

# デフォルト企業の正常復帰に関する 要因分析と正常復帰確率推定モデル

田上 悠太<sup>1,2</sup>・山下 智志<sup>3</sup>

(受付 2014 年 5 月 15 日 ; 改訂 9 月 23 日 ; 採択 11 月 18 日)

## 要 旨

信用リスク研究は、非デフォルト債権に対するデフォルト確率の研究が中心に行われており、デフォルト債権の研究は限られている。デフォルト債権の研究においては、正常復帰確率の分析が重要で、本研究ではデフォルト債権の正常復帰確率推定モデルを作成し、正常復帰に影響を与える要因について分析した。2007 年 3 月から 2012 年 3 月までのある地方銀行の法人向け貸付債権に対して、担保情報、保証情報、財務情報、時間情報を説明変数に用いたロジットモデルを作成し、正常復帰に影響を与える要因について分析した。また、本研究では Box-Cox 変換の負の値への拡張である Yeo-Johnson 変換 (YJ 変換) をモデルの説明力向上のために用いた。AIC, AUC, Hosmer-Lemeshow 統計量をもとにモデル評価を行い、モデルの説明変数として、YJ 変換済みの財務変数、財務評点、時間情報カテゴリー変数などを選択した。また、YJ 変換を施したモデルと施していないモデルを AIC, AUC, Hosmer-Lemeshow 統計量に基づいて比較を行い、YJ 変換によるモデルの説明力の向上を確認した。

キーワード：信用リスク、正常復帰確率、LGD。

## 1. はじめに

### 1.1 正常復帰分析の背景と重要性

これまでの信用リスク研究は正常状態の債権の分析が中心に行われてきた。正常状態の債権とは、債務に対して約束通りの利息の支払いが滞りなく行われている債権を言う。正常状態の債権の分析では、デフォルト確率推定の研究が中心に行われ、ロジットモデル (山下・安道, 2006; 森平・岡崎, 2009), ハザードモデル (Lane et al., 1986; 乾・室町, 2000), ニューラルネットワークモデル (Atiya, 2001), 決定木モデル (Sung et al., 1999; 白田, 1999), 構造モデル (Kijima and Suzuki, 2001; Merton, 1974), 誘導モデル (Duffie and Singleton, 1999) など様々なモデルが提案されている。また、近年はデフォルト時損失率 (LGD) (デフォルトした債権のデフォルト時点の与信額のうち生じた損失額の割合を表した数値のこと) の研究も盛んに行われており、線形回帰モデル (Caselli et al., 2008; Citron et al., 2003; Grunert and Weber, 2009), ロジットモデル (Caselli et al., 2008; 三浦 他, 2010), log-log モデル (Bastos, 2010; Dermine and Neto de Carvalho, 2006), 回帰木モデル (Bellotti and Crook, 2012; 川田・山下, 2012; 森平, 2009; Gürtler and Hibbeln,

<sup>1</sup> 総合研究大学院大学 複合科学研究科統計科学専攻：〒190-8562 東京都立川市緑町 10-3

<sup>2</sup> 日本学術振興会特別研究員 DC：〒102-0083 東京都千代田区麹町 5-3-1 麹町ビジネスセンター

<sup>3</sup> 統計数理研究所：〒190-8562 東京都立川市緑町 10-3

2011)など様々なモデルが提案されている。一方、デフォルト状態の債権に関しては、一般に公表され利用可能なデータが少ないため研究がほとんど進んでいない。

以下で説明するように、デフォルト状態の債権の分析においては、正常復帰の分析が重要になる。正常復帰(詳細な定義は2章で行う)とはデフォルト状態の債権が、正常状態に戻ることを言う。本研究で用いたデータでは968件の債権がデフォルト状態に陥り、そのうち144件が正常復帰した。デフォルト状態の債権の分析において、正常復帰は、(1)詳細なLGD分析、(2)銀行の貸出意思決定、(3)デフォルト状態の債権の現在価値の算出、に必要とされている。それぞれについて説明する。

(1) 詳細なLGD分析に関して、正常復帰する債権からも損失が生じることがあるが、これまでのLGD研究の中で正常復帰する債権はLGDが小さいことが示されている(Gürtler and Hibbeln, 2011; 川田・山下, 2012等)。実際、本研究で用いたデータにおいては、144件の正常復帰をした債権の全てが損失を生じていない。このように、より精緻なLGD研究のために正常復帰の分析が必要とされている。

(2) 銀行の貸出意思決定に関して、通常、デフォルト状態の債権は利息を払わないが、正常復帰する場合、今後もその債権から利息収入が期待できる。そのため、正常復帰する可能性が高い債権の債務者に対して追加融資などの経営支援することには合理性がある。このように、銀行の貸付意思決定において、正常復帰の分析は重要である。

(3) デフォルト状態の債権の現在価値の算出に関して、今後強制適用が予定されるIAS(国際会計基準)のIAS39(国際会計基準の一部で、金融資産、金融負債および非金融商品の売買契約についての認識および測定に関して定めている)では、デフォルト状態の債権の適切な現在価値評価が要求される。債権の適切な現在価値評価のためにはその債権からのキャッシュフローの予測が必要とされる。デフォルト債権のキャッシュフロー予測のためにはLGDおよび利息の予測が必要であり、上述のとおり、LGDおよび利息の予測のためには、正常復帰の分析が必要となる。

このように、デフォルト状態の債権の分析において正常復帰の分析が必要とされている。

## 1.2 正常復帰分析に関する先行研究

これまで、正常復帰分析は主にLGD推定モデルの中で正常復帰確率推定モデルを作成することを通して行われてきた。

正常復帰確率を推定したモデルには、(1)状態遷移確率推定モデル(青沼・市川, 2008; 三浦 他, 2010)、(2)デフォルト終了判定モデル(Gürtler and Hibbeln, 2011; 川田・山下, 2012)、の大きく2種類が存在する。上述の通り(1)状態遷移推定モデル、(2)デフォルト終了判定モデルは、それぞれLGD推定モデルの中で正常復帰確率を推定しているが、以下ではLGD推定モデル全体ではなく、その中で用いられている正常復帰確率推定モデル(また、それに該当する部分)について説明する。

### 1.2.1 状態遷移確率モデル

状態遷移確率モデルでは各債権を(1)完済、(2)正常状態、(3)非吸収デフォルト状態、(4)吸収デフォルト状態、の4つに分類し、各状態の債権が、次の時点において各状態へ遷移する確率を推定している。それぞれの状態は表1のようになっている。(1)完済と(4)吸収デフォルト状態は銀行との取引終了状態であり、一度到達したら抜け出すことのできない吸収状態となっている。そのため、状態遷移確率過程は吸収マルコフ連鎖となっている。そして、状態遷移確率行列のうち、(3)非吸収デフォルト状態から(1)完済または(2)正常状態への遷移確率が正常復帰確率を表している。

表 1. 銀行内部格付の状態の分類.

