

## デフォルト実績データによるデフォルト依存関係の推定 - 2ファクターモデルによるアセット相関の最尤推定 -

北野 利幸  
東京工業大学

(受理 2005年9月26日; 再受理 2006年10月9日)

**和文概要** 本論では、デフォルト実績データを用いたデフォルト依存関係の推定を論じる。債務者の信用力を代理する潜在変数の変動モデルとして正規分布の2ファクターモデルを用い、債務者の属するカテゴリ内外の相関を考慮できるようにした。さらに複数カテゴリ間の相関に単純化の仮定を加えて、アセット相関を表すファクター負荷の最尤法による推定計算を実装した。モンテカルロ法による推定値の頑強性の検討および実際の日本のカテゴリ別デフォルトデータを用いた実証分析により、この方法は既存の方法より計算負荷が少なく、また実際の推定を前提としてもカテゴリを考慮しないモデルより頑強であることが確認できた。このことはあるカテゴリ内の債務者間の相関(カテゴリ内相関)のみを推定する目的においても、カテゴリ間相関の推定を通じて他のカテゴリの情報も用い推定することが重要であることを示している。

**キーワード:** 確率モデル, 金融, 信用リスク管理, デフォルト依存, パラメータ推定

### 1. はじめに

信用プールの評価問題において中心的な課題の一つは、債務者間の信用力ひいてはデフォルト事象発生依存性の実証である。デフォルト依存関係をモデル化することはできて、稀なデフォルト事象の実績データが少なく、そのパラメータ推定は非常に難しい問題となっている。本論では、デフォルトに関する潜在変数法を基礎として、デフォルト債務者数の履歴データからこの依存性を最尤法により推定する方法を議論する。本論で用いた依存関係のモデルは、債務者の属するカテゴリによってカテゴリ間、カテゴリ内の債務者の潜在変数の依存関係が変わることを想定した、潜在変数に関する正規分布の2ファクターモデルである<sup>1</sup>。本論では、複数カテゴリ間の相関に単純化の仮定を用いて、Gordy & Heitfield [11] のフレームワークに沿いながらも、実際のパラメータの推定計算を可能とした。

この方法を用いて、推定誤差のモンテカルロ・シミュレーションを行うことにより、カテゴリ間の相関に強い仮定を必要とする1ファクター形式のモデルの適用では、推定値にバイアスを引き起こす可能性があり、カテゴリ内外の部分的な相関を考慮できる2ファクターモデルによる推定がより頑強であることを示した。さらに、日本の企業の資本金別倒産数の時系列データを用いてそれぞれのパラメータを推定した結果、カテゴリ間の相関は時期によって変動しており、実際のデータからも、2ファクターモデルによる複数カテゴリの情報を使った推定が好ましいことが裏付けられる。

本論における分析対象は、市場性の低い中小企業向け債権や住宅ローンなどの個人向け債

<sup>1</sup>正規分布に従う確率変数間では独立と無相関は同値であり、本論では相関をもって依存性の尺度とする議論となる。

権を想定している。このような債権は、株価変動や信用格付け推移などの信用力の代理変数によるデフォルトの依存構造の推定ができない。デフォルトの依存構造は、信用プールの最大損失に大きく影響を及ぼす重要な要素のひとつである一方で、その実証については特に本邦では研究成果がない。本論のような推定手法の検討により、様々なデータでの実証研究が進み、ひいては金融機関のリスク管理や流動化資産の評価の高度化を通して、信用リスク市場の効率化に資することが期待される。

以降、2節では既存研究とともに潜在変数に正規分布を想定したアセット相関推定の位置づけを説明する。3節では既存の推定法、次の4節で本論の推定法の定式化と説明を行う。5節ではモンテカルロ法による推定誤差とバイアスの検討を行う。6節では実際のデータによりパラメータ推定を行った結果を示し、5節の内容が実用上にも影響することを見る。7節で結論としてまとめと今後の課題を述べる。

## 2. 先行研究

デフォルトに関する潜在変数法とは、Merton [18] に端を発する、負債の満期時に企業価値（アセット価値）がある閾値を越えて小さい場合にデフォルトすると考える構造（Structural）モデルのフレームワークに沿うものである。Zhou [27] ではMertonモデルの発展形である初期到達時点型の構造モデルを基礎としてデフォルト相関のモデル化をしているが、その中でアセット相関パラメータが導入されデフォルト相関との関係が議論されている<sup>2</sup>。潜在変数による相関のモデル化に当たっては、複数の潜在変数＝アセット価値（の変化率）の挙動をどう考えるかが問題となる。

