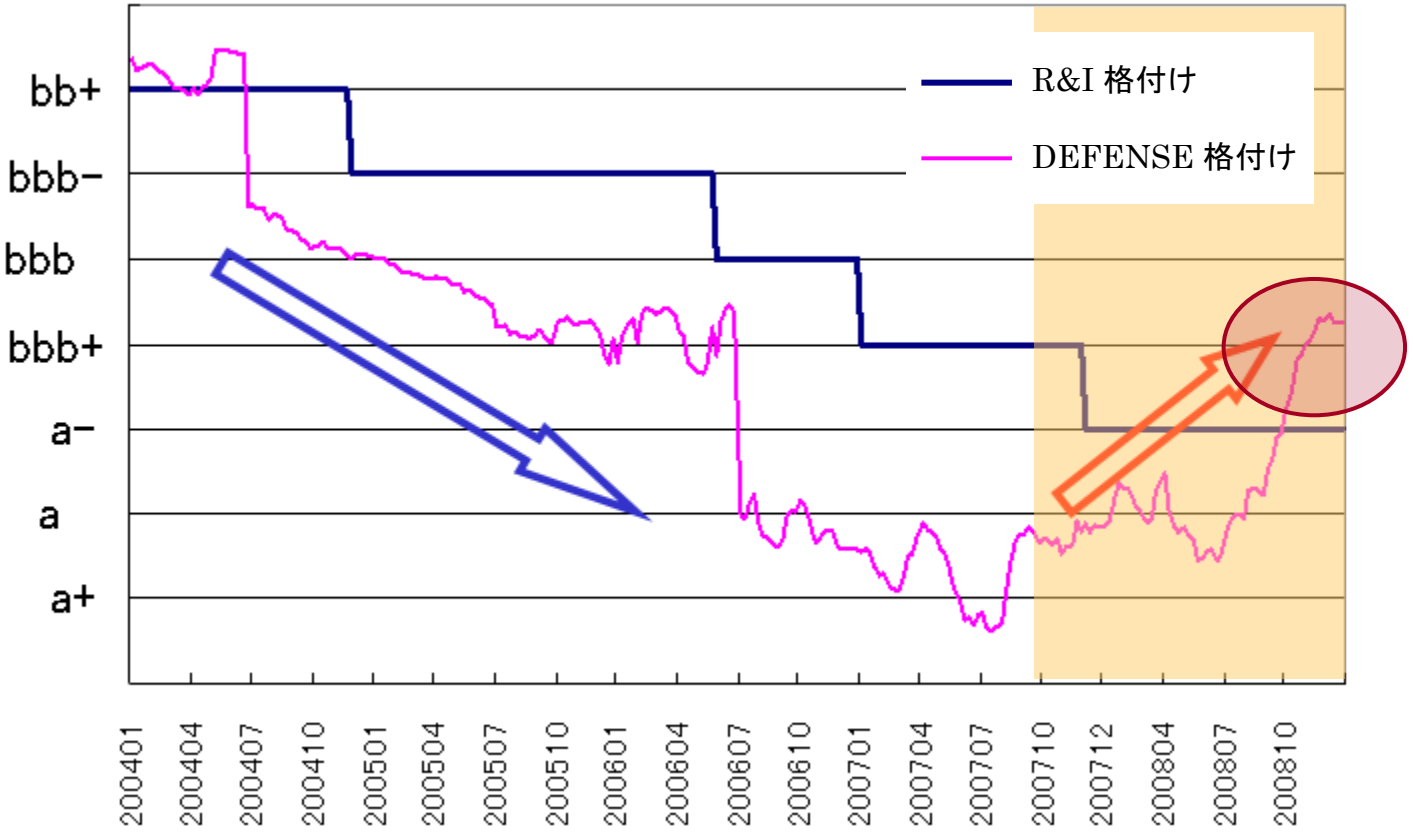
### 信用リスクのわずかな変化も早期に捕捉

株価に関する情報と会計情報を利用して、信用リスクを算出します。評価結果は週1回* 更新されるので、財務データによるスコアリング結果と比べて信用リスクの変化を早期に捕捉することが可能です。（*オフライン版は月2回の提供となります。）