	非デフォルト状態	デフォルト状態
取引中	(2) 正常状態	(3) 非吸収デフォルト状態
取引終了	(1) 取引終了	(4) 吸収デフォルト状態

状態遷移確率モデルでは、状態遷移確率に斉時性を仮定し、複数年の実績状態遷移数をもとに、遷移確率の平均を状態遷移確率の推定値としている。そのため、(3)非吸収デフォルト状態の全ての債権は全て同一の正常復帰確率を持ち、個別債権に対する LGD 分析、貸出意思決定、現在価値評価を行うために用いるには、不十分である。

### 1.2.2 デフォルト終了判定モデル

デフォルト終了とは、債権のデフォルト状態が終了することである。デフォルト状態は、(1)正常復帰する、(2)デフォルト状態のまま取引終了する、のいずれかで終了する。デフォルト終了判定モデルは、正常状態の債権がデフォルト状態に陥った場合に、そのデフォルト終了が(1)正常復帰によるものか、デフォルト状態のまま取引終了することなのかを、個別債権の正常状態時点(デフォルト状態に陥る以前)でのデータを用いて推定している。デフォルト終了判定モデルは正常状態の債権の LGD 推定モデルの一部として用いられている。そのため、モデルの説明変数として債権の正常状態時点でのデータを用いている。

Gürtler and Hibbeln(2011)では、詳細は記述されていないが、債務者の属性データ、担保の属性データ、債権の属性データ、を用いてデフォルト終了確率を推定している。川田・山下(2012)では、CRITS 標準スコア(CRITS 標準スコアは地方銀行から提出された大量のサンプルデータに基づき、地方銀行協会と(株)金融工学研究所が共同で開発した統計的スコアリングモデルのアウトプットで、各債務者を信用力に応じて1点から100点で評価するスコアである。全国地方銀行協会の全会員銀行に提供されており、会員銀行において債務者の信用力を評価するために広く用いられている(社団法人全国地方銀行協会信用リスク管理高度化支援室, 2005)。モデルの詳細に関しては全国地方銀行協会会員以外には非公開となっている。エクスポージャー、担保情報、保証情報を用いて分析を行っており、また、2004年9月期と2009年3月期において正常復帰確率が高いことが示された。これは条件緩和措置の影響であると指摘している。

デフォルト状態に陥る以前と以後では債権の状態は著しく異なる。そのため、正常状態時点での説明変数を用いているデフォルト終了判定モデルは、デフォルト状態の債権の分析としては十分に機能しない。また、Gürtler and Hibbeln(2011)では用いられている説明変数が言及されておらず、川田・山下(2012)では説明力を有すると考えられる個別債権の財務情報を用いていないという問題もある。

### 1.3 本稿の目的

上記のように、先行研究は正常状態でのデフォルト状態の債権の正常復帰の分析としては不十分なことがわかった。

そこで、本研究ではデフォルト状態の債権の個別データ(財務、担保保証、時間情報)を用いて、その債権が次の時点(半年後)に正常復帰する確率をロジットモデルで推定し、正常復帰に与える要因を分析する。

以下では、2章で本研究で必要になるデフォルト、正常復帰、債務者区分の定義をし、また、本研究で使用したデータについて説明する。3章で、本研究の分析方法とその分析結果を示す。そして、4章でまとめと研究の今後の課題について説明する。

## 2. デフォルト・正常復帰の定義と使用データ

ここでは、デフォルトと正常復帰、デフォルト状態と非デフォルト状態を定義する。そのために、まずこれらの定義の基礎となる債務者区分について説明する。続いて、本研究で用いたデータと Yeo-Johnson 変換について説明する。

### 2.1 債務者区分の定義

債務者区分とは「金融検査マニュアル(預金等受入金融機関に係る検査マニュアル)」の自己査定(別紙 1) (平成 25 年 8 月に改定されたもの。 <http://www.fsa.go.jp/manual/manualj/manual-yokin/14.pdf>) で定義される債務者の区分であり、債務者は財務状態などの信用力によって、正常先、要注意先、要管理先、破綻懸念先、実質破綻先、破綻先、の 6 区分のいずれかに区分される。厳密には要管理先は要注意先の中に含まれる区分であるが、本稿では要管理先を要注意先から独立した一つの区分として扱う。

なお、本研究では各債権は債務者単位で集計されているため、債権と債務者(企業)を同じ意味で用いることがある。

### 2.2 デフォルト、正常復帰、デフォルト状態、非デフォルト状態の定義

本研究ではデフォルトと正常復帰を債務者区分の遷移により定義する。日本銀行金融機構局(2005)を参考に、ある債権の債務者である企業の債務者区分が要注意先以上(正常先、要注意先)から要管理先以下(要管理先、破綻懸念先、実質破綻先、破綻先)に遷移することをデフォルト、その逆を正常復帰と定義する。正常復帰とは債権の債務者である企業の債務者区分が正常先になることではなく、要注意先以上になることである。なお、今回の分析ではデータの制約、分析の簡素化のために、金融庁告示第十九号(銀行法第十四条の二の規定に基づき、銀行がその保有する資産等に照らし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)のデフォルトの定義は用いなかった。

債務者の債権者区分が要管理先以下の債権をデフォルト状態の債権といい、債務者の債権者区分が要注意先以上の債権を非デフォルト状態の債権という。

### 2.3 データの概要

分析に用いたデータは銀行貸付債権で各債権に対して一つ以上のレコードが記録されているパネルデータである。各レコードには、(1)正常復帰フラグ、(2)顧客番号、(3)債務者区分、(4)デフォルト状態、非デフォルト状態、(5)基準日、(6)説明変数(担保情報など)、が含まれている。(1)の正常復帰フラグはあるレコードがデフォルト状態にあり、次の基準日に非デフォルト状態になるときに 1 で、それ以外の場合は 0 のフラグである。

表 2 は、分析に用いたデータの例である。

### 2.4 分析単位と分析対象

本研究では、分析単位を債権単位ではなくレコード単位とする。表 2 は債務者 A のデータであるが、このデータにはレコードが 12 個あり、各レコードを独立な分析仮想単位として扱う。また、本研究ではデフォルト状態の債権のレコードのみを分析対象とするため、表 2 においては、レコード 3 から 5 と 9 から 12 のみが分析対象となる。

### 2.5 データの基本情報

本研究で用いたデータは地方銀行の法人向け貸付債権データで、データは 2007 年 3 月から

表 2. 分析データの例.

正常復帰フラグ	顧客番号	債務者区分 (格付)	状態	基準日	説明変数	
1	0	A	正常先	非デフォルト	2005年3月	$X_{A,2005}$ 年3月
2	0	A	正常先	非デフォルト	2005年9月	$X_{A,2005}$ 年9月
3	0	A	要管理先	デフォルト	2006年3月	$X_{A,2006}$ 年3月
4	0	A	要管理先	デフォルト	2006年9月	$X_{A,2006}$ 年9月
5	1	A	要管理先	デフォルト	2007年3月	$X_{A,2007}$ 年3月
6	0	A	要注意先	非デフォルト	2007年9月	$X_{A,2007}$ 年9月
7	0	A	正常先	非デフォルト	2008年3月	$X_{A,2008}$ 年3月
8	0	A	正常先	非デフォルト	2008年9月	$X_{A,2008}$ 年9月
9	0	A	要管理先	デフォルト	2009年3月	$X_{A,2009}$ 年3月
10	0	A	破綻懸念先	デフォルト	2009年9月	$X_{A,2009}$ 年9月
11	0	A	実質破綻先	デフォルト	2010年3月	$X_{A,2010}$ 年3月
12	0	A	破綻先	デフォルト	2010年9月	$X_{A,2010}$ 年9月

2012年3月までの半年ごとに記録されている。

分析に用いたデフォルト状態の債権のレコード数は3312で、デフォルト発生件数は968、正常復帰件数は144である。なお、分析に用いたデータには2012年9月時点(データの最終時点)でデフォルト状態の債権も含まれている。これらの債権は2012年9月以降に正常復帰する可能性がある。そのため、実際には968件発生したデフォルトに対して、正常復帰件数は144件以上である。

