

CRDモデル概要書

.モデル3 (法人モデル)

.モデル4 (個人事業主モデル)

- BS モデル -

平成 18 年 3 月

有限責任中間法人 CRD協会

本概要書は、信用保証協会の保証料率決定において活用いただく CRD デフォルト確率推計モデルの特徴、構造、モデルの精度 (パフォーマンス) 等に関して、概要をとりまとめたものです。

目 次

.モデル3 (法人モデル)

第1章 モデル3の特徴	-----	2
第2章 モデル3の構造	-----	2
2.1. モデル3構築に使用したデータ		
2.2. デフォルト確率算出の手順		
2.3. 1次分割モデル		
2.4. 総合デフォルト指数		
2.5. モデル業種分類		
2.6. 2次スコアリングモデルによるデフォルト確率の計算		
2.7. 複数年累積デフォルト確率の算出		
2.8. モデル3でのスコアの算出		
第3章 モデル3の検証	-----	6
3.1. デフォルト捕捉力についての検証		
3.2. 推計デフォルト確率と実績デフォルト率の一致性についての検証		

.モデル4 (個人事業主モデル) -BS モデル-

第1章 モデル4 (BS モデル)の特徴	-----	7
第2章 モデル4 (BS モデル)の構造	-----	8
2.1. モデル4構築に使用したデータ		
2.2. デフォルト確率算出の手順		
2.3. 個人事業主の財務データとデータの再整備		
2.4. 説明変数の生成		
2.5. モデル4でのスコアの算出		
第3章 モデル4 (BS モデル)の検証	-----	10
3.1. デフォルト捕捉力についての検証		
3.2. 推計デフォルト確率と実績デフォルト率の一致性についての検証		

モデル3 (法人モデル)

第1章 モデル3の特徴

モデル3は、CRD 法人モデル2を改定した「デフォルト確率推計モデル」です。モデル3の主な特徴は図表1の通りです。

図表1】モデル3の主な特徴

	主な特徴	具体的内容
1	1年のみならず、複数年(累積2年、累積3年)デフォルト確率の算出	会員の要望などを踏まえて、累積2年、累積3年のデフォルト確率も推計可能とした。
2	中業種分類でのデフォルト確率推計	細かい業種でもデフォルト確率の推計を行いたいとの要望などを踏まえて、中業種分類でのPD推計を実施した。
3	高い精度によるデフォルト企業の捕捉	長期・大量に蓄積されたCRDデータベースにより、精度の高いモデルを構築することができた。
4	理解しやすいモデル構造の構築	モデルの構造や財務指標別の寄与度について、比較的わかりやすいものとした。

第2章 モデル3の構造

2.1. モデル3構築に使用したデータ

モデル3構築に当たっては、CRDに蓄積されている図表2のとおりデータ数を使用しました。豊富なデータ量により、トータルの精度の高さはもちろんのこと、業種別にモデルを構築しても十分に高い精度を確保することができました。