フロントオフィスからミドルオフィス、そしてバックオフィスへ…
横断／垂直的な信用リスク管理が実現します。

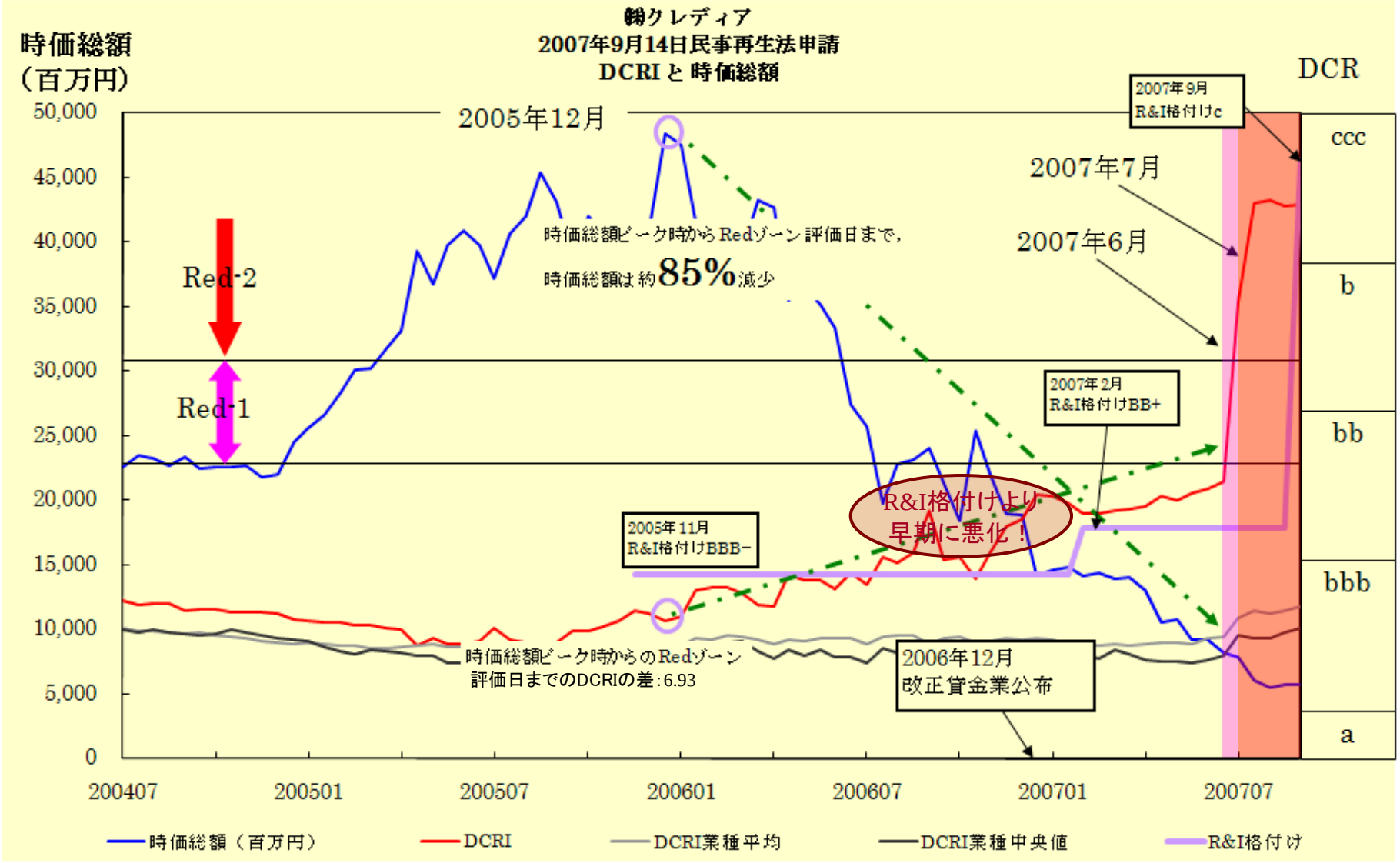
# ②最近の評価傾向

- X社
  - 資源価格の高騰などによる収益力向上による信用力の改善傾向が、足元の燃料価格下落や景気悪化による信用力悪化を受けて急速に悪化している。



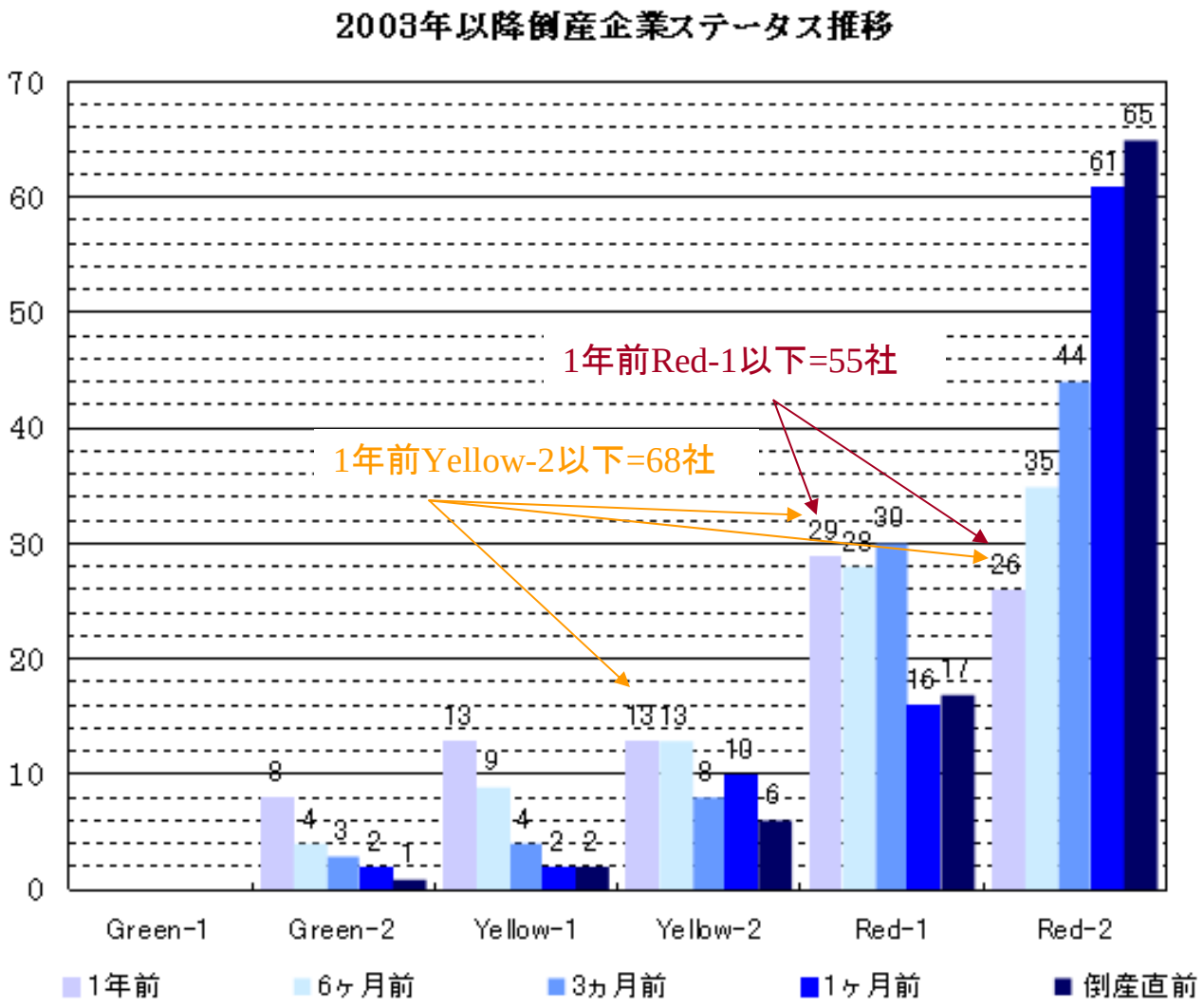
# ③倒産企業の事例



# ④倒産企業のステータス推移

- ステータス推移
  - 2003～08年に倒産した企業91社が対象。
  - 倒産に近づくにつれて、ステータスが悪化。
  - 1年前においても、55社がRed-1以下、68社がYellow-2以下と精度が高い。



# ⑤まとめと今後の展開

- DEFENSEの評価指標（DCRI、DCR）は、一般にR&I格付に先行して、評価が変動しているため、期中モニタリング等で効果を発揮する。

- 倒産企業の具体的な事例からは、倒産に至るまで時価総額は減少傾向であり、3年前の価値のほとんどが失われてしまっている。

- 1年前は比較的評価が高かった企業においても、倒産日に近づくにつれて評価を下げていく様子が確認できた。

- 今後は、四半期決算情報も取り込むことで、よりタイムリーな評価を可能とするよう開発中である。

# 5.格付およびモデルの検証

新BIS規制の要件にあることからも、近年、金融機関の内部監査セクションには、現場セクションである信用リスク管理部署が設定したモデルおよび格付について、独立に検証・評価し、牽制していくというミッションが与えられてきている。つまり、内部監査セクションといえども、現場セクション同様のレベル感での理解度が必要であり、数理的な側面も含めての対応が必要と考えられる。

本章で扱う検証方法は、極めて技術的な数理的検証に留まる。しかし、他にも格付制度そのもの、運用・管理状況の調査等、内部監査セクションに求められる課題が多数あることだけは指摘しておきたい。

# ①新BIS規制におけるモデル検証

【金融庁告示第19号（平成18年3月27日）第百八十九条】
内部格付手法採用行は、債務者格付若しくは案件格付の付与又はPD、LGD 及びEAD の推計に統計的モデルその他の機械的な手法（以下「モデル」と総称する。）を用いる場合は、次に掲げるすべての要件を満たさなければならない。
（略）
六　モデルの運用実績及び安定性の評価、モデルとモデルの前提となっている状況の関連性の見直し、実績値とモデルの予測値の対照その他のモデルの検証が定期的に行われること。

新BISの内部格付手法では、格付に利用するモデルを定期的に検証することが求められており、検証の第一義的な責任は銀行にある[1]とされています。

銀行が定期的にモデルを検証する上で重要な要素は？

■定量的処理を行なう人材、またはシステムの配置　■検証作業の安定運用

■検証ノウハウの蓄積と継承　■新検証方法への柔軟な対応、過去への遡及

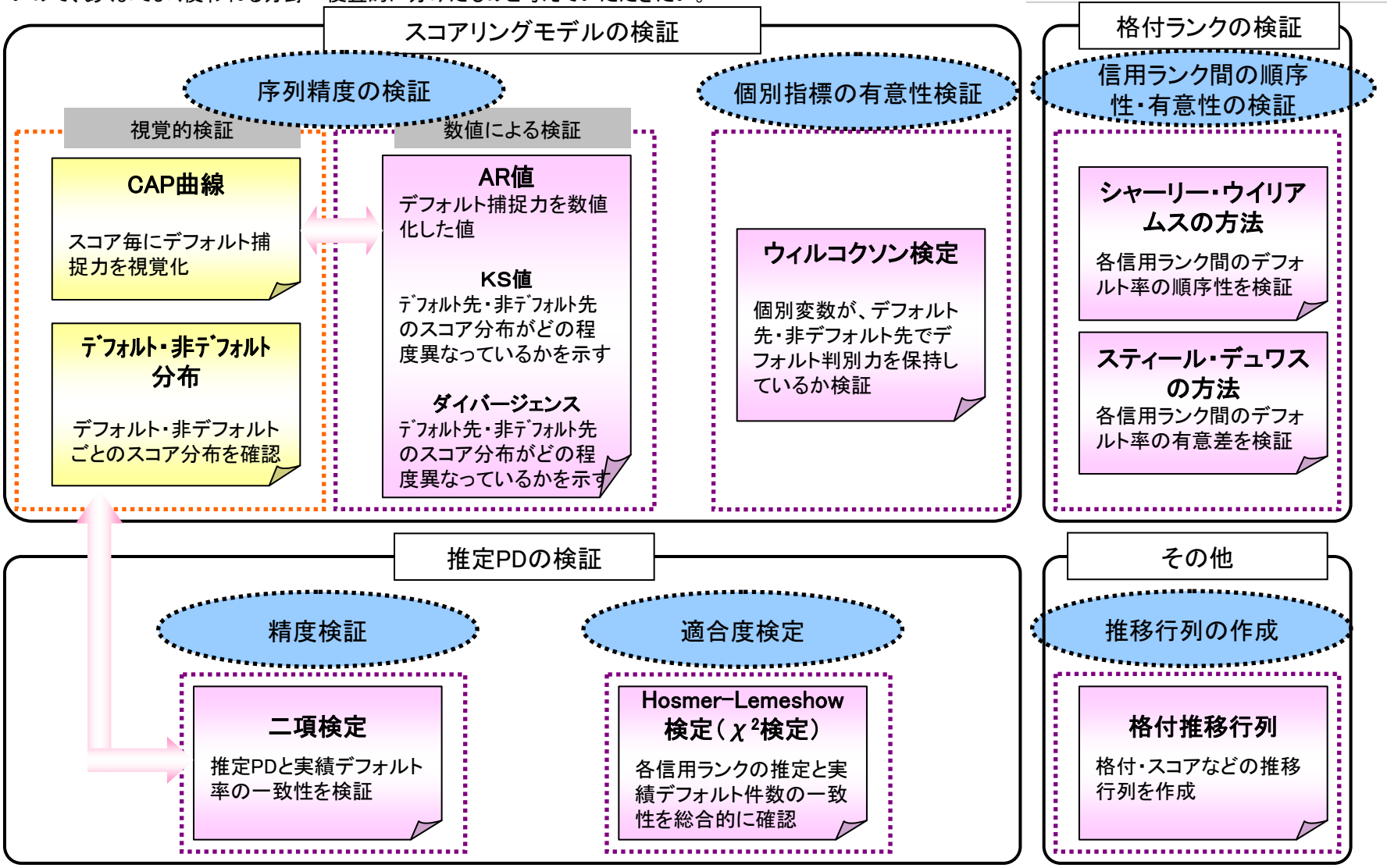
1 Basel Committee Newsletter No.4(January 2005) Principle2：The bank has primary responsibility for validation.