説明変数として、担保情報、保証情報、財務情報を用いた。また、データの基準日情報(データが記録された基準日のこと)、デフォルトからの経過時間情報をカテゴリカル変数として用いた。説明変数の一覧を表3に掲載した。さらに以下で説明する Yeo-Johnson 変換をそれぞれの変数に施した変数も説明変数として使用した。

## 2.6 Yeo-Johnson 変換

本稿では、モデルの説明力の向上のために Yeo and Johnson (2000) により導入された Yeo-Johnson 変換(以下、YJ 変換と呼ぶ)を用いた。

YJ 変換は式(2.1)で定義される。

$$(2.1) \quad \psi(\lambda, x) = \begin{cases} \{(x+1)^\lambda - 1\}/\lambda & (x \geq 0, \lambda \neq 0) \\ \log(x+1) & (x \geq 0, \lambda = 0) \\ -\{(x+1)^{2-\lambda} - 1\}/(2-\lambda) & (x < 0, \lambda \neq 2) \\ -\log(-x+1) & (x < 0, \lambda = 2) \end{cases}$$

ただし、 $\lambda \in \mathbb{R}$ 。

YJ 変換は式(2.2)で定義される Box-Cox 変換を負の値(本研究で用いた説明変数では、売上総利益/資産合計、経常利益/資産合計、営業利益/資産合計などの変数が負の値を取る)に拡張した変換である。 $x$  が正の時には  $(x+1)$  の Box-Cox 変換と同等になり、 $x$  が負の時には  $\lambda$  が  $(2-\lambda)$  の  $(-x+1)$  の Box-Cox 変換と同等になる。

$$(2.2) \quad \text{BoxCox}(x) = \begin{cases} \frac{1}{\lambda}(x^\lambda - 1) & (\lambda \neq 0) \\ \log(x) & (\lambda = 0) \end{cases}$$

また、YJ 変換は式(2.3)で定義される neglog 変換の拡張でもある。 $x$  が正の時には neglog 変

表 3. 説明変数の一覧.

デフォルト判定	負債純資産合計/資産合計	前払費用/与信額	売上事業キャッシュフロー率
基準日	売上高/資産合計	仮払金/与信額	総資本回転率
日銀業種	売上総利益/資産合計	流動資産計/与信額	売上債権回転日数
地銀業種	営業利益/資産合計	土地/与信額	棚卸資産回転日数
地域	受取利息配当金/資産合計	有形固定資産計/与信額	売上高有形固定資産率
与信額	支払利息割引料/資産合計	長期貸付金/与信額	買入債務回転日数
貸出金額/与信額	経常利益/資産合計	固定資産計/与信額	流動比率
支承/与信額	税引前当期利益/資産合計	繰延資産計/与信額	当座比率
外為/与信額	当期利益/資産合計	支払手形/与信額	支払準備率
保全合計/与信額	割引手形/資産合計	買掛金/与信額	現預金比率
担保預金/与信額	裏書譲渡手形/資産合計	短期借入金/与信額	金利対現金預金比率
担保有価証券/与信額	減価償却実施額/資産合計	未払金/与信額	自己資本比率
担保不動産/与信額	期末役員従業員数	未払費用/与信額	純資産倍率
担保担保手形/与信額	預貸率	前受金/与信額	固定長期適合率
担保商手/与信額	流動預貸率	未成工事受入金/与信額	固定比率
保証保証協会/与信額	資金利益額/資産合計	前受収益/与信額	借入金依存度
担保譲渡担保/与信額	直接費差引後利益/資産合計	流動負債計/与信額	デッドキャパシティーレシオ
保証機関保証/与信額	手数料外為利益/資産合計	社債/与信額	預借率
保証私的保証/与信額	取引年数	長期借入金/与信額	借入金月商倍率
決算年月	融資シェア	固定負債計/与信額	売上高支払利息割引料率
モデル業種	法人メイン化区分	負債合計/与信額	有利子負債比率
CRITS 標準スコア	代表者取引	資本金/与信額	現金預金対利子割引料率
当行シェア	担保余力	純資産合計/与信額	支払利息割引料対総利益率
現金預金/資産合計	取引格付	負債純資産合計/与信額	事業キャッシュフロー有利子負債比率
受取手形/資産合計	預貸率評点	売上高/与信額	インタレストカバレッジ
売掛金/資産合計	流動預貸率評点	売上総利益/与信額	事業キャッシュフロー金利負担率
有価証券/資産合計	資金利益額評点	営業利益/与信額	減価償却率
当座資産計/資産合計	直接費差引後利益評点	受取利息配当金/与信額	売上高減価償却率
棚卸資産計/資産合計	手数料外為利益評点	支払利息割引料/与信額	売上高
前渡金/資産合計	取引年数評点	経常利益/与信額	資産合計土地比率
未収入金/資産合計	融資シェア評点	税引前当期利益/与信額	資産合計有形固定資産比率
前払費用/資産合計	法人メイン化評点	当期利益/与信額	資産合計前払費用比率
仮払金/資産合計	代表者取引評点	割引手形/与信額	資産合計流動資産比率
流動資産計/資産合計	担保余力評点	裏書譲渡手形/与信額	資産合計固定資産比率
土地/資産合計	毀損額	減価償却実施額/与信額	資産合計支払手形比率
有形固定資産計/資産合計	経常運転資金額/資産合計	資金利益額/与信額	資産合計受取手形比率
長期貸付金/資産合計	長期運転資金額/資産合計	直接費差引後利益/与信額	資産合計当座資産比率
固定資産計/資産合計	一時運転資金額/資産合計	手数料外為利益/与信額	資産合計棚卸資産比率
繰延資産計/資産合計	その他運転資金額/資産合計	合計/与信額	基準日
資産合計	土地建物購入資金額/資産合計	経常運転資金額/与信額	デフォルトからの経過時間
支払手形/資産合計	事業設備資金額/資産合計	長期運転資金額/与信額	
買掛金/資産合計	その他資金額/資産合計	一時運転資金額/与信額	
短期借入金/資産合計	毀損合計	その他運転資金額/与信額	
未払金/資産合計	累積毀損額	土地建物購入資金額/与信額	
未払費用/資産合計	デフォルト回数	事業設備資金額/与信額	
前受金/資産合計	デフォルトからの経過時間	その他資金額/与信額	
未成工事受入金/資産合計	担保余力割る/与信額	与信額/EAD	
前受収益/資産合計	現金預金/与信額	総資本営業利益率	
流動負債計/資産合計	受取手形/与信額	総資本経常利益率	
社債/資産合計	売掛金/与信額	ROA	
長期借入金/資産合計	有価証券/与信額	ROE	
固定負債計/資産合計	当座資産計/与信額	売上高総利益率	
負債合計/資産合計	棚卸資産計/与信額	売上高営業利益率	
資本金/資産合計	前渡金/与信額	売上高経常利益率	
純資産合計/資産合計	未収入金/与信額	売上高当期利益率	

換と  $\lambda = 0$  の YJ 変換は同等で,  $x$  が負の時には  $\text{neglog}$  変換と  $\lambda = 2$  の YJ 変換は同等になる.

$$(2.3) \quad \text{neglog}(x) = \begin{cases} \log(x+1) & (x \geq 0) \\ -\log(-x+1) & (x < 0) \end{cases}$$