以下では、デフォルトに関する潜在変数法を基礎としたアセット相関の実証研究のうち、本論と同様にデフォルト実績データから推定したものについてまとめる。上場企業や社債発行企業など信用力の高い債務者同士のアセット相関を推定しようとした場合、過去に観測されたデフォルトではサンプル数が十分でないことが多い。このためデフォルト実績から直接的に相関を推定するのではなく、株価変動や格付け推移など信用力を代理すると考えうる情報からこれらとデフォルト発生との関係をモデル化した上で、アセット相関構造を推定しようとする研究や実務的取り組みがある。しかしながら、本論では市場性が低く、間接金融（銀行・ノンバンクからの借り入れ）によって債務者となる中小企業・個人の債権を主な適用対象としているため、株価や格付け推移などの市場情報をつかった相関構造の実証は議論の対象外としている。

以下に述べるアセット相関の実証研究は、全て潜在変数に正規分布を想定したものであり、このモデルのもたらす潜在変数間の依存構造は、正規コピュラ（Gaussian Copula）となる。前述のように、潜在変数間の依存構造に正規コピュラ以外のものを適用し信用ポートフォリオの損失分布に与える影響を議論した研究があり、実証研究においても正規コピュラの適用についてはそれ自身吟味が必要である。しかしながら、実証研究自体が未だ豊富でない現状において、いずれにしる実証的根拠に乏しい他の依存構造（コピュラ）は適用せず、まずはパラメータ数が一つと少なく推定も容易であり、また資産価値変動のモデルとしてオーソドックスな多変量正規分布の依存構造である正規コピュラを適用した過去の学術的・実務的取り組みは妥当と考えられる。様々な依存構造によるモデル化の可能性がいわば無限にある中で、本論のような実証研究はその依存構造の解明にあり、正規コピュラを適用する正当性・妥当性も、本論のような実証研究を通してのみ論じ得ると考えられる。

<sup>2</sup>デフォルト相関並びにアセット相関の定義は付録 A 参照。

## 2.1. デフォルト履歴を使った推定方法

Lucas [15] は、Moody's 社の 25 年間の格付けごとのデフォルト実績を用い、各格付けにある債務者が同じデフォルト率・デフォルト相関であると仮定し、同時デフォルト確率を簡便な方法で推定することで、一定期間のデフォルト相関係数を初めて推定した。Nagpal & Bahar [20] は、ほぼ同じ考え方によって、Standard & Poor's 社の 19 年間のデータを用い、投資適格・投資不適格のそれぞれの階級同士のデフォルト相関係数を業種別に推定した。さらに Servigny & Olivier [24] は、やや改良された推定方法によるデフォルト相関係数推定値と、アセット変化率が多変量正規分布に従うという前提の下で株価を用いて推定した値とを比較し、双方の隔たりは大きいと主張した。

Gordy [10] は、潜在変数の 1 ファクターモデル<sup>3</sup>を前提とした場合の、デフォルト率のボラティリティとアセット相関係数の関係を示した。加えて格付け毎のデフォルト率の時系列推移データからデフォルト率のボラティリティの推定値を得て、この上でアセット相関係数を初めて推定している（モーメント法<sup>4</sup>）。Gordy & Heitfield [11] では、自らが提案する最尤法による推定が、少ないサンプルにおいてモーメント法と比較してどの程度頑強であるかを Moody's 社並びに Standard & Poor's 社のデータにより検証し、特に現実のデータで問題となるサンプル数（データ期間）が充分にない場合の最尤法の優位性を主張している。

## 2.2. 多数均一分散プールを対象とした実証研究

Gordy [10] では信用格付け毎のデフォルト率データを用いてアセット相関を推定した。格付けを用いていることから、比較的大きな規模の企業の信用リスクを適用対象としているが、一方でこの推定方法は、観測対象が信用力・相関の面で均一であると考えうるような分散の強い観測対象から得られたデフォルト数・デフォルト率の履歴データに最も適した方法である。また正規分布の 1 ファクターモデルがバーゼル II の銀行自己資本賦課額の計算ロジックにも用いられたこともあり、リース、住宅ローン、中小企業向けローンなど、個人や中小企業を対象とした均一性の高い様々な信用プールに対して適用した実証研究がある<sup>5</sup>。