※（新）BIS規制とは、国際的に業務展開を行う銀行の健全性を保つために定められたルールであり、象徴的には、分母をリスク量とした自己資本比率を8%以上（国際業務を行わない国内銀行においては4%以上）なければならないというものである。問題は、分母のリスク量を計算する方法に幾つか選択肢があり、信用リスク分野においては内部格付手法採用行（いうなればモデルを利用する銀行）は、モデル検証する義務があるとされている。なお、日本では、「銀行法第14条の2の規程に基づき、銀行がその保有する資産等に照らした自己資本の充実の状況が適当であるかどうかを判断するための基準」という金融庁告示にその規制内容が記されている。

# ②検証手法一覧

※たとえば、AR値は「スコアリングモデルの検証」に分類されているが、「格付ランクの検証」に使えないわけではないので、あくまでよく使われる分野へ便宜的に分けたものと考えていただきたい。



# ③序列精度の検証（1）AR値の信頼区間とデフォルト件数の関係

- 事例1：Englemann,Hayden,Tascheの推計例
  - データ：1987～1993年
  - デフォルト先数が異なるポートフォリオに対するAR値の信頼区間を計算。
- 事例2：BISペーパーの計算例。
  - AR値一定の下でデフォルト数を変化させた場合の考察。
- 計算方法
  - AR値の信頼区間は、ROC曲線から導かれるAUC（Area Under Curve）をベースに計算される。デフォルト先、非デフォルト先をランダムサンプリングしAUC信頼区間を計算。以下の式よりAUCからAR値に変換。

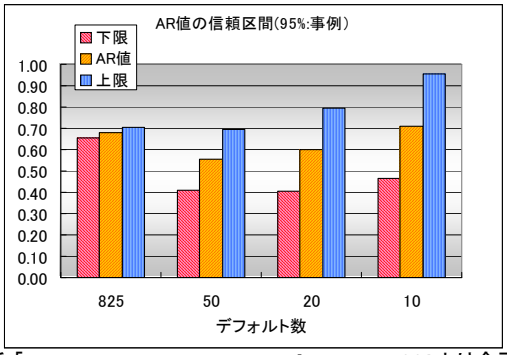
$$AR値=2\times AUC-1$$

※実はランダムサンプリングといったシミュレーションなしに、近似的にではあるが、理論的に信頼区間を構成することが可能である。

AR値はデフォルト数に依存する。デフォルト件数が少ないほど信頼区間は広く、得られたAR値が安定的とは言えなくなることに留意する。

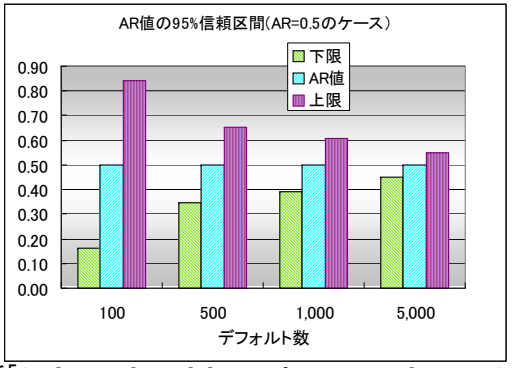
- 事例1

| ケース | 債務者総数 | 正常先 | デフォルト先 | 実績PD | AR値 | 95%信頼区間 | | 99%信頼区間 | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | 下限 | 上限 | 下限 | 上限 |
| 1 | 200,000 | 199,175 | 825 | 0.41% | 0.680 | 0.656 | 0.705 | 0.650 | 0.712 |
| 2 | 500 | 450 | 50 | 10.00% | 0.554 | 0.409 | 0.697 | 0.364 | 0.742 |
| 3 | 500 | 480 | 20 | 4.00% | 0.602 | 0.406 | 0.796 | 0.347 | 0.855 |
| 4 | 500 | 490 | 10 | 2.00% | 0.710 | 0.463 | 0.955 | 0.386 | 1.000 |




出所:「Testing rating accuracy」 Risk january2003より金融工学研究所が計算

- 事例2




出所「Studies on the Validation of International Rating Systems “ BIS 2005年5月」より金融工学研究所が計算

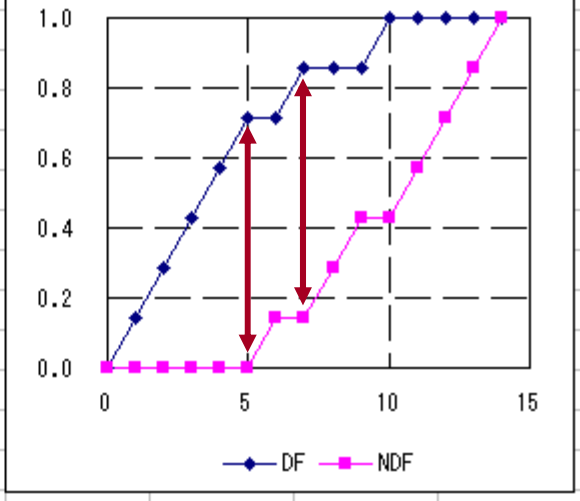
# ③序列精度の検証（2）KS値：Kolmogorov-Smirnov

Microsoft Excel - Book1.xls

ファイル(F) 編集(E) 表示(V) 挿入(I) 書式(O) ツール(T) データ(D) ウィンドウ(W) ヘルプ(H) Adobe PDF(B)

L3 =L2+(1-B3)/($B$18-$B$17)

| | A | B | C | D | J | K | L | M | N |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | k | DF | NDF | DF - NDF | |
| 2 | i | δ i | x1 i | pi | 0 | 0.000 | 0.000 | 0.000 | |
| 3 | 1 | 1 | 1.932 | 1.000 | 1 | 0.143 | 0.000 | 0.143 | |
| 4 | 3 | 1 | 1.053 | 0.993 | 2 | 0.286 | 0.000 | 0.286 | |
| 5 | 2 | 1 | 0.993 | 0.983 | 3 | 0.429 | 0.000 | 0.429 | |
| 6 | 5 | 1 | 0.896 | 0.939 | 4 | 0.571 | 0.000 | 0.571 | |
| 7 | 4 | 1 | 0.881 | 0.926 | 5 | 0.714 | 0.000 | 0.714 | * |
| 8 | 13 | 0 | 0.720 | 0.576 | 6 | 0.714 | 0.143 | 0.571 | |
| 9 | 6 | 1 | 0.693 | 0.483 | 7 | 0.857 | 0.143 | 0.714 | * |
| 10 | 14 | 0 | 0.670 | 0.405 | 8 | 0.857 | 0.286 | 0.571 | |
| 11 | 9 | 0 | 0.617 | 0.247 | 9 | 0.857 | 0.429 | 0.429 | |
| 12 | 7 | 1 | 0.588 | 0.180 | 10 | 1.000 | 0.429 | 0.571 | |
| 13 | 10 | 0 | 0.563 | 0.134 | 11 | 1.000 | 0.571 | 0.429 | |
| 14 | 12 | 0 | 0.550 | 0.115 | 12 | 1.000 | 0.714 | 0.286 | |
| 15 | 8 | 0 | 0.403 | 0.017 | 13 | 1.000 | 0.857 | 0.143 | |
| 16 | 11 | 0 | 0.264 | 0.002 | 14 | 1.000 | 1.000 | 0.000 | |
| 17 | DF | 7 | | | | | | | |
| 18 | AL | 14 | | | | | KS | 0.714 | |




KS値=max(DF-NDF)