本研究では、説明変数の説明力を向上させるために、YJ変換を施していない説明変数に加えてYJ変換を施した変数も説明変数として分析に用いた。YJ変換を施すことで、データの分布の歪みを補正し、異常値を丸め込むことでモデルの説明力の向上が期待できる。YJ変換の $\lambda$ はYeo and Johnson (2000)に従って、変換後変数 $\Psi(\lambda, x)$ が正規分布に従っていると仮定し、式(2.4)の対数尤度が最大になる $\lambda$ (ならびに平均 $\mu$ と標準偏差 $\sigma$ )を用いた。

$$(2.4) \quad l_n(\lambda, \mu, \sigma|x) = -\frac{1}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n \Psi(\lambda, x_i) - \mu^2 \\ + (\lambda - 1) \sum_{i=1}^n \text{sign}(x_i) \log(|x_i| + 1)$$

ただし、 $\text{sign}(x)$ は、 $x \geq 0$ の時1、 $x < 0$ の時-1となる関数とする。

なお、本研究はYJ変換を財務変数に対して適用した最初の例であるが、本研究以外のYJ変換の適用例として、Schwalm et al.(2006)、Wang et al.(2009)、Wang and Robertson(2011)などがある。これらのいずれの研究においても本研究と同様にYJ変換を変数の分布を正規分布に近づけるために用いている。実際に財務変数に対してYJ変換を適用した際の $\lambda$ の値の結果を表4にまとめた。なお、表4の $\lambda$ が空欄になっている変数はフラグ変数、または、カテゴリ変数であり、これらの変数にはYJ変換は適用していない。

### 3. 分析結果

#### 3.1 正常復帰確率推定モデル

正常復帰確率推定モデルには、以下の式(3.1)のロジットモデルを用いた。係数の推定には最尤法を用いた。

$$(3.1) \quad P = \frac{1}{1 + \exp(Z)} \\ Z = \alpha + \sum_i \beta_i x_i$$

式(3.1)において、 $P$ は推定正常復帰確率、 $x_i$ はモデルに用いる説明変数、 $\alpha$ 、 $\beta_i$ はデータから推定される係数である。

モデル選択は、AIC、AUC、Hosmer-Lemeshow 統計量、 $z$ 値を用いて行った。AICは絶対的な基準がない、また、以下で説明するが、AUCは順序性の指標であり正常復帰確率の水準を評価できない、Hosmer-Lemeshow 統計量はモデルの適合度指標であるがその値に絶対的な基準がない、さらに、 $z$ 値は各説明変数に対する指標でありモデル全体の評価ができない、などの欠点を持っている。そのため、いずれか一つの指標のみを基準にモデル選択を行うことは妥当ではない。そこで、本研究ではこれらの4つの指標を総合的に用いてモデル選択を行った。以下でAUC、Hosmer-Lemeshow 統計量について説明する。

#### 3.2 AUCの定義

AUCは正常復帰レコードの集合を $S_R$ 、非正常復帰レコードの集合を $S_{NR}$ 、正常復帰レコード数を $N_R$ 、非正常復帰レコード数を $N_{NR}$ 、 $P_i$ をレコード $i$ の推定正常復帰確率とし、 $I(\cdot)$ をヘビサイド関数( $I(x)$ に対して、 $x > 0$ の時1をとり、それ以外の際は0を取る関数)とする時、

$$(3.2) \quad AUC = \frac{1}{N_R N_{NR}} \sum_{i \in S_R} \sum_{j \in S_{NR}} I(P_i - P_j)$$

表 4. YJ 変換の  $\lambda$ .

デフォルト判定	6.74	負債純資産合計/資産合計	0.77	前払費用/与信額	-6.82	売上事業キャッシュフロー率	1.11
基準日		売上高/資産合計	-0.76	仮払金/与信額	-8.88	総資本回転率	-0.76
日銀業種		売上総利益/資産合計	0.53	流動資産計/与信額	-0.50	売上債権回転日数	1.12
地銀業種		営業利益/資産合計	3.23	土地/与信額	-1.18	棚卸資産回転日数	0.62
地域		受取利息配当金/資産合計	-534.03	有形固定資産計/与信額	-0.59	売上高有形固定資産率	-0.58
与信額	0.04	支払利息割引料/資産合計	-23.76	長期貸付金/与信額	-5.37	買入債務回転日数	0.67
貸出金額/与信額	303.92	経常利益/資産合計	2.59	固定資産計/与信額	-0.50	流動比率	-0.83
支承/与信額	-246.63	税引前当期利益/資産合計	2.19	繰延資産計/与信額	-14.51	当座比率	-1.32
外為/与信額	-6467.01	当期利益/資産合計	2.21	支払手形/与信額	-3.56	支払準備率	-2.77
保全合計/与信額	-0.85	割引手形/資産合計	-58.79	買掛金/与信額	-1.64	現預金比率	-7.80
担保預金/与信額	-203.75	裏書譲渡手形/資産合計	-382.68	短期借入金/与信額	-0.78	金利対現金預金比率	0.43
担保有価証券/与信額	-371.40	減価償却実施額/資産合計	-26.01	未払金/与信額	-2.75	自己資本比率	2.80
担保不動産/与信額	-5.75	期末役員従業員数	-0.31	未払費用/与信額	-6.13	純資産倍率	1.04
担保担保手形/与信額	-1495.64	預貸率	-0.51	前受金/与信額	-15.34	固定長期適合率	1.12
担保商手/与信額	-105.65	流動預貸率	-0.64	未成工事受入金/与信額	-26.35	固定比率	1.05
保証保証協会/与信額	0.65	資金利益額/資産合計	-46.62	前受収益/与信額	-434.59	借入金依存度	-1.39
担保譲渡担保/与信額	-5586.31	直接費差引後利益/資産合計	-0.61	流動負債計/与信額	-0.52	デッドキャパシティーレシオ	-0.87
保証機関保証/与信額	-516.87	手数料外為利益/資産合計	-862.14	社債/与信額	-11.57	預信率	-6.61
保証私的保証/与信額	-1.04	取引年数	0.31	長期借入金/与信額	-0.48	借入金月商倍率	0.72
決算年月		融資シェア	0.34	固定負債計/与信額	-0.46	売上高支払利息割引率	-7.88
モデル業種		法人メイン化区分		負債合計/与信額	-0.47	有利子負債利率	-19.20
CRITS 標準スコア	0.57	代表者取引	-0.17	資本金/与信額	-1.36	現金預金対利子割引利率	0.21
当行シェア	0.39	担保余力	1.10	純資産合計/与信額	0.94	支払利息割引料対総利益率	0.71
現金預金/資産合計	-8.94	取引格付	2.53	負債純資産合計/与信額	-0.41	事業キャッシュフロー有利子負債	1.81
受取手形/資産合計	-20.11	預貸率評点	0.24	売上高/与信額	-0.31	インタレストカバレッジ	1.29
売掛金/資産合計	-4.70	流動預貸率評点	0.21	売上総利益/与信額	0.42	事業キャッシュフロー金利負担率	0.96
有価証券/資産合計	-205.17	資金利益額評点	0.13	営業利益/与信額	0.95	減価償却率	0.34
当座資産計/資産合計	-2.62	直接費差引後利益評点	1.68	受取利息配当金/与信額	-21.95	売上高減価償却率	-17.97
棚卸資産計/資産合計	-5.12	手数料外為利益評点	0.51	支払利息割引料/与信額	-3.68	売上高	0.74
前渡金/資産合計	-246.46	取引年数評点	-0.34	経常利益/与信額	1.00	資産合計土地比率	-5.89
未収入金/資産合計	-34.38	融資シェア評点	0.56	税引前当期利益/与信額	1.00	資産合計有形固定資産比率	-1.02
前払費用/資産合計	-53.11	法人メイン化評点	-1.21	当期利益/与信額	1.03	資産合計前払費用比率	-53.11
仮払金/資産合計	-73.25	代表者取引評点	-9.78	割引手形/与信額	-7.42	資産合計流動資産比率	0.96
流動資産計/資産合計	0.96	担保余力評点	0.11	裏書譲渡手形/与信額	-40.13	資産合計固定資産比率	0.47
土地/資産合計	-5.89	買掛額	-0.02	減価償却実施額/与信額	-3.37	資産合計支払手形比率	-29.60
有形固定資産計/資産合計	-1.02	経常運転資金額/資産合計	-14.53	資金利益額/与信額	-17.98	資産合計受取手形比率	-20.11
長期貸付金/資産合計	-30.19	長期運転資金額/資産合計	-3.76	直接費差引後利益/与信額	0.58	資産合計当座資産比率	-2.62
固定資産計/資産合計	0.47	一時運転資金額/資産合計	-50.58	手数料外為利益/与信額	-162.88	資産合計棚卸資産比率	-5.12
繰延資産計/資産合計	-96.97	その他運転資金額/資産合計	-75.78	合計/与信額	0.52	基準日	
資産合計	-0.04	土地建物購入資金額/資産合計	-34.58	経常運転資金額/与信額	-5.17	デフォルトからの経過時間	
支払手形/資産合計	-29.60	事業設備資金額/資産合計	-32.73	長期運転資金額/与信額	1.07		
買掛金/資産合計	-9.25	その他資金額/資産合計	-355.16	一時運転資金額/与信額	-21.80		
短期借入金/資産合計	-2.71	毀損合計	-0.26	その他運転資金額/与信額	-31.45		
未払金/資産合計	-14.92	累積毀損額	-0.64	土地建物購入資金額/与信額	-16.01		
未払費用/資産合計	-37.06	デフォルト回数		事業設備資金額/与信額	-19.75		
前受金/資産合計	-114.01	デフォルトからの経過時間	-0.42	その他資金額/与信額	-164.14		
未成工事受入金/資産合計	-160.07	担保余力割る/与信額	1.20	与信額/EAD	0.05		
前受収益/資産合計	-8736.06	現金預金/与信額	-0.93	総資本営業利益率	3.23		
流動負債計/資産合計	-1.63	受取手形/与信額	-0.19	総資本経常利益率	2.59		
社債/資産合計	-188.03	売掛金/与信額	-1.10	ROA	2.21		
長期借入金/資産合計	-1.15	有価証券/与信額	-22.36	ROE	1.19		
固定負債計/資産合計	-1.14	当座資産計/与信額	-0.81	売上高総利益率	0.74		
負債合計/資産合計	-1.60	棚卸資産計/与信額	-1.11	売上高営業利益率	1.11		
資本金/資産合計	-5.07	前渡金/与信額	-23.74	売上高経常利益率	1.85		
純資産合計/資産合計	2.80	未収入金/与信額	-4.17	売上高当期利益率	1.92		