特に中小企業については、Dietsch & Petey [7] が、大手金融情報プロバイダーであるフランスの Coface とドイツの Creditreform より得た中小企業のパネルデータを元に、前述の Gordy [10] のモーメント法を用い、信用クラス別、企業規模別のデフォルト率とアセット相関係数の推定を行った。この結果、大企業より中小企業のほうがアセット相関係数は小さくなること、推定値は 1~3% と非常に小さい上、企業規模の小さいほうが小さくなること等を報告した。さらにデフォルト率とアセット相関係数との間には正の関係がある可能性があることを示唆した。

Dullmann & Scheule [6] はより広範囲・長期間のドイツのデータを用いて、デフォルト相関係数の推定に、モーメント法に加え Gordy & Heitfield [11] の最尤法とその派生形も適用した。その上で同様にアセット相関係数とデフォルト率並びに企業規模との関係について分析した。

Hemerle et al. [13] は、Boegelein et al. [1] で用いられたデフォルト率の時系列データ（倒産会社数と全体会社数の各期の比率）から G7 国のデータのみを用いて、最尤法による方法で各国・各業種でのデフォルト率並びにアセット相関係数の推定を行った。その結果、ア

<sup>3</sup>潜在変数が後述する正規分布ファクターモデルで表現される形式は、Vasicek [26] が提案したもので、実務モデルである CreditMetrics [12] 等で用いられる一般的な形式である。

<sup>4</sup>モーメント法については付録 B 参照。

<sup>5</sup>リースについては Duchemin [5]、Schmit [22]、住宅ローンについては Calem [2] 等。

セット相関係数は高くても 2.3 % (0.023) どもりであったこと、彼らの提案する 1 期前のデフォルト率の実現値が今期のモデル中のデフォルト率に影響を与えるというモデルによれば、デフォルトの変動は時系列的な変動によって説明される部分が増えるため、モデル中のアセット相関係数の推定値は時系列構造を考えない Nagpal & Bahar [20] 等の方法での推定値より大幅に小さくなること、等を主張した。

### 3. デフォルトモデルとパラメータ推定

本論のモデルの基礎となる、デフォルトモデル (デフォルトに関する潜在変数法、ファクターモデル) と最尤法によるパラメータ推定について説明する。

#### 3.1. 潜在変数のファクターモデル

デフォルトに関する潜在変数法では、デフォルト状態は確率変数である潜在変数がある閾値を下回った場合に対応すると考える。ここでは、複数債務者の潜在変数を以下のようなファクターモデルで表現し、各債務者のデフォルト確率間に依存関係を導入する<sup>6</sup>。

デフォルトに関する潜在変数法

債権プール中の債務者の数を  $n$  とする。デフォルト・非デフォルトなどの 2 状態を扱うデフォルトに関する潜在変数法は次のように定義される (Frey and McNeil [9])。

潜在変数ベクトル  $V = (V_1, \dots, V_n)$  を  $n$  次元の確率変数ベクトルとする。 $i$  番目の要素は、債務者  $i$  に関する潜在変数である。債務者  $i$  についてのデフォルトに関する閾値  $\theta_{0,i}$  を  $-\infty = \theta_{-1,i} < \theta_{0,i} < \theta_{1,i} = \infty$  とすると、状態ベクトル  $S$  の要素  $S_i$  が次を満たすときデフォルトに関する潜在変数法という。

$$S_i = j \iff V_i \in (\theta_{j-1,i}, \theta_{j,i}], \quad j \in \{0, 1\}, \quad i \in \{1, \dots, n\} \quad (1)$$

$S_i = 0$  をデフォルト、 $S_i = 1$  を非デフォルトとすれば、潜在変数  $V_i$  が閾値  $\theta_{0,i}$  を下回った場合がデフォルト状態に対応する<sup>7</sup>。確率変数  $V_i$  が分布  $G_{0,i}$  に従う ( $V_i \simeq G_{0,i}$ ) として、デフォルト確率  $p_i$  は次のようになる。