Kolmogorov-Smirnov

PDの降順にソート

*x*座標：高PD側からの順位

*y*座標：DF ＝デフォルトサンプルの累積割合
NDF＝非デフォルトサンプルの累積割合

※*x*座標は、PDでもZスコアでも可。悪い順にならべばよい。KS値は曲線間の開き具合のみわかればよいため。

※KS値は、2つの曲線の間が最も広がったときの値であり、その値が大きければ良いと判断する指標である。また、次のような解釈を与えていた時代もある。左図において5番目までの顧客への融資を否決していたら、機会損失なくデフォルト先の71%を否決していたと考えることができるし、7番目までの顧客を否決していれば、非デフォルト先の否決は14%出てしまうものの、デフォルト先の実に86%を否決できたと考えることができる。現在でこそAR値がデファクトとなってしまったが、解釈を与える上では、左図やKS値はわかりやすい。

# ③序列精度の検証（3）ダイバージェンス：Divergence

Microsoft Excel - Book1.xls

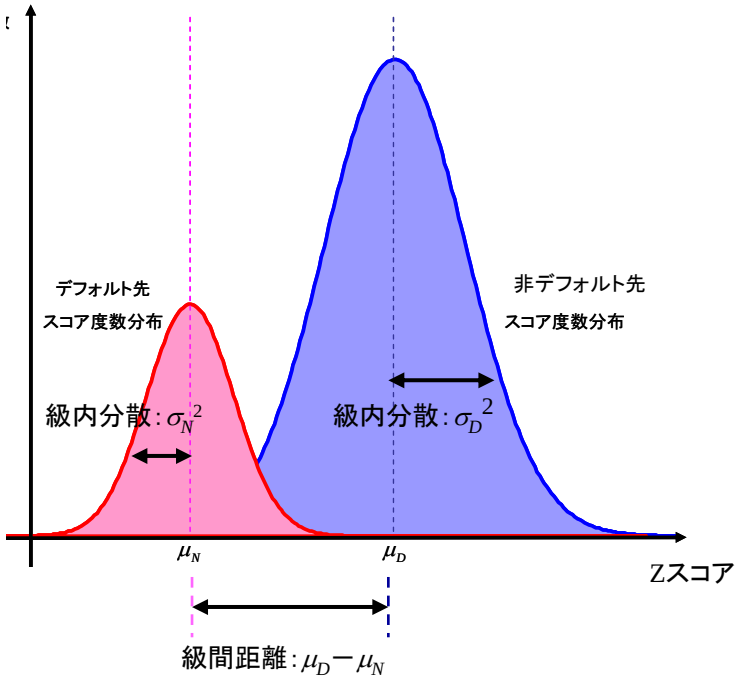
ファイル(F) 編集(E) 表示(V) 挿入(I) 書式(O) ツール(T) データ(D)
ウィンドウ(W) ヘルプ(H) Adobe PDF(B)

I8 =2*(I4-I6)^2/(J4^2+J6^2)

| | A | B | C | D | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|
| 2 | i | δ i | x1 i | pi | Zi | | | | |
| 3 | 1 | 1 | 1.932 | 1.000 | 17.053 | | Zi: | μ D | σ D |
| 4 | 2 | 1 | 0.993 | 0.983 | 4.078 | | | 4.246 | 6.076 |
| 5 | 3 | 1 | 1.053 | 0.993 | 4.907 | | | μ N | σ N |
| 6 | 4 | 1 | 0.881 | 0.926 | 2.530 | | | -2.167 | 2.191 |
| 7 | 5 | 1 | 0.896 | 0.939 | 2.738 | | | Div_Zi | |
| 8 | 6 | 1 | 0.693 | 0.483 | -0.067 | | | 1.9718 | |
| 9 | 7 | 1 | 0.588 | 0.180 | -1.518 | | | | |
| 10 | 8 | 0 | 0.403 | 0.017 | -4.074 | | Pi: | μ D | σ D |
| 11 | 9 | 0 | 0.617 | 0.247 | -1.117 | | | 0.786 | 0.324 |
| 12 | 10 | 0 | 0.563 | 0.134 | -1.863 | | | μ N | σ N |
| 13 | 11 | 0 | 0.264 | 0.002 | -5.995 | | | 0.214 | 0.211 |
| 14 | 12 | 0 | 0.550 | 0.115 | -2.043 | | | Div_pi | |
| 15 | 13 | 0 | 0.720 | 0.576 | 0.306 | | | 4.3821 | |
| 16 | 14 | 0 | 0.670 | 0.405 | -0.385 | | | | |

DIV

コマンド NUM




$$Divergence: Div = \frac{2 \times (\mu_D - \mu_N)^2}{\sigma_D^2 + \sigma_N^2}$$

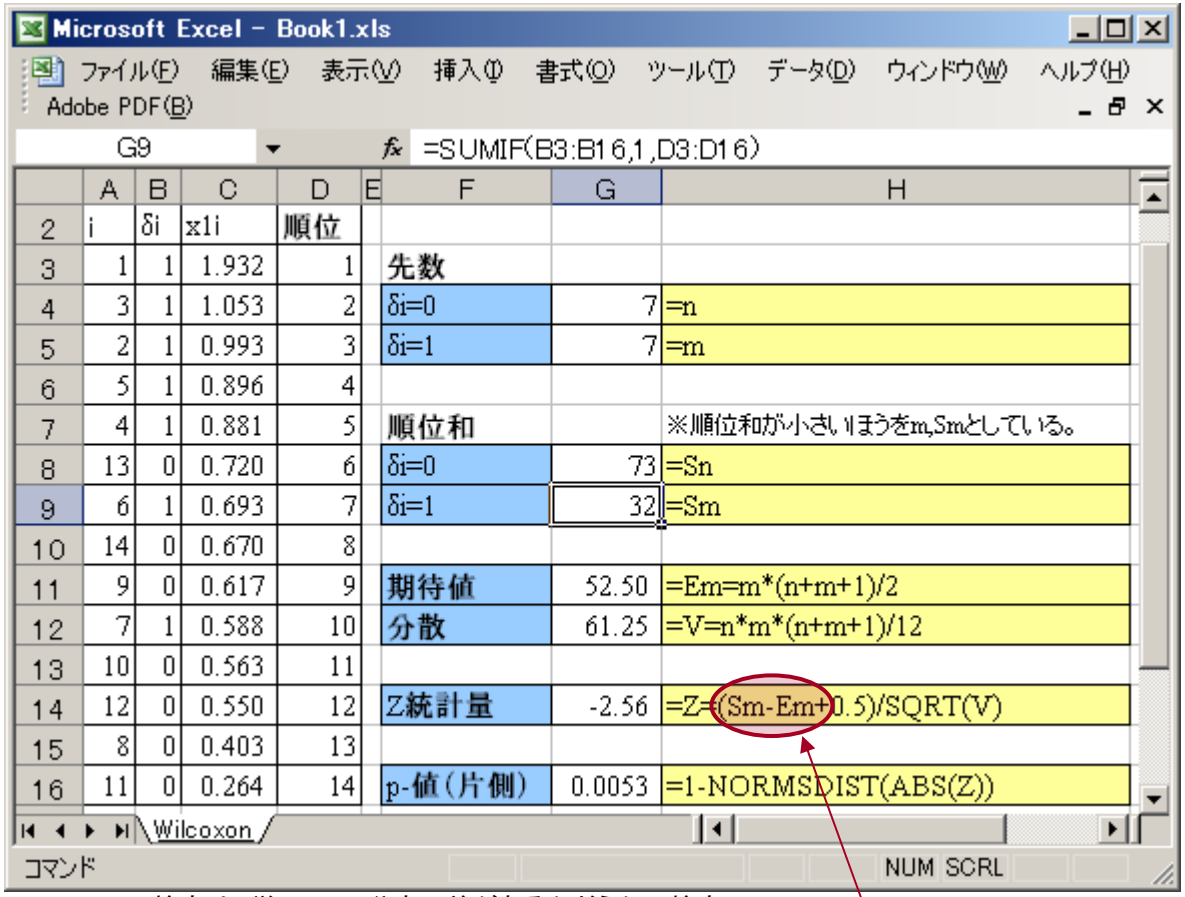
※ダイバージェンスは、デフォルトサンプルのZスコア分布と非デフォルトサンプルのZスコアの分布が正規分布に従うと仮定して計算される指標であるので、パラメトリック手法の一種である。ダイバージェンス値は、2つの分布の違いを図る尺度であり、非デフォルトサンプルの分布とデフォルトサンプルの分布がより大きく離れている場合に大きなダイバージェンス値をとる。