で定義される。  $S_R$ ,  $S_{NR}$  それぞれからランダムに抽出したレコードの推定正常復帰確率を比べた時,  $S_R$  から抽出されたレコードの推定正常復帰確率が  $S_{NR}$  から抽出されたレコードの推定正常復帰確率より高い確率を示している。

AUC は順序性に関する指標である(山下・三浦, 2011)。式(3.2)から分かるように, AUC は推定正常復帰確率  $P$  のアフィン変換に関して不変であり, 推定正常復帰確率  $P$  の水準には依存せず, 任意の  $P_i$ ,  $P_j$  の大小関係, 順序関係のみ依存して算出される。

なお, AUC と同等の概念として AR 値が存在する。両者は数学的には式(3.3)の様な等式が成り立っており, 金融実務におけるモデル評価に用いられる指標である。

$$(3.3) \quad AR \text{ 値} = 2AUC - 1$$



### 3.3 Hosmer-Lemeshow 統計量の定義

次に, Hosmer-Lemeshow 統計量について定義する.

Hosmer-Lemeshow 検定は, Pearson の適合度検定の拡張の一種であり, 観測された頻度分布がモデルによる理論分布と同じかどうかの検定である. 推定正常復帰確率をもとにデータを適当なグループ(通常グループ数は 10 である (Hosmer and Lemeshow, 1989). 本稿でも 10 とした)に分け, 各グループの期待正常復帰件数(各グループに属するデータの推定正常復帰確率の和)と実績正常復帰件数を比べることで, モデルの適合度検定を行う.

Hosmer-Lemeshow 検定の統計量は式(3.4)の  $HL$  で定義される.  $G$  をグループ数,  $O_g$  をグループ  $g$  の実績正常復帰件数,  $N$  を全レコード数,  $N_g$  をグループ  $g$  に属すレコード数,  $\pi_g$  をグループ  $g$  の平均推定正常復帰確率とする時, グループは, 以下のように作る. まず, レコードを推定正常復帰確率が小さいものから順に並べ, 一番小さいものから  $N/G$  番目までのレコードをグループ 1 に,  $(N/G) + 1$  番目から  $2N/G$  番目までのレコードをグループ 2 に入れる. 以下同様にグループを作っていく.

$$(3.4) \quad HL = \sum_{g=1}^G \frac{(O_g - N_g \pi_g)^2}{N_g \pi_g (1 - \pi_g)}$$

式(3.4)の分子は各グループの実績正常復帰件数とモデルの期待正常復帰件数の差の二乗で, 分母は各グループの期待正常復帰件数の分散である. 分母の期待正常復帰件数の分散によって, 分子の各グループの実績正常復帰件数とモデルの期待正常復帰件数の差の二乗を標準化している.

Hosmer-Lemeshow 検定の統計量はモデルが真のモデルであるとき, 自由度  $G - 2$  の  $\chi^2$  分布に従うことがシミュレーションにより示されており (Hosmer and Lemeshow, 1980), この性質をもとに Pearson の適合度検定と同様に観測された頻度分布がモデルによる理論分布と同じかどうかを検定する.

式(3.4)から分かるように, 各グループにおける実績正常復帰件数( $O_g$ )とモデルによる期待正常復帰件数( $N_g \pi_g$ )の差が小さいほど統計量  $HL$  が小さくなる. そこで, 本研究ではこの性質を利用し, Hosmer-Lemeshow 統計量をモデルの評価指標として用いる.

表 5, 6 にデータセット 1 のインサンプルとアウトオブサンプルの Hosmer-Lemeshow 統計量の算出例をまとめた. 表 5 のデータセット 1 のインサンプルでは, 推定正常復帰確率( $P_i$ )が高いグループほど実績正常復帰件数( $O_g$ )が多い. また, 各グループの期待正常復帰件数( $N_g \pi_g$ )と実績正常復帰件数( $O_g$ )が比較的近い値をとっており, それぞれの合計はほぼ等しい. 表 6 の

表 5. データセット 1 のインサンプル Hosmer-Lemeshow 統計量の算出例.

$g$	$P_i$	$O_g$	$N_g \pi_g$	$\frac{(O_g - N_g \pi_g)^2}{N_g \pi_g (1 - \pi_g)}$	$N_g$
1	(0.000,0.002]	0	0.185	0.185	206
2	(0.002,0.003]	1	0.476	0.576	205
3	(0.003,0.005]	2	0.849	1.566	205
4	(0.005,0.008]	2	1.362	0.300	205
5	(0.008,0.013]	3	2.190	0.302	206
6	(0.013,0.022]	5	3.526	0.627	205
7	(0.022,0.036]	4	5.879	0.617	205
8	(0.036,0.064]	9	10.027	0.110	205
9	(0.064,0.153]	18	20.161	0.257	205
10	(0.153,1.000]	75	74.347	0.008	206
合計		119	118.999	4.553	2053

表 6. データセット 1 のアウトサンプル Hosmer-Lemeshow 統計量の算出例.