図表2】モデル3の構築データ

	使用データ量
モデル3構築データ	約293万件
デフォルトデータ	約4万件

2.2. デフォルト確率算出の手順

モデル3のデフォルト確率算出に当たり、大きくは次の3つのステップをとっています(図表3参照)。

ステップ1:1次分割モデルによる企業分割

分析対象企業を、信用力を基準に、「優良な企業」と「通常の企業」に2分割。

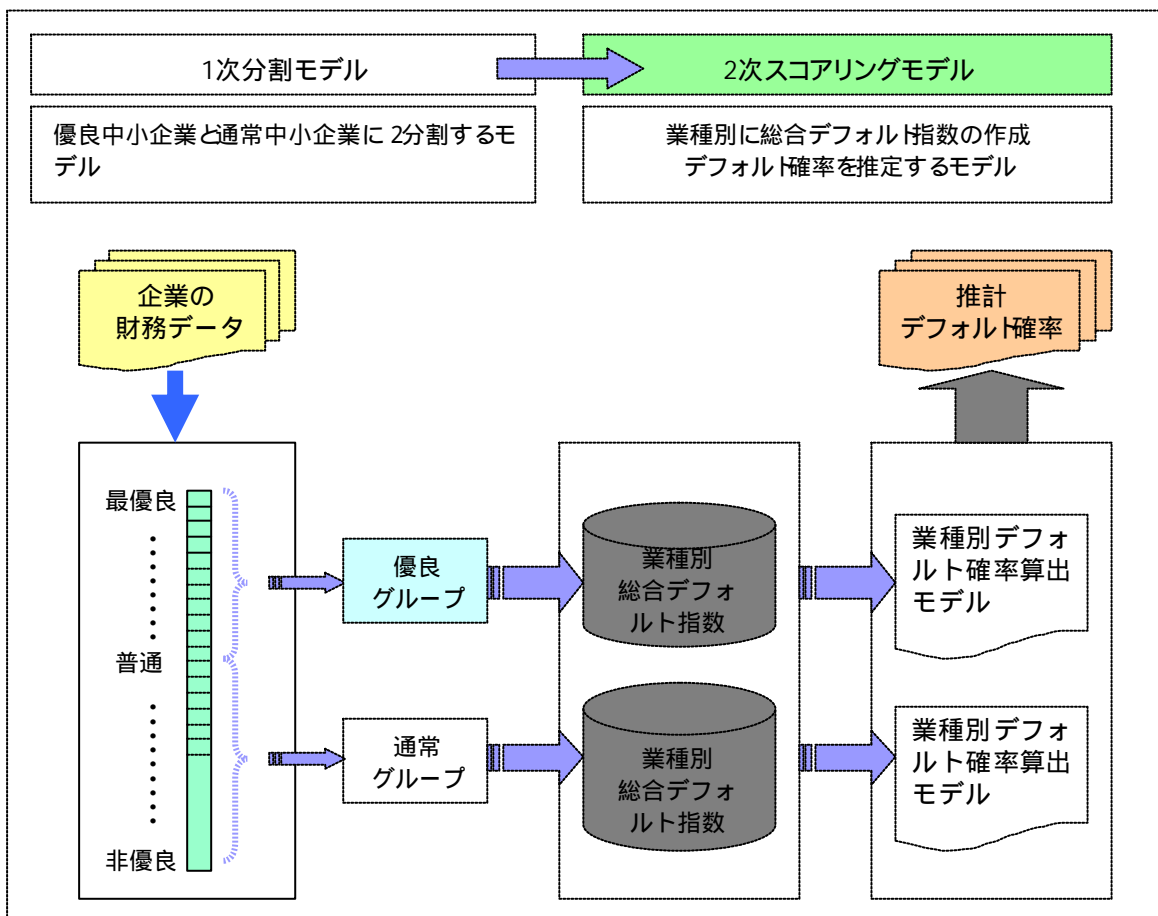
ステップ2 業種別「総合デフォルト指数」の作成

業種別に各種の財務指標を組み合わせることで、企業のデフォルト確率を説明するための変数「総合デフォルト指数」を作成。

ステップ3 2次スコアリングモデルによるデフォルト確率の計算

業種別の「総合デフォルト指数」を利用して、企業のデフォルト確率を計算。

【図表3】デフォルト確率算出の流れ



2.3. 1次分割モデル

企業が倒産に至るプロセスには、「優良な企業」と「通常の企業」で違いがあると考えられます。従って、これらの二種類の企業では、デフォルト判断に有効な財務指標も異なり、検証の結果からも異なるモデルを使うことにより、より精度の高い推計が可能になります。

そこで、企業の財務内容から、優良グループと通常グループに大まかに分割することを目的に導入したのが1次分割モデルです。この分割モデルにより、企業を2種類のタイプに分割することで、より精度の高いデフォルト確率の推定が可能になっています。

2.4. 総合デフォルト指数

企業のデフォルト確率を説明するために、業種別に各種の財務指標を組み合わせて作成する変数が「総合デフォルト指数」です。

欠損値を含むデータを処理する場合、判別関数分析など従来の統計手法では、欠損値の多い財務項目を分析の対象から外すか、欠損値を他の値で代用するなどの加工が必要となっていました。このため、データにゆがみが発生したり、利用できる指標が制限されるなど、モデルを構築する上で障害となることが多くなっていました。