※ダイバージェンスは、何を対象として計算するかで結果が異なることに注意が必要である。

# ③序列精度の検証（4）ウィルコクソン順位和検定：Wilcoxon

- 概要
  - スコアリングモデルにおける財務比率等の説明変数がデフォルト判別に有効であるか検定するための手法。
  - 財務比率の値ではなく、財務比率の大きさの順位情報のみ利用。
  - 順位の和に偏在性があるかどうかを正規分布を利用して検定（右図におけるZ統計量がサンプル数が大きければ、すなわち漸近的に正規分布に従う）。
  - p-値が$x$%のとき、有意水準が$x$%超の検定では、差があるということになり、当該財務比率はデフォルト判別に有効とされる。




| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 2 | i | δi | x1i | 順位 | | | | |
| 3 | 1 | 1 | 1.932 | 1 | | 先数 | | |
| 4 | 3 | 1 | 1.053 | 2 | | δi=0 | 7 | =n |
| 5 | 2 | 1 | 0.993 | 3 | | δi=1 | 7 | =m |
| 6 | 5 | 1 | 0.896 | 4 | | | | |
| 7 | 4 | 1 | 0.881 | 5 | | 順位和 | | ※順位和が小さいほうをm,Smとしている。 |
| 8 | 13 | 0 | 0.720 | 6 | | δi=0 | 73 | =Sn |
| 9 | 6 | 1 | 0.693 | 7 | | δi=1 | 32 | =Sm |
| 10 | 14 | 0 | 0.670 | 8 | | | | |
| 11 | 9 | 0 | 0.617 | 9 | | 期待値 | 52.50 | =Em=m*(n+m+1)/2 |
| 12 | 7 | 1 | 0.588 | 10 | | 分散 | 61.25 | =V=n*m*(n+m+1)/12 |
| 13 | 10 | 0 | 0.563 | 11 | | | | |
| 14 | 12 | 0 | 0.550 | 12 | | Z統計量 | -2.56 | =Z=(Sm-Em+0.5)/SQRT(V) |
| 15 | 8 | 0 | 0.403 | 13 | | | | |
| 16 | 11 | 0 | 0.264 | 14 | | p-値（片側） | 0.0053 | =1-NORMSDIST(ABS(Z)) |

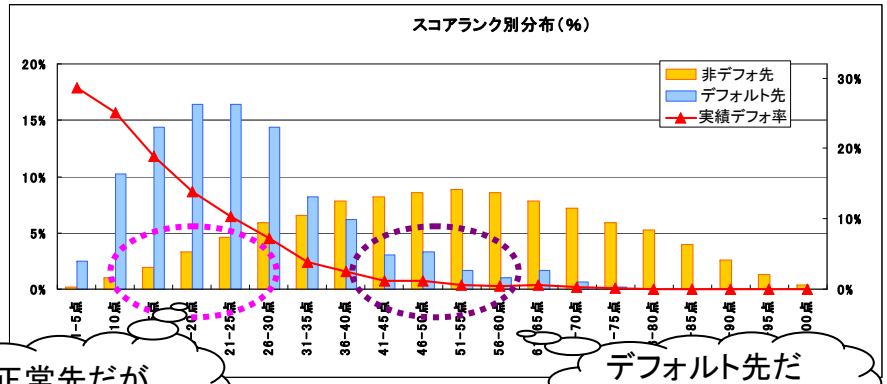
※Wilcoxon検定は、単に2つの分布に差があるかどうかの検定であり、デフォルト先のほうが非デフォルト先に比較して当該指標が悪化する側に偏在しているか確認が必要。
※$x_1$が同一の値をもつ場合は、順位は平均し、分散の計算式は、分散を小さくする補正項が付加（詳細は省略）。
※正規近似を用いず行う正確検定（Exact Test）もあり、実際のp-値は0.0035と厳しい有意水準でも耐えうる結果となった。

実際の順位和$S_m$と、ランダムな場合の順位和$E_m$との差が広いほど、有効な指標と言える。

# ④推定PDの検証（1）デフォルト・非デフォルト分布

■ PDランク別のデフォルト・非デフォルト分布と実績デフォルト率




正常先だが低スコア

デフォルト先だが高スコア

格付・モデルがデフォルトした債務者とデフォルトしなかった債務者を、正しく判別しているか？

これらの分析を、より定量的に実施し、評価するにはどうすればよいか？

⇒二項検定、Hosmer-Lemeshow検定

■ 各カテゴリ別に推定PDと実績デフォルト率の比較

| | | | サンプル数 | 非デフォルト数 | デフォルト数 | 実績デフォ率 | 推定PD |
|---|---|---|---|---|---|---|---|
| 全体 | | | 41,160 | 40,000 | 1,160 | 2.82% | 3.12% |
| 業種別 | 1 | 製造業 | 10,250 | 10,000 | 250 | 2.44% | 3.00% |
| | 2 | 建設業 | 6,200 | 6,000 | 200 | 3.23% | 4.50% |
| | 3 | 卸売業 | 5,100 | 5,000 | 100 | 1.96% | 1.50% |
| | 4 | 小売業 | 8,250 | 8,000 | 250 | 3.03% | 3.00% |
| | 5 | 不動産業 | 2,080 | 2,000 | 80 | 3.85% | 4.50% |
| | 6 | サービス業 | 6,180 | 6,000 | 180 | 2.91% | 3.00% |
| | 7 | その他 | 3,100 | 3,000 | 100 | 3.23% | 3.00% |
| 規模別 | 1 | 大企業・中堅企業 | 15,350 | 15,000 | 350 | 2.28% | 2.50% |
| | 2 | 中小企業 | 14,450 | 14,000 | 450 | 3.11% | 3.64% |
| | 3 | 零細企業 | 11,360 | 11,000 | 360 | 3.17% | 3.27% |
| 地域 | 1 | ○県 | 13,350 | 13,000 | 350 | 2.62% | 3.00% |
| | 2 | △県 | 12,380 | 12,000 | 380 | 3.07% | 3.75% |
| | 3 | □県 | 15,430 | 15,000 | 430 | 2.79% | 2.71% |
| 時系列 | | 2001年 | 6,997 | 6,809 | 188 | 2.69% | 2.40% |
| | | 2002年 | 6,537 | 6,321 | 216 | 3.30% | 2.52% |
| | | 2003年 | 2,193 | 2,124 | 69 | 3.15% | 2.41% |

各属性別に、モデルが算出したPDが実績デフォルト率と整合的であり大きな乖離がないか？

推定PDと実績デフォルト率の乖離

# ④推定PDの検証（2）二項検定

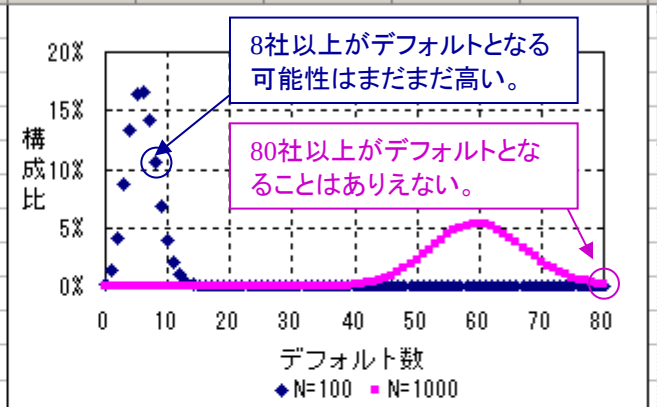
実績デフォルト率が推定PDの信頼区間内になければ両者の間には「差がある」と捉える。
サンプル数が増加すると、信頼区間は狭くなる。

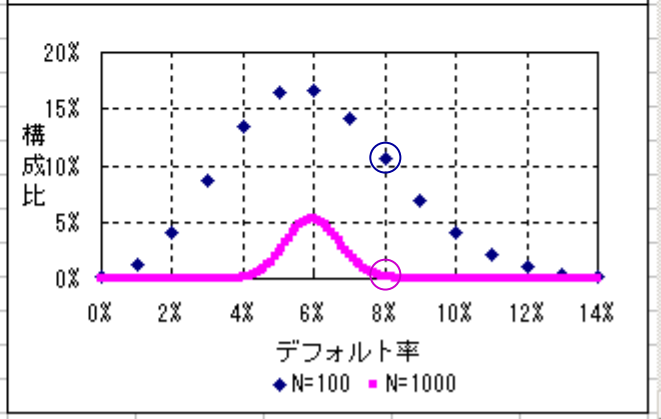
Microsoft Excel - Book1.xls

ファイル(F) 編集(E) 表示(V) 挿入(I) 書式(O) ツール(T) データ(D) ウィンドウ(W) ヘルプ(H) Adobe PDF(B) 質問を入力してください