$g$	$P_i$	$O_g$	$N_g \pi_g$	$\frac{(O_g - N_g \pi_g)^2}{N_g \pi_g (1 - \pi_g)}$	$N_g$
1	(0.000,0.002]	0	0.171	0.171	126
2	(0.002,0.003]	0	0.351	0.352	126
3	(0.003,0.005]	1	0.506	0.474	126
4	(0.005,0.007]	2	0.718	2.341	126
5	(0.007,0.009]	3	0.9910	4.085	126
6	(0.009,0.013]	3	1.401	1.849	125
7	(0.013,0.021]	3	2.191	0.303	127
8	(0.021,0.037]	2	3.566	0.708	125
9	(0.037,0.060]	2	5.839	2.637	126
10	(0.060,1.000]	9	13.178	1.479	126
合計		25	28.903	14.397	1259

表 7. データセット 1 の推定結果.

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-7.101	0.708	-10.032	0.000
毀損の有無	-2.665	1.035	-2.574	0.010
YJ 資産合計土地比率	5.463	1.983	2.755	0.006
YJ 受取手形/与信額	0.951	0.388	2.454	0.014
YJ 担保余力評点	-0.323	0.133	-2.424	0.015
YJ 期末役員従業員数	1.009	0.274	3.677	0.000
売上高	-0.000	0.000	-2.613	0.009
資金利益額評点	0.066	0.033	1.983	0.047
CRITS 標準スコア	0.030	0.008	3.635	0.000
2007 年 3 月	-0.427	0.666	-0.640	0.522
2007 年 9 月	-1.489	0.901	-1.653	0.098
2008 年 3 月	-1.026	0.547	-1.875	0.061
2008 年 9 月	2.377	0.238	9.993	0.000
2009 年 3 月	0.175	0.394	0.445	0.656
2009 年 9 月	0.660	0.320	2.066	0.039
2010 年 3 月	0.402	0.351	1.146	0.252
2010 年 9 月	-0.673	0.485	-1.388	0.165

データセット 1 のアウトオブサンプルでは、グループ 9 の期待正常復帰件数 ( $N_g \pi_g$ ) と観測正常復帰件数 ( $O_g$ ) に大きな差があるが、他のグループに関しては大きな差はなく、それぞれの合計も比較的近い値をとっている。

### 3.4 回帰係数の推定結果

2010 年 9 月, 2011 年 3 月, 2011 年 9 月を境に、それぞれの時点以前のデータをインサンプル (学習データ) とし、それ以降のデータをアウトオブサンプル (検証データ) とする、3 つのデータセット (それぞれデータセット 1, 2, 3 とする) に関して正常復帰確率推定モデルを作成した。

データセット 1, 2, 3 による回帰係数の推定結果をそれぞれ表 7, 8, 9 に示した。なお、表 7, 8, 9 の「YJ」は YJ 変換済みの変数であることを意味する。また、AIC, AUC, Hosmer-Lemeshow 統計量は、それぞれ表 10 のようになった。

AIC, AUC, Hosmer-Lemeshow 統計量,  $z$  値を総合的に用いてモデル評価を行い、説明変数

表 8. データセット 2 の推定結果.

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-7.266	0.663	-10.957	0.000
毀損の有無	-2.780	1.035	-2.685	0.007
YJ 資産合計土地比率	5.020	1.854	2.708	0.007
YJ 受取手形/与信額	0.872	0.357	2.439	0.015
YJ 担保余力評点	-0.184	0.120	-1.532	0.126
YJ 期末役員従業員数	0.991	0.256	3.877	0.000
売上高	-0.000	0.000	-2.597	0.009
資金利益額評点	0.071	0.031	2.255	0.024
CRITS 標準スコア	0.036	0.008	4.546	0.000
2007 年 3 月	-0.486	0.672	-0.724	0.469
2007 年 9 月	-1.550	0.914	-1.697	0.090
2008 年 3 月	-1.063	0.549	-1.936	0.053
2008 年 9 月	2.301	0.222	10.341	0.000
2009 年 3 月	0.110	0.390	0.281	0.779
2009 年 9 月	0.589	0.311	1.896	0.058
2010 年 3 月	0.354	0.344	1.030	0.303
2010 年 9 月	-0.719	0.484	-1.485	0.137
2011 年 3 月	0.466	0.326	1.426	0.154

表 9. データセット 3 の推定結果.

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-7.263	0.639	-11.373	0.000
毀損の有無	-2.831	1.033	-2.740	0.006
YJ 資産合計土地比率	4.889	1.790	2.731	0.006
YJ 受取手形/与信額	0.913	0.345	2.647	0.008
YJ 担保余力評点	-0.170	0.115	-1.470	0.141
YJ 期末役員従業員数	1.018	0.248	4.105	0.000
売上高	-0.000	0.000	-2.749	0.006
資金利益額評点	0.067	0.030	2.221	0.026
CRITS 標準スコア	0.034	0.008	4.583	0.000
2007 年 3 月	-0.447	0.677	-0.661	0.509
2007 年 9 月	-1.504	0.923	-1.630	0.103
2008 年 3 月	-1.022	0.552	-1.852	0.064
2008 年 9 月	2.335	0.213	10.978	0.000
2009 年 3 月	0.140	0.388	0.361	0.718
2009 年 9 月	0.623	0.306	2.040	0.041
2010 年 3 月	0.381	0.340	1.123	0.261
2010 年 9 月	-0.688	0.485	-1.419	0.156
2011 年 3 月	0.488	0.321	1.517	0.129
2011 年 9 月	-0.305	0.386	-0.791	0.429

として毀損の有無フラグ変数(デフォルト発生時点からデータの基準日時点までに毀損が生じている場合は 1, そうでない場合には 0 になるフラグ変数), YJ 資産合計土地比率, YJ 受取手形/与信額, YJ 担保余力評点(0 点から 15 点の 5 段階の点数で担保余力を評価する指標, 分析銀行独自の指標), YJ 期末役員従業員数, 売上高, 資金利益額評点(3 点から 15 点の 5 段階の点数で利益を評価する指標, 担保余力評点と同じく分析銀行独自の指標), CRITS 標準スコア, 基準日カテゴリー変数を選択した。

表 10. 各データセットにおける AUC と Hosmer-Lemeshow 検定の  $p$  値.

	インサンプル	アウトオブサンプル
データセット 1 AUC	0.888	0.713
データセット 2 AUC	0.874	0.739
データセット 3 AUC	0.871	0.733
データセット 1 HL-p	4.551	14.398
データセット 2 HL-p	13.153	8.175
データセット 3 HL-p	5.627	8.912
データセット 1 AIC	636.894	
データセット 2 AIC	738.181	
データセット 3 AIC	805.583	

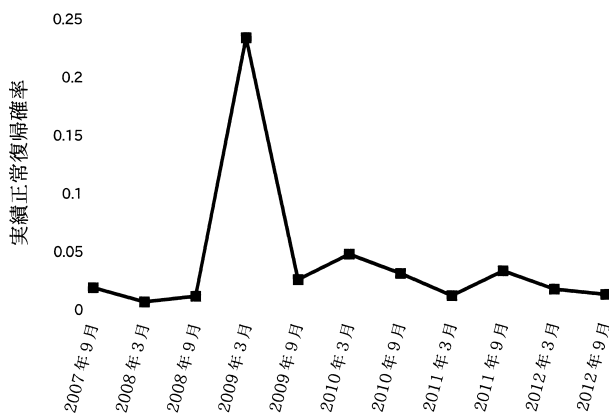


図 1. 実績正常復帰確率.