そこで、本モデルでは、こういった問題点を回避するために、デフォルトを説明する変数として「総合デフォルト指数」を導入しました。モデル3では、業種別、優良・通常別に、22～28種類の財務指標を使用して「総合デフォルト指数」を算出しており、この「総合デフォルト指数」は次の様な特徴を持っています。

特定の財務指標に依存せず、様々な財務指標を評価対象に取り込むことができる

従来の統計的手法では欠損値の多い財務項目を外さざるを得ないため、せいぜい10程度の指標しか評価の対象となりえなかったのですが、本手法では、欠損値の存在を気にすることなく指標として採用することが出来るため、多くの財務指標を取り込むことが出来ます。

外れ値や欠損値による影響を受けにくい

各債務者の財務指標値を連続した数値として評価せずに、いくつかの区分に区切って区分ごとの評価を行うため、特定の財務指標が極端に良い(悪い)値であったとしても、その指標値により結果全体が大きく影響を受けることはないため、安定した評価が可能となります。

また、財務指標値の中で欠損値のないものみのデフォルト指数を平均処理して(総合デフォルト指数として)取り扱うため、欠損値がある場合でも、欠損値による影響を受けにくい評価方法です。

2.5. モデル業種分類

モデル3では、業種分類(大分類)を図表4のとおり7つのモデル業種に分けています。企業の財務データを優良・通常に判別した後、該当する業種グループに対応する2次スコアリングモデルによりデフォルト確率を算出します。

また、この7業種区分でのデフォルト確率推計に加え、さらに細かい中分類で、デフォルト確率を調整する方式を採用しています。

具体的には、中業種分類毎に推計デフォルト確率と実績デフォルト率を対比し、得られるサンプル数や水準のずれから、中分類でデフォルト率の水準調整を行った方が精度が改善される業種について、調整を行いました。

図表 4]モデル 3の業種分類

モデル業種		摘要 (大分類)
1	製造業	製造業
2	建設業	建設業
3	不動産業	不動産業
4	卸売業	卸売業
5	小売業	小売業
6	サービス業	飲食店、サービス業
7	その他業種	農林漁業、鉱業、電気・ガス 熱供給・水道業、運輸・通信業、金融・保険業、公務、分類不能の産業

2.6. 2次スコアリングモデルによるデフォルト確率の計算

2次スコアリングモデルは、個々の企業のデフォルト確率を求めるモデル(計算式)です。このモデルでは、優良・通常別に、7業種毎に「総合デフォルト指数」を用いてデフォルト確率を計算しています。

計算に当っては、 x 説明変数 (= 総合デフォルト指数)として、以下のロジスティック関数を用いてデフォルト確率 (PD) を算出します(ロジットモデル)。ただしパラメータ β 、 γ は業種モデル毎に異なる数値であり、実績のデフォルト率との整合性を高めるように統計的な手法により推定しています。

$$PD = \frac{1}{1 + \exp(-Z)} \quad Z = \beta + x^{\gamma}$$

2.7. 複数年累積デフォルト確率の算出

モデル3では、2年以内にデフォルトする確率である累積2年PD、3年以内にデフォルトする確率である累積3年PDも算出可能です。

決算書データを入力することで債務者毎に1年、累積2年、累積3年の各PDが同時に計算されます。複数年累積PDが計算可能となったことにより、長期の貸出期間中の損失額の計算等にも活用することが可能となります。

2.8. モデル3でのスコアの算出

モデル3では、企業の財務内容に応じた個社別の「デフォルト確率」が得られます。しかし、信用保証協会の提携保証のようにスコアを基準とする分野があることから、「デフォルト確率」だけでなく、「60点」などの点数による表示の必要性も考えられます。そこで、モデル3では「デフォルト確率」に一定の基準で切れ目を入れて、0点から100点に読み替えたスコアも同時に計算することとしています。

なお、デフォルト確率とスコアの対応がモデルや業種ごとに異なると、スコア間の読み替えが必要となり、運用上煩雑となるため、デフォルト確率とスコアは全業種・全モデルで共通としています。