H3 =BINOMDIST($F3,C$3,C$6,FALSE)

| | B | C | D | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | N=100 | | | |
| 2 | ＜前提＞ | | | j | PD | 構成比 | *1 | *2 |
| 3 | サンプル数＝N | 100 | 1,000 | 0 | 0.0% | 0.21% | 0.21% | 100.00% |
| 4 | デフォルト数 | 8 | 80 | 1 | 1.0% | 1.31% | 1.52% | 99.79% |
| 5 | 実績デフォルト率 | 8% | 8% | 2 | 2.0% | 4.14% | 5.66% | 98.48% |
| 6 | 平均PD | 6% | 6% | 3 | 3.0% | 8.64% | 14.30% | 94.34% |
| 7 | 有意水準 | 5% | 5% | 4 | 4.0% | 13.38% | 27.68% | 85.70% |
| 8 | | | | 5 | 5.0% | 16.39% | 44.07% | 72.32% |
| 9 | ＜結果：信頼区間＞ | | | 6 | 6.0% | 16.57% | 60.64% | 55.93% |
| 10 | デフォルト数 下限 | 1 | 45 | 7 | 7.0% | 14.20% | 74.83% | 39.36% |
| 11 | デフォルト数 上限 | 12 | 76 | 8 | 8.0% | 10.54% | 85.37% | 25.17% |
| 12 | デフォルト率 下限 | 1.0% | 4.5% | 9 | 9.0% | 6.87% | 92.25% | 14.63% |
| 13 | デフォルト率 上限 | 12.0% | 7.6% | 10 | 10.0% | 3.99% | 96.24% | 7.75% |
| 14 | 構成比（確率関数）は、平均PDを前提 | | | 11 | 11.0% | 2.09% | 98.32% | 3.76% |
| 15 | にサンプル数がNの場合に得られるデ | | | 12 | 12.0% | 0.99% | 99.31% | 1.68% |
| 16 | フォルト数jの分布。jをNで除して、PDも | | | 13 | 13.0% | 0.43% | 99.74% | 0.69% |
| 17 | 表示した。 | | | 14 | 14.0% | 0.17% | 99.91% | 0.26% |
| 18 | *1は低PD側からの累積で、*2は高いほ | | | 15 | 15.0% | 0.06% | 99.97% | 0.09% |
| 19 | うからの累積。特に*1は分布関数と呼 | | | 16 | 16.0% | 0.02% | 99.99% | 0.03% |
| 20 | ばれる。 | | | 17 | 17.0% | 0.01% | 100.00% | 0.01% |
| 21 | | | | 18 | 18.0% | 0.00% | 100.00% | 0.00% |
| 22 | 有意水準5%なので、片側2.5%の検定 | | | 19 | 19.0% | 0.00% | 100.00% | 0.00% |
| 23 | の場合、*1で2.5%以下の最大のjがデ | | | 20 | 20.0% | 0.00% | 100.00% | 0.00% |
| 24 | フォルト数下限、*2で2.5%以下の最小 | | | 21 | 21.0% | 0.00% | 100.00% | 0.00% |
| | のjがデフォルト数の上限になる。 | | | | | | | |




二項検定

コマンド NUM

※この検定では、各々のサンプルは独立、すなわちサンプルと別のサンプルの間のデフォルト、非デフォルトに相関はないということが前提になっている。これらの間の相関を考慮した二項検定というのもあるが、相関係数をどう与えるかという問題が残る。基本的に、正相関が強くなるほど信頼区間は広がる。

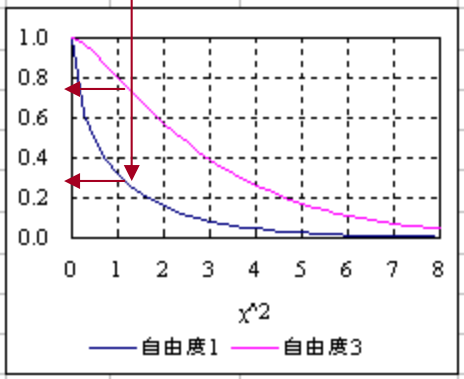
# ④推定PDの検証（3）Hosmer-Lemeshow検定

Microsoft Excel - Book1.xls

ファイル(F) 編集(E) 表示(V) 挿入(I) 書式(O) ツール(T) データ(D) ウィンドウ(W) ヘルプ(H) Adobe PDF(B)

R9 =CHIDIST(R6,1)

| | A | B | C | D | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | ① | ② | ③ | ④ | ⑤ | ⑥ | ⑦ |
| 2 | i | δi | x1i | pi | | k | ALL | DF | pi平均 | =①×③ | =(②－④)^2 | =①×③×(1-③) | =⑤/⑥ |
| 3 | 11 | 0 | 0.264 | 0.002 | | 1 | 5 | 1 | 0.090 | 0.448 | 0.305 | 0.408 | 0.747 |
| 4 | 8 | 0 | 0.403 | 0.017 | | 2 | 5 | 2 | 0.527 | 2.637 | 0.405 | 1.246 | 0.325 |
| 5 | 12 | 0 | 0.550 | 0.115 | | 3 | 4 | 4 | 0.979 | 3.915 | 0.007 | 0.083 | 0.087 |
| 6 | 10 | 0 | 0.563 | 0.134 | | | | | | | 合計 | χ^2= | 1.159 |
| 7 | 7 | 1 | 0.588 | 0.180 | | | | | | | | | |
| 8 | 9 | 0 | 0.617 | 0.247 | | | | | | | | 自由度1のp-値 | |
| 9 | 14 | 0 | 0.670 | 0.405 | | | | | | | | CHIDIST(χ^2,1)= | 0.282 |
| 10 | 6 | 1 | 0.693 | 0.483 | | | | | | | | | |
| 11 | 13 | 0 | 0.720 | 0.576 | | | | | | | | 自由度3のp-値 | |
| 12 | 4 | 1 | 0.881 | 0.926 | | | | | | | | CHIDIST(χ^2,3)= | 0.763 |
| 13 | 5 | 1 | 0.896 | 0.939 | | | | | | | | | |
| 14 | 2 | 1 | 0.993 | 0.983 | | | | | | | | | |
| 15 | 3 | 1 | 1.053 | 0.993 | | | | | | | | | |
| 16 | 1 | 1 | 1.932 | 1.000 | | | | | | | | | |




Hosmer-Lemeshow

コマンド

格付別に、実績デフォルト数と想定デフォルト数の間の差異を見ている。

$$\chi^2_{HL} = \sum_{k=1}^{K} \frac{(DF_k - N_k PD_k)^2}{N_k PD_k (1 - PD_k)}$$

$DF_k$ :ランク　kの実績デフォルト数

$PD_k$ :ランク　kの平均PD

$N_k$ :ランク　kのサンプル数

格付別の実績デフォルト率と推定PDの間に乖離があるかを検定。乖離があってはいけないのであるから、p-値は大きいほどよい。

なお、モデル構築用データにて検定を行う場合は、ペナルティーとして自由度を2減じて実施。

※検証用データと考えれば、格付数が3なので、自由度3の$\chi^2$分布からp-値は0.763。しかし、本例の場合、モデル構築に利用したデータであるため、PDの順序が揃っているのは当然であり、その分のペナルティー2を減じた自由度1の$\chi^2$分布からp-値を計算すれば0.282。いずれにせよ、有意水準10%でも乖離がないと言えるので、全体として格付設定がうまくなされたと捉えることができる。

※なお、本例は説明の都合上、前述の数値例を引き続いて利用しているものであるが、実際にはサンプル数も多く格付は10ランク程度に区分して検定がなされるのが通常である。

# ⑤格付ランクの検証：スティール・デュワス、シャーリー・ウィリアムズ

- 多重比較とは
  - デフォルト先と非デフォルト先ではなく、3つ以上の格付間比較に必要な方法。
  - 2つの格付間の差異の検証（たとえばウィルコクソン順位和検定）を繰り返したのでは、3つ以上の格付間で相互に差異があることの検証にはならない。

※1つの部品の故障率が0.1%だとしても、10個もあれば全体で$1-(1-0.1\%)^{10}\fallingdotseq 1\%$になってしまう。よくロケット部品でなされる議論である。これとのアナロジーで任意の2つの格付間で5%有意だとしても、3格あれば $1-(1-5\%)^3\fallingdotseq 14\%$になってしまう（格付間は独立という前提）。

  - そこで、全体で必要な有意水準が確保できるよう、2つ（または2群）の格付間での比較時に、厳しめの有意水準（一種のペナルティー付き状況）下で検定していく必要。

※以下では、ウィルコクソン順位和検定同様、順位情報のみ利用した（ノンパラメトリックな）検定法のみ簡単に紹介する。検定結果の表示例は、モデルチェッカーEXの節にて紹介する。

- スティール・デュワス（有意差）検定
  - 任意の2つの格付$i,j$間の検証対象変数の順位和に偏在性があるかどうかを検定。
  - 検定統計量は、ウィルコクソン順位和検定時のZ統計量と同様（$t_{ij}$）。
  - 自由度∞、格付数$K$、有意水準$\alpha$のスチューデント化された分布$q(K,\infty;\alpha)$との比較を行う。
  - $|t_{ij}| \geqq q/\mathrm{SQRT}(2)$であれば差があると判断。