なお、基準日カテゴリー変数に関しては、係数の合計が 0 になるように制約をかけて推定した。そして、アウトオブサンプルにおいてモデルをフィッティングさせる際には、基準日カテゴリー変数以外の説明変数の係数に関しては、インサンプルにおける推定係数を用い、基準日カテゴリー変数に関しては係数を 0 とした。

表 7, 8, 9 から毀損が生じている債権は正常復帰しにくく、また、YJ 変換済みの資産合計土地比率、YJ 変換済みの受け取り手形/与信額、YJ 変換済み期末役員従業員数、資金利益評点、CRITS 標準スコアそれぞれが高いほど正常復帰しやすく、さらに、YJ 変換済み担保余力評点売上高それぞれが高いほど正常復帰しにくいことが分かる。

また、時間情報に関しては、デフォルトからの経過時間は  $p$  値が非常に大きかったため説明変数として選ばなかったが、基準日カテゴリー変数には  $p$  値が非常に小さいカテゴリーがあり、AIC, AUC, Hosmer-Lemeshow 統計量、 $z$  値を鑑みてモデルの説明変数として選択した。表 7, 8, 9 から分かるように、2008 年 9 月変数の  $p$  値が他の基準日に比べて特に小さいことが分かる。これは、川田・山下(2012)にあるように条件緩和措置の影響であると考えられる。なお、本研究では次の時点(半年後)の正常復帰を説明するモデルを作成しているため、2008 年 9 月変数は 2009 年 3 月の正常復帰の予測に用いられる変数である。図 1 からも、2009 年 3 月の実績正常復帰確率が高いことがわかる。また、2008 年 9 月以外に関しても有効な基準日が確認された。以上から、政策やマクロ経済の状態が正常復帰に影響を与えることが確認された。

### 3.5 YJ 変換による説明力の向上

YJ 変換の効果を検証するために、各データセットにおいて、提案モデル(「YJ 変換ありモデル」と呼ぶ)と提案モデルにおいて YJ 変換を施している変数の代わりに YJ 変換を施していない変数を用いてパラメータ推定を行ったモデル(「YJ 変換なしモデル」と呼ぶ)を、AIC, AUC, Hosmer-Lemeshow 統計量に基づいて比較し、YJ 変換による説明力の向上について分析した。

AIC に関しての比較結果を検討する。表 11 に YJ 変換ありモデルの AIC, YJ 変換なしモデルの AIC, YJ 変換ありモデルと YJ 変換なしモデルによる AIC をまとめた。表 11 から分かるように、YJ 変換ありモデルのほうが AIC が小さく、モデルの説明力が向上することが分かる。

AUC に関しての比較結果を検討する。表 12 に YJ 変換ありモデルの AUC, YJ 変換なしモデルの AUC, YJ 変換ありモデルと YJ 変換なしモデルによる AUC の上側 DeLong 検定(DeLong et al., 1988)による  $p$  値をまとめた。表 12 から分かるように、全てのデータセットのインサンプル、アウトオブサンプルにおいて YJ 変換ありモデルの AUC は YJ 変換なし済みモデルに比べて高い。また、DeLong 検定の  $p$  値も小さく、YJ 変換ありモデルは YJ 変換なしモデルに比べて AUC が統計的に有意に高いことが確認された。以上から、AUC に関して YJ 変換ありモデルの YJ 変換なしモデルに対する優位性が確認された。

Hosmer-Lemeshow 検定に関して比較結果を検討する。表 13 に YJ 変換モデル, YJ 変換なしモデルの Hosmer-Lemeshow 統計量をまとめた。表 13 から分かるように、データセット 2 のインサンプルを除いた全てのケースにおいて、YJ 変換モデルの方が Hosmer-Lemeshow 統計量が

表 11. YJ 変換ありモデルと YJ 変換なしモデルの AIC の比較.

	YJ 変換あり	YJ 変換なし
データセット 1	636.894	652.395
データセット 2	738.181	754.552
データセット 3	805.583	824.362

表 12. YJ 変換ありモデルと YJ 変換なしモデルの AUC の比較.

	YJ 変換あり	YJ 変換なし	$p$ 値
データセット 1 インサンプル	0.888	0.879	0.017
データセット 1 アウトオブサンプル	0.713	0.681	0.030
データセット 2 インサンプル	0.874	0.865	0.018
データセット 2 アウトオブサンプル	0.739	0.686	0.013
データセット 3 インサンプル	0.871	0.860	0.006
データセット 3 アウトオブサンプル	0.733	0.674	0.062

表 13. YJ 変換ありモデルと YJ 変換なしモデルの Hosmer-Lemeshow 統計量の比較.

	YJ 変換あり	YJ 変換なし
データセット 1 インサンプル	4.551	9.139
データセット 1 アウトオブサンプル	14.398	16.934
データセット 2 インサンプル	13.153	7.498
データセット 2 アウトオブサンプル	8.175	16.389
データセット 3 インサンプル	5.627	10.751
データセット 3 アウトオブサンプル	8.912	16.318

小さい。以上から、Hosmer-Lemeshow 検定に関しても YJ 変換ありモデルの YJ 変換なしモデルに対する優位性が確認された。

以上のように、AIC, AUC, Hosmer-Lemeshow 統計量の結果の比較により、YJ 変換ありモデルの YJ 変換なしモデルに対する優位性が確認され、YJ 変換によりモデルの説明力が向上することが分かる。

#### 4. まとめと今後の課題

##### 4.1 まとめ

2007年3月から2012年3月までのある地方銀行の法人向け貸付債権に対して、担保情報、保証情報、財務情報、時間情報を用いたロジットモデルで、デフォルト状態の債権の正常復帰確率推定モデルを作成し、正常復帰に影響を与える要因について分析した。その結果、毀損が生じている債権は正常復帰しにくく、また、YJ変換済みの資産合計土地比率、YJ変換済みの受け取り手形/与信額、YJ変換済み期末役員従業員数、資金利益評点、CRITS標準スコアそれぞれが高いほど正常復帰しやすく、さらに、YJ変換済み担保余力評点売上高それぞれが高いほど正常復帰しにくいことがわかった。さらに、説明変数にYJ変換を施すことによって、モデルの説明力が向上することが確認された。

##### 4.2 今後の課題

###### 4.2.1 モデルの精度向上のための課題

さらにモデルの予測精度を向上させるためには、(1)基準日カテゴリー変数でなく、マクロ経済変数をモデルに組み込む、(2)各企業の各レコードを独立に分析するのではなく、生存データ解析などの手法を用いてレコードの時系列変化情報を用いて分析を行う、(3)さらなる説明変数の変換、合成などを試みるなどの工夫が必要である。

###### 4.2.2 推定方法に関する課題

また、本研究ではモデルの推定に最尤法を用いたが、モデルの利用目的によって最尤法以外の方法を検討する必要がある。モーメント法、AUC(Miura et al., 2010)、Hosmer-Lemeshow 統計量を最適化する推定方法などが考えられる。

###### 4.2.3 モデルの利用に関する課題

1.1節で説明したように、デフォルト状態の債権の分析は、(1)詳細なLGD分析、(2)銀行の貸出意思決定、(3)デフォルト状態の債権の現在価値の算出、に必要とされている。