第3章 モデル3の検証

3.1. デフォルト捕捉力についての検証

全ての評価対象債務者をデフォルト確率(評点)順に並べたとき、優秀なモデルであれば、デフォルトする可能性の高い企業ほど高いデフォルト確率(低得点)が付与されます。

この点につき、どの程度正確性が確保されているかを統計的に検討するに当たっては、通常「AR 値」と呼ばれる指標値で評価されます。「AR 値」を計算し、この数値が大きければ大きいほど、デフォルト捕捉精度が高い(モデルとして優秀)ことを示すことになります。

モデル3のデフォルト捕捉力をAR 値を用いて評価した結果、いずれの年の1年、累積2年、および累積3年デフォルト捕捉においても、従来のCRDモデルよりも優れた結果を示すことが確認できました。

3.2. 推計デフォルト確率と実績デフォルト率の一致性についての検証

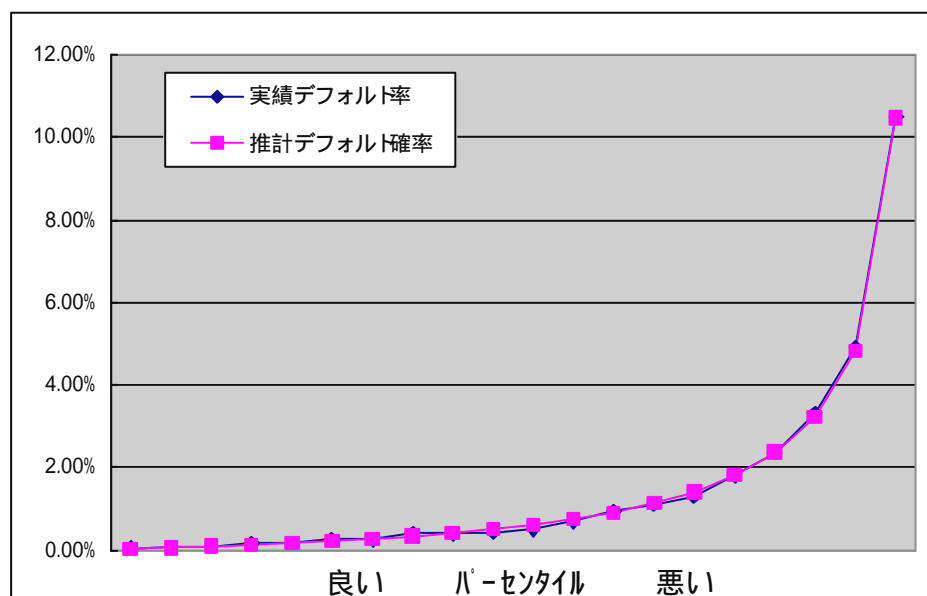
たとえ「AR 値」が高いとしても、計算されたデフォルト確率が実態を表さない数値である場合、実務上問題が発生します。

極端な例では、推計されたデフォルト確率の序列には整合性があり「AR 値が高い」としても、その推計されたデフォルト確率が、全て50%以上で100%に近い債務者も多数あるようなケースの場合、推計されたデフォルト確率はそのままでは実務に利用できません。

従って、推計されたデフォルト確率と実績のデフォルト率とがどの程度一致しているかの確認も重要となります。

図表5は、モデル3で推計された全業種にわたるデフォルト確率の結果を、デフォルト確率の高い側から低い側にサンプル数を均等に20分割し、各分割ごとの推計デフォルト確率の平均値と実績デフォルト率を対比したものです。推計デフォルト確率と実績デフォルト率は良好な一致を示しています。

図表5 推計デフォルト確率と実績デフォルト率の一致性の検証



.モデル4 (個人事業主モデル) -BS モデル-

第1章 モデル4 (BS モデル)の特徴

モデル4は、現行の CRD 個人事業主モデルを踏まえつつ新たに構築した、個人事業主のみを対象とした「デフォルト確率推計モデル」です。

モデル4では、従来モデルがほぼ BS、PL データのみからなる説明変数のモデルとなっていた点を改めて、属性 (定性) データを加えたほか、業種分類と実績デフォルトの相関に着目して業種をひとつの説明変数に加えた構成としています。