※ウィルコクソン順位和検定は正規分布による検定だが、本方法では、それに全体の有意水準確保のため、スチューデント化された分布を用いているのが特徴である。

※シャーリー・ウィリアムズ検定は、任意の格付に対し最下位格付から順に差があることを調べていく検定とも言える。

- シャーリー・ウィリアムズ（順序性）検定
  - 格付ランクと検証対象変数が単調増加・減少であることを前提として、どの格付から差異があるか検定。
  - 検定統計量は、格付$i$と格付$p$（$>i$）では、$t_{ip}=(M_p-U_{ip})/\mathrm{SQRT}(V_p(1/n_p+1/n_1))$。
  - $M_p$:＝格付$i$+1～$p$, $i$+2 ～$p$,…,$p$の平均順位和最大値、$U_{ip}$:＝格付$i$の平均順位和、$V_p$:=分散、$n_i,n_p$:=格付$i$と格付$p$の先数。
  - 自由度∞ 、参照格付数$p$-$i$+1、有意水準$\alpha$のウィリアムズの方法のための分布$w(p\text{-}i\text{+}1,\infty;\alpha)$との比較を行う。
  - $t_{ip} \geqq$ wであれば差があると判断。
  - 上記がOKなら、格付$p$を$p$-1にして繰り返し。

# ⑥格付推移行列

Microsoft Excel - Book2.xls

D11　81.36%

t+1期

| | 1年 | 1格 | 2格 | 3格 | DF |
|---|---|---|---|---|---|
| t期 | 1格 | 90% | 8% | 2% | 0.00% |
| | 2格 | 4% | 86% | 8% | 2.00% |
| | 3格 | 2% | 8% | 85% | 5.00% |
| | DF | 0% | 0% | 0% | 100% |

A

累積デフォルト率

| | t期 | t+1期 | t+2期 | t+3期 | t+4期 | t+5期 |
|---|---|---|---|---|---|---|
| 1格 | | 0.00% | 0.26% | 0.75% | 1.45% | 2.32% |
| 2格 | | 2.00% | 4.12% | 6.31% | 8.52% | 10.73% |
| 3格 | | 5.00% | 9.41% | 13.33% | 16.85% | 20.04% |
| DF | | 100.00% | 100.00% | 100.00% | 100.00% | 100.00% |

※t+1期にデフォルトゼロでも、格付低下により、次期以降デフォルトが発生。いわゆるLDP先のPDとしては累積デフォルト率を年換算すべきというのは、この議論から得られる帰結。

t+2期

| | 1年 | 1格 | 2格 | 3格 | DF |
|---|---|---|---|---|---|
| t期 | 1格 | 81% | 14% | 4% | 0.26% |
| | 2格 | 7% | 75% | 14% | 4.12% |
| | 3格 | 4% | 14% | 73% | 9.41% |
| | DF | 0% | 0% | 0% | 100% |

AA

先数推移

| | t期 | t+1期 | t+2期 | t+3期 | t+4期 | t+5期 |
|---|---|---|---|---|---|---|
| 1格 | 200 | 206 | 210 | 213 | 214 | 215 |
| 2格 | 500 | 470 | 445 | 423 | 404 | 387 |
| 3格 | 300 | 299 | 296 | 291 | 286 | 279 |
| DF | 0 | 25 | 49 | 73 | 96 | 118 |

$x(t)$　$x(t+1)$

t+3期

| | 1年 | 1格 | 2格 | 3格 | DF |
|---|---|---|---|---|---|
| t期 | 1格 | 74% | 19% | 6% | 0.75% |
| | 2格 | 10% | 66% | 18% | 6.31% |
| | 3格 | 5% | 18% | 63% | 13.33% |
| | DF | 0% | 0% | 0% | 100% |

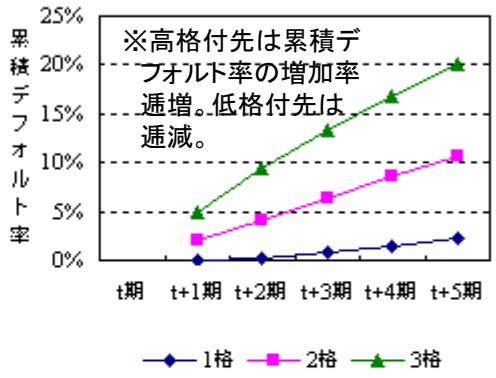
AAA

割合推移

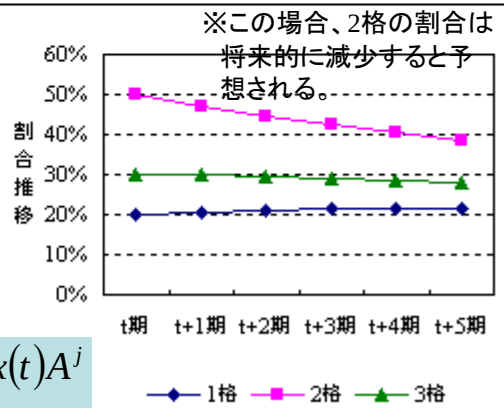
| | t期 | t+1期 | t+2期 | t+3期 | t+4期 | t+5期 |
|---|---|---|---|---|---|---|
| 1格 | 20% | 20.6% | 21.0% | 21.3% | 21.4% | 21.5% |
| 2格 | 50% | 47.0% | 44.5% | 42.3% | 40.4% | 38.7% |
| 3格 | 30% | 29.9% | 29.6% | 29.1% | 28.6% | 27.9% |
| DF | | | | | | |

t+4期

| | 1年 | 1格 | 2格 | 3格 | DF |
|---|---|---|---|---|---|




※高格付先は累積デフォルト率の増加率逓増。低格付先は逓減。




※この場合、2格の割合は将来的に減少すると予想される。

$${}^{T}x(t+j) = {}^{T}x(t+j-1)A = {}^{T}x(t+j-2)AA = \Lambda = {}^{T}x(t)A^{j}$$

$A^2$

※LDP:Low Default Portfolio

# ⑦検証にあたっての留意点

- 使用目的に合致した手法の選択
  - 格付・モデルの判別力 or デフォルト率の水準感
  - インサンプル（モデル構築時データ） or アウトサンプル（検証用データ）
  - デフォルト確率推計モデル or 格付推計モデル
- 検証は数値（いわゆる統計量）だけに頼らず視覚的なチェックを行う
- ひとつの数値に頼らずに複数の数値を使う
- 検証手法の長所・短所を抑える
- 検証手法が暗黙のうちに前提としている仮定に留意
  （デフォルト相関の有無、分布の仮定など）
- 検証結果はサンプルサイズに依存する
- 検証用データは複数年度を用いて、各々検証数値を比較する

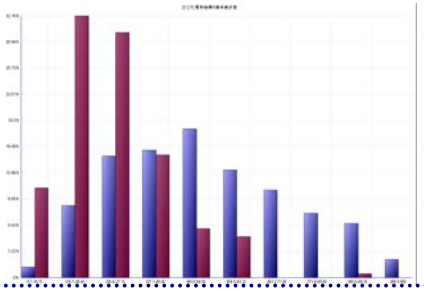
# ⑦モデルチェッカーEX

## ■ 概要

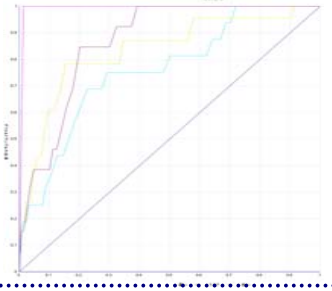
- スコアリングモデルと格付モデルのパフォーマンス検証機能に特化した統計、データベース、レポート出力ツール。
- 当ツールにより、モデル検証体制の効率的な構築が可能。

### 序列精度の検証

【デフォルト・非デフォルト分布】



【CAP曲線】



### 個別指標の有意性検定【ウィルコクソン検定（ノンパラメトリック）】

| 業種 | 非デフォルト | デフォルト | 統計量 | 閾値:Z | P値 | 有意水準 | 平均値比較 | 中央値比較 |
|---|---|---|---|---|---|---|---|---|
| 1 | 1877 | 23 | 8294 | 5.189 | 0.000% | ◎ | > | > |
| 2 | 1868 | 32 | 15702 | 4.783 | 0.000% | ◎ | > | > |
| 3 | 1864 | 36 | 16364 | 5.477 | 0.000% | ◎ | > | > |
| 4 | 1874 | 26 | 8194 | 5.947 | 0.000% | ◎ | > | > |
| 5 | 1874 | 26 | 10964 | 4.950 | 0.000% | ◎ | > | > |
| 6 | 1874 | 26 | 7548 | 6.179 | 0.000% | ◎ | > | > |
| 7 | 1868 | 30 | 9997 | 6.209 | 0.000% | ◎ | > | > |
| All | 13099 | 199 | 499613 | 15.326 | 0.000% | ◎ | > | > |