(1)詳細なLGD分析に関して、本研究で正常復帰に影響を与える要因が判明した。そこで、これらの要因をLGD分析に応用する方法を考える必要がある。

(2)銀行の貸出意思決定に関して、本研究で正常復帰しやすい債権の特徴がわかったが、実際にどのような貸出意思決定をすべきかに関しては研究を行わなかった。そこで、実際の貸出意思決定において、どういった意思決定を行うべきなのかを検討する必要がある。

(3)デフォルト状態の債権の現在価値の算出に関して、1.1節で説明したように、今後強制適用が予定されるIAS(国際会計基準)のIAS39では、デフォルト状態の債権の適切な現在価値評価が要求され、デフォルト状態の債権の適切な現在価値評価のためには正常復帰確率の推定が必要である。本研究では次の時点(半年後)までの正常復帰確率を推定するモデルを開発したが、より精緻なデフォルト状態の債権の評価に用いるためには、割引を考慮した復帰までの時間を組み込んだ正常復帰確率推定モデルの開発が必要となる。これも今後の課題とする。

## 参 考 文 献

- 青沼君明, 市川伸子(2008). 『EXCEL で学ぶバーゼル II と信用リスク評価手法』, 金融財政事情研究会, 東京.
- Atiya, A. F. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results, *IEEE Transactions on Neural Networks*, **12**, 929–935.
- Bastos, J. A. (2010). Forecasting bank loans loss-given-default, *Journal of Banking and Finance*, **34**, 2510–2517.
- Bellotti, T. and Crook, J. (2012). Loss given default models incorporating macroeconomic variables for credit cards, *International Journal of Forecasting*, **28**, 171–182.
- Caselli, S., Gatti, S. and Querci, F. (2008). The sensitivity of the loss given default rate to systematic risk: New empirical evidence on bank loans, *Journal of Financial Services Research*, **34**, 1–34.
- Citron, D., Wright, M., Ball, R. and Rippington, F. (2003). Secured creditor recovery rates from management buy-outs in distress, *European Financial Management*, **9**, 141–161.
- DeLong, E. R., DeLong, D. M. and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach, *Biometrics*, **44**, 837–845.
- Dermine, J. and Neto de Carvalho, C. (2006). Bank loan losses-given-default: A case study, *Journal of Banking and Finance*, **30**, 1243–1291.
- Duffie, D. and Singleton, K. J. (1999). Modeling term structures of defaultable bonds, *Review of Financial Studies*, **12**, 687–720.
- Grunert, J. and Weber, M. (2009). Recovery rates of commercial lending: Empirical evidence for German companies, *Journal of Banking and Finance*, **33**, 505–513.
- Gürtler, M. and Hibbeln, M. (2011). Pitfalls in modeling loss given default of bank loans, Working Paper Series, IF35V1, Department of Finance, Institute of Technology, University of Braunschweig.
- Hosmer, D. and Lemeshow, S. (1980). Goodness of fit tests for the multiple logistic regression model, *Communications in Statistics*, **9**, 1043–1069.
- Hosmer, D. and Lemeshow, S. (1989). *Applied Logistic Regression*, Wiley, New York.
- 乾 孝治, 室町幸雄(2000). 『金融モデルにおける推定と最適化(シリーズ現代金融工学)』, 朝倉書店, 東京.
- 川田章広, 山下智志(2012). 回収実績データに基づく LGD の要因分析と多段階モデルによる LGD および EL 推計, 金融庁金融研究センター ディスカッションペーパー DP2012-6.
- Kijima, M. and Suzuki, T. (2001). A jump-diffusion model for pricing corporate debt securities in a complex capital structure, *Quantitative Finance*, **32**, 611–620.
- Lane, W. R., Looney, S. W. and Wansley, J. W. (1986). An application of the Cox proportional hazards model to bank failure, *Journal of Banking and Finance*, **10**, 511–531.
- Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance*, **29**, 449–470.
- Miura, K., Yamashita, S. and Eguchi, S. (2010). AUC maximization method in credit scoring, *Risk Model Validation*, **4**, 3–25.
- 三浦 翔, 山下智志, 江口真透(2010). 内部格付手法における回収率・期待損失の統計型モデル, 金融庁金融研修センター・リサーチレビュー 2010, 174–205.
- 森平爽一郎(2009). 『信用リスクモデリング—測定と管理(応用ファイナンス講座)』, 朝倉書店, 東京.
- 森平爽一郎, 岡崎貴治(2009). マクロ経済効果を考慮したデフォルト確率の期間構造推定, ワーキングペーパーシリーズ, WIF-09-004, 早稲田大学ファイナンス総合研究所.
- 日本銀行金融機構局(2005). 内部格付制度に基づく信用リスク管理の高度化, リスク管理高度化と金融機関経営に関するペーパーシリーズ.
- Schwalm, C. R., Black, T. A., Amiro, B. D., Arain, M. A., Barr, A. G., Bourque, C. P.-A., Dunn, A., Flanagan, L. B., Giasson, M.-A., Lafleur, P. M., Margolis, H. A., McCaughey, J. H., Orchansky,

- A. L. and Wofsy, S. C. (2006). Photosynthetic light use efficiency of three biomes across an east-west continental-scale transect in Canada, *Agricultural and Forest Meteorology*, **140**, 269–286.
- 社団法人全国地方銀行協会信用リスク管理高度化支援室(2005). 信用リスク情報統合システム[CRITS]テクニカルガイダンス第 2.1 版.
- 白田佳子(1999). 『企業倒産予知情報の形成—会計理論と統計技術の応用』, 中央経済社, 東京.
- Sung, T. K., Chang, N. and Lee, G. (1999). Dynamics of modeling in data mining: Interpretive approach to bankruptcy prediction, *Journal of Management Information Systems*, **16**, 63–85.
- 山下智志, 安道知寛(2006). 時間依存共変量を用いたハザードモデルによるデフォルト確率期間構造の推計手法, 統計数理, **54**, 23–38.
- 山下智志, 三浦 翔(2011). 『信用リスクモデルの予測精度—AR 値と評価指標—』, 朝倉書店, 東京.
- Yeo, I. and Johnson, R. (2000). A new family of power transformations to improve normality or symmetry, *Biometrika*, **87**, 954–959.
- Wang, Q. J. and Robertson, D. (2011). Multisite probabilistic forecasting of seasonal flows for streams with zero value occurrences, *Water Resources Research*, **47**, 3–11.
- Wang, Q. J., Robertson, D. E. and Chiew, F. H. S. (2009). A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites, *Water Resources Research*, **45**, W05407.



## Probability of Recovery from Default Loan to Performing Loan

Yuta Tanoue<sup>1,2</sup> and Satoshi Yamashita<sup>3</sup>

<sup>1</sup>Department of Statistical Science, School of Multidisciplinary Sciences,  
The Graduate University for Advanced Studies

<sup>2</sup>Research Fellow of Japan Society for the Promotion of Science

<sup>3</sup>The Institute of Statistical Mathematics

We made a model which estimates the probability of recovery of a default company, and analyse processes of recovery. We use corporate loan data of a regional bank which were recorded every half year from September 2007 to March 2012. Borrower characteristics, financial characteristics, and time period categorical variables are used as explanatory variables in the model choice. We propose Yeo-Johnson (YJ) transformation for explanatory variables, which is an expansion of the Box-Cox transformation to negative values. After the model choice, we advance a logistic regression in the model with some YJ transformed borrower characteristic, borrower characteristic, time period categorical variables and so on. We choose these variables by comparing model fitness using AIC, AUC and Hosmer-Lemeshow test statistics. We thus confirmed that YJ transformation improves model estimating performance.