(1) 属性 (定性) データの活用

CRD では、会員の皆様のご協力により、近時個人事業主データの充実が図られています。モデル4の構築に当たっては、従来より蓄積されていた CRD データのうち、特に、これまであまり活用されてこなかった属性データ (定性データ) に注目し、モデル構築にあたって最大限活用しています。

(2) きめ細かい業種分類の活用

業種分類とデフォルトの相関性に注目してきめ細かい業種分類を反映させた変数を、デフォルト確率算出のための説明変数の1つとして使用しています。

モデル4 (BS モデル) の主要ポイントをまとめると【図表6】の通りです。

【図表6】モデル4 (BS モデル) の主要ポイント

モデル4 (BS モデル) の主要ポイント		本文中の主な記載箇所
(1)	属性 (定性) データの活用	2.4.2.属性説明変数 (定性情報) の生成 2.4.3.属性 (定性) データの反映について 3.1.デフォルト捕捉力についての検証
(2)	きめ細かい業種分類の活用	2.4.2.属性説明変数 (定性情報) の生成[後段]

以下、信用保証協会の保証料率の決定において使用される BS モデルについて、その概要を解説します。

第 2 章 モデル 4 (BS モデル) の構造

2.1. モデル 4 構築に使用したデータ

モデル 4 構築に当たっては、CRD に蓄積されている図表 7 のとおりのデータ数から抽出を行って使用しました。豊富なデータ量により、トータルの精度の高さを確保することが出来ました。

図表 7】モデル 4 の構築データ

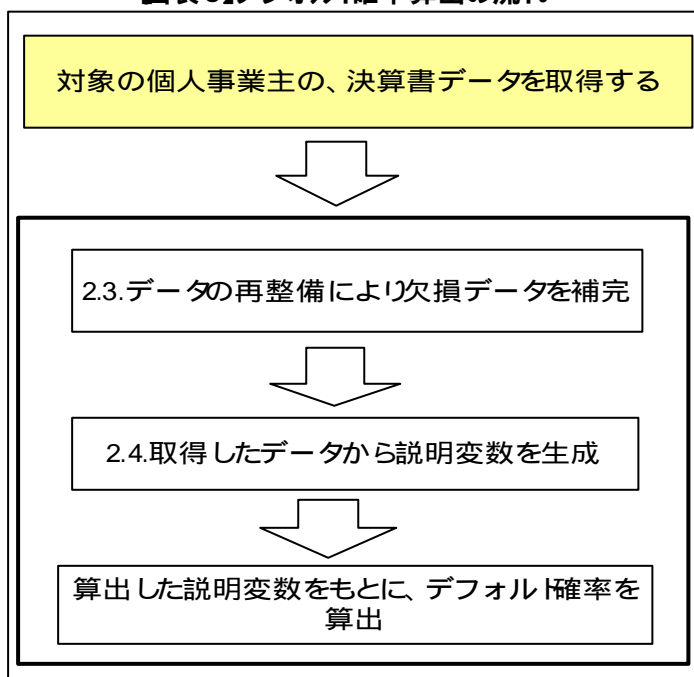
	使用データ量
モデル 4 構築データ	約 120 万件
デフォルトデータ	約 2 万件

2.2. デフォルト確率算出の手順

モデル 4 は、個人事業主ごとに 1 年内のデフォルト確率 (PD) を算出する「デフォルト確率推計モデル」であり、その推計式には、信用リスクモデルの分野で広く使用されている一般的な手法であるロジスティック回帰モデル (ロジットモデル) を採用しています。

デフォルト確率を算出する手順の流れにつきましては、**図表 8】**のとおりです。

図表 8】デフォルト確率算出の流れ



2.3. 個人事業主の財務データとデータの再整備

個人事業主の場合、BS をそもそも作成していない先が多いため、法人の場合と異なり 精度の高いモデルの構築が難しいといわれています。そういった状況を踏まえると、個人事業主データにおいて、欠損値への対応は重要課題です。そのため、モデル 4 では、以下のような合理的な手法に基づいて自

動的に再整備を行っています。

(1) BS上の「借入金」とPL上の「利子割引料」は、償還能力を算出する際において使用されることから、デフォルト確率推計の上で説明変数としての使用頻度が高いものといえます。そこで、これらの2つのデータの重要度に鑑み、もしこの一方が欠損値の場合に、片方からもう一方を推計する仕組みを組み入れています。