デフォルトと非デフォルト先について分布を仮定できない場合に、個別指標・変数がデフォルト説明力を持っているかを検定します。なお、両分布について正規性・等分散性を仮定できる場合は、F検定を行ないます。（F検定も当ツールで行なえます。）検定結果のCSV出力可。

### 格付ランクの順序性・有意差検定

【シャーリー・ウイリアムズ（順序性）検定】

| 信用ランク | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| | | | | | | |
| 2 | △ | | | | | |
| | -0.212 | | | | | |
| 3 | ◎ | ◎ | | | | |
| | 2.96 | 3.689 | | | | |
| 4 | ◎ | ◎ | △ | | | |
| | 2.544 | 3.284 | -0.068 | | | |
| 5 | ◎ | ◎ | ◎ | ◎ | | |
| | 7.844 | 9.929 | 6.096 | 5.796 | | |
| 6 | ◎ | ◎ | ◎ | ◎ | ◎ | |
| | 12.146 | 14.62 | 10.76 | 10.101 | 5.033 | |

【スティール・デュワス（有意差）検定】

| 信用ランク | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| | | | | | | |
| 2 | △ | | | | | |
| | 0.212 | | | | | |
| 3 | △ | ◎ | | | | |
| | 2.288 | 3.689 | | | | |
| 4 | △ | ◎ | △ | | | |
| | 2.26 | 3.662 | 0.068 | | | |
| 5 | ◎ | ◎ | ◎ | ◎ | | |
| | 5.427 | 8.713 | 5.379 | 5.796 | | |
| 6 | ◎ | ◎ | ◎ | ◎ | ◎ | |
| | 8.143 | 12.997 | 9.426 | 10.132 | 5.033 | |

多重比較法により、格付ランク間の順序性、有意差性を検証します。上図は当ツールレポート機能の出力例です。

### デフォルトフラグ付与機能

自己査定結果、延滞月（日）数、回収情報について、デフォルトフラグの付与を設定することができます。

決算月とデフォルト観測期間の関係を定義することができます。

このほか、デフォルト日からフラグを付与する機能もございます。

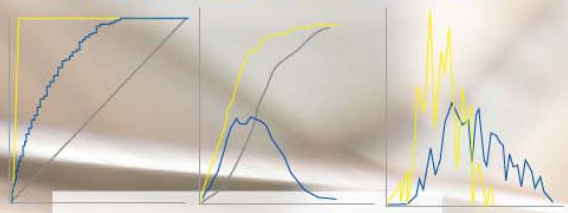
検定・検証機能だけでなく、自己査定結果、延滞月、デフォルト日等のデータにより、デフォルトフラグを付与する機能も用意しています。

その他、AR値（信頼区間含む）・KS・ダイバージェンス計算、推計PDの検証（二項検定・HL検定）、格付等の推移行列作成機能も利用することにより、モデル検証作業の効率化を図ることができます。

※本商品にはBlue Bookとよばれる解説書（いわば計算算法書）がセットされている。また、毎年、ユーザー向けセミナーを実施し、統計学的な側面からのフォローアップも万全な体制を整えている。

モデルチェッカー EX®は、銀行等の金融機関を中心に多くのユーザー様にご利用いただいております。

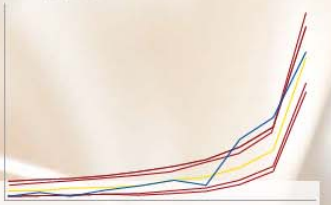
## AR, KS, Divの計測、CAP曲線の作成



デフォルトした債務者はより悪く、デフォルトしていない債務者はより良く評価しているか、つまり債務者の序列を正しく判定しているかを検証します。捕捉力の検証は、視覚的・統計的の両面からチェックします。具体的にはCAP曲線により視覚的な検証を行い、統計的側面からは、AR、KS、ダイバージェンスにより検証します。

**序列精度の検証**

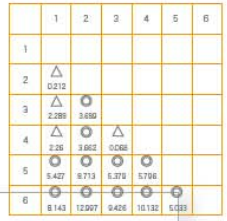
## Hosmer-Lemeshow 検定
## 二項検定



推計されたデフォルト件数（推計PD）が実際にデフォルトした件数（実績デフォルト率）を適切に予測できたかを確認します。二項検定を実施して検証を行います。さらにHosmer-Lemeshow 検定により、信用ランク別の推計デフォルト件数（推計PD）の適合度を総合的に検証します。

**推計PD・実績デフォルト率の水準感に関する検証**

## スティール・デュワス検定（有意差検定）
## シャーリー・ウィリアムズ検定（順序性検定）
## ノッチ調整の評価

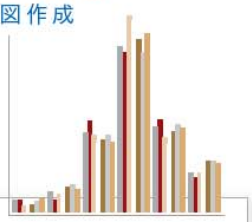


| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| 2 | △ 0.212 | | | | | |
| 3 | △ 2.289 | ○ 3.699 | | | | |
| 4 | △ 2.26 | ○ 3.662 | △ 0.068 | | | |
| 5 | ○ 5.427 | ○ 8.713 | ○ 6.379 | ○ 5.796 | | |
| 6 | ○ 8.143 | ○ 12.997 | ○ 9.426 | ○ 10.132 | ○ 5.033 | |

信用ランクについては、通常、信用ランクが良ければデフォルト率が低く、信用ランクが悪ければデフォルト率が高くなることが期待されます。信用ランク別のデフォルト率を算出し、本来期待される順序について、統計的に有意差と順序性が整合的に確保されているかどうか確認を行います。

**信用ランクの有意差・順序性の検証**

## プール評価
## 箱ひげ図の作成
## 外れ値の検定
## 正規性の検定
## 推移行列の作成
## 二次元マトリックス作成
## 基本統計量計算
## 相関係数算出
## 散布図作成



その他、推移行列作成やプール評価のための機能を搭載しています。

**その他**

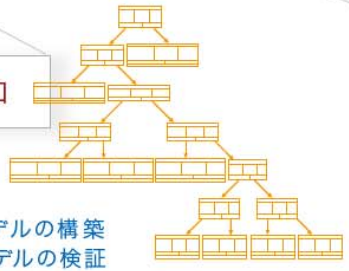
| 格付 | 企業数 | デフォルト | 累計率 | 調整 | P値 | 有意水準 | 順位相関比率 | 中央値比較 |
|---|---|---|---|---|---|---|---|---|
| 1 | 1877 | 23 | 8294 | 5.189 | 0.000% | ○ | > | > |
| 2 | 1868 | 32 | 16072 | 4.783 | 0.000% | ○ | > | > |
| 3 | 1864 | 36 | 16364 | 5.477 | 0.000% | ○ | > | > |
| 4 | 1874 | 28 | 8194 | 5.947 | 0.000% | ○ | > | > |
| 5 | 1874 | 26 | 10864 | 4.950 | 0.000% | ○ | > | > |
| 6 | 1874 | 28 | 7548 | 6.119 | 0.000% | ○ | > | > |
| 7 | 1868 | 30 | 9697 | 6.209 | 0.000% | ○ | > | > |
| AR | 13309 | 109 | 406613 | 15.326 | 0.000% | ○ | > | > |

## エラー率の計測
## ウィルコクソン検定
## F検定
## カイ二乗検定
## クラスカル・ウォリス検定

モデル内で利用されている個別変数のデフォルト説明力の検証です。具体的には、モデル構築時にデフォルトに対して説明力を十分有していた各財務指標が、構築後のアウトサンプルデータに対しても依然有効性を維持しているか検証します。分布（例：正規性）を仮定できない場合は、例えばノンパラメトリック手法であるウィルコクソン検定を用いて、非デフォルト先母集団とデフォルト先母集団の間に有意な差があるか確認します。

**個別指標の検証**

## モデル構築機能の追加



ロジットモデルの構築
ツリーモデルの構築
ハイブリッド（ロジット＆ツリー）モデルの構築
クロスバリデーションによる推計モデルの検証

モデルチェッカー EX® Ver.2.0 では、従来のモデル検証機能に加え、新たにモデル構築機能が追加されました。

モデルチェッカー EX® は、内部格付体系、大中小企業・事業性個人スコアリングモデル、リテールモデルなどの精度検証のためのツール提供だけでなく、解説書（金工研 Blue Book）やユーザーセミナー開催等によって検証手法の理論解説もします。貴社の格付体系・モデル検証体制の効率的な構築をサポートします。

# ◎Information



連絡先

株式会社　金融工学研究所
〒103－0027　東京都中央区日本橋1-4-1
日本橋一丁目ビルディング
TEL:03-3276-3440