(2) このほか、「営業利益」のデータが欠損値の場合には、損益計算書上の前後の財務項目を参考にして、次のとおり補正を行っています。

$$\text{営業利益} = \text{売上総利益} - \text{経費計}$$

このような補正は、PL、BS あわせて合計 20 個にのぼる財務項目について行っています。

2.4. 説明変数の生成

2.4.1. 財務説明変数の生成

財務データから作成される財務説明変数については、計 272 個の財務指標について分析を行い、どの指標の有意性が高く、説明変数としてふさわしいか分析を行ったうえで選定しています。

2.4.2. 属性説明変数 (定性情報) の生成

財務データのほか、属性データを利用して説明変数を生成しています。

属性データの抽出にあたっては、CRD データとしてすでに収録されている属性情報をもとに生成した計 27 項目の指標を分析しました。その結果、デフォルト事象との関係性が確認できた属性データについて、モデルへの反映が可能であるとの結論が得られました。

ここでデフォルトとの関係性が確認された CRD 属性データ、ならびに審査実務において意識される利益状況と売上規模を加えた属性情報を得点化して、説明変数としての「総合定性得点」の生成を行いました。

生成した「総合定性得点」について、財務説明変数とともに、デフォルト確率算出のための説明変数の一つとして使用します。

なお、属性 (定性) データのうち、業種分類については、前述のとおり、従来のモデル以上にきめ細かい対応を行い、小分類業種に基づいて配点を行い、「総合定性得点」に反映させています。

2.4.3. 属性 (定性) データの反映について

入力されている財務データおよび属性 (定性) データから説明変数を生成した上で、ロジット関数を用いてデフォルト確率を算出します。

モデル 4では、デフォルト確率を推計する際に、属性 (定性) データを加味するかどうかで、2種類の結果を算出しています。属性 (定性) データを加味した場合を「総合モデル」、加味せずに財務データのみから算出する場合を「財務モデル」と呼んでいます。この2種類の結果は、両方とも出力されます。

2.5. モデル4でのスコアの算出

モデル4では、企業の財務内容に応じた個人事業主ごとの「推定デフォルト確率 (PD)」が得られます。しかし、企業の優劣比較の理解しやすさの観点等から、「デフォルト確率」だけでなく、「60点」などの点数による表示の必要性も考えられます。そこで、モデル4では「デフォルト確率」に一定の基準で切れ目を入れて、0点から100点に読み替えた「スコア」も同時に算出することとしています。

なお、デフォルト確率とスコアの対応がモデルや業種ごとに異なると、スコア間の読み替えが必要となり、運用上煩雑となるため、デフォルト確率とスコアは全業種 全モデルで共通としています。

第3章 モデル4 (BSモデル)の検証

3.1. デフォルト捕捉力についての検証

デフォルト捕捉力について、モデル3と同様、「AR値」を用いて検証しました。

BSデータがある場合のデフォルト捕捉力をAR値を用いて評価した結果、モデル4の捕捉力が従来のCRDモデルを上回っていることが確認できました。

さらに、CRDの属性(定性)データを加味して判断した場合(総合モデル)については、財務データのみで判断した場合(財務モデル)よりもデフォルト捕捉力が上回っており、属性データの活用が一定程度有効であることを示しています。

3.2. 推計デフォルト確率と実績デフォルト率の一致性についての検証

AR値が高いモデルであっても、算出されたデフォルト確率の実態から大幅に乖離した数値である場合は精度が高いとは言えません。従って、推計されたデフォルト確率と実績のデフォルト率とがどの程度一致しているかを検証することも重要となります。

【図表9】は、モデル4で推計されたデフォルト確率の結果を、デフォルト確率の低い側から高い側にサンプル数を均等に10分割し、各分割ごとの推計デフォルト確率の平均値と実績デフォルト率を対比したものです。これをみると、推計デフォルト確率と実績デフォルト率は良い一致を示していることがわかります。

図表9 推計デフォルト確率と実績デフォルト率の一致性の検証