$$p_i = P[S_i = 1] = P[V_i \leq \theta_{0,i}] = G_i(\theta_{0,i}) \quad (2)$$

#### ファクターモデル

債務者  $i$  の時点  $t$  における潜在変数  $V_{i,t} \simeq G_i$  について次のような 1 ファクターモデルを考える。

$$V_{i,t} = \rho_i X_t + \sqrt{1 - \rho_i^2} \varepsilon_{i,t} \quad (3)$$

ここで  $X_t \simeq G_1$  は全体の変動を表す共通ファクター、 $\varepsilon_{i,t} \simeq G_{2,i}$  は債務者  $i$  独自の要因を表す固有ファクターであり、任意の分布をとる確率変数である。 $\rho_i$  はファクターの加重和のウェイトを表す 0 以上 1 以下の実数である。ファクターモデルを導入することにより、潜在変数の分布  $G_{0,i}$  は、共通ファクターと固有ファクターの従う分布  $G_1$ 、 $G_{2,i}$  とウェイト  $\rho_i$  から畳み込みによって事後的に定まる。分布  $G_1$ 、 $G_{2,i}$  の分散が等しい場合には、債務者  $i, j$  間のアセット相関係数は  $\rho_i \rho_j$  となる。

<sup>6</sup>ファクターモデルによる相関のモデル化の概要は Schönbucher [23] 等。ここでの整理は北野 [14] と同様。

<sup>7</sup>CreditMetrics は格付け状態間の推移も潜在変数法によるモデル化の対象としているが、これをデフォルト・非デフォルトの 2 状態に簡素化したこの形式が Gordy [10]、Finger [8] など広く応用されている。

### カテゴリと条件付きデフォルト確率

債務者  $i$  がカテゴリ  $g$  に属するとする．ファクターウェイト  $\rho_i$ 、閾値  $\theta_{0,i}$  はカテゴリ  $g$  内の債務者については同じと仮定し、それぞれ  $\rho_g, \theta_g$  と表す．カテゴリ内共通ファクター  $X_{g,t}$  の分布  $G_1$  は標準正規分布  $N(0, 1)$ 、固有ファクター  $\varepsilon_{i,t}$  の分布  $G_{2,i}$  も債務者によらず標準正規分布とする．正規分布の再生性により、このとき潜在変数  $V_{g,i,t}$  の分布  $G_0$  も  $i$  によらず全て標準正規分布  $N(0, 1)$  となる．ファクターモデルは次のように書き直される．

$$V_{g,i,t} = \rho_g X_{g,t} + \sqrt{1 - \rho_g^2} \varepsilon_{i,t} \quad (4)$$

このとき、カテゴリ  $g$  に属する債務者間のアセット相関係数は  $\rho_g^2$  となる．これをカテゴリ内相関と呼ぶ．

カテゴリ内共通ファクター  $X_{g,t} = x_{g,t}$  の下での条件付きのデフォルト確率  $p_{g,t}$  を  $p_g(x_{g,t})$  と表記すると、(1) 式、(3) 式より次のようになる．

$$\begin{aligned} p_g(x_{g,t}) &= P \left[ \rho_g X_{g,t} + \sqrt{1 - \rho_g^2} \varepsilon_{i,t} \leq \theta_g \mid X_{g,t} = x_{g,t} \right] \\ &= P \left[ \varepsilon_{i,t} \leq \frac{\theta_g - \rho_g x_{g,t}}{\sqrt{1 - \rho_g^2}} \mid X_{g,t} = x_{g,t} \right] = G_2 \left( \frac{\theta_g - \rho_g x_{g,t}}{\sqrt{1 - \rho_g^2}} \right) \end{aligned} \quad (5)$$

### 3.2. 最尤法によるパラメータ推定

$\rho_g$  の推定方法は、Gordy & Heitfield [11] における最尤法を基礎とする．この方法においては、(4) 式のように各カテゴリに潜在変数の 1 ファクターモデルによるデフォルトの発生がモデル化された上で、各カテゴリのカテゴリ内共通ファクター  $X_{g,t}$  間に相関が存在しうる形となっている．

カテゴリ数  $G$  の第  $g$  カテゴリに属する  $t$  期初 ( $t = 1, \dots, T$ ) の正常債務者数を  $N_{g,t-1}$  とする． $t$  期中に観測されたデフォルト債務者数を  $D_{g,t}$  とする． $X_{g,t}$  は  $t$  に関して独立同一の標準正規分布  $N(0, 1)$  に従うとする．つまり、カテゴリ内共通ファクター  $X_{g,t}$  は時系列的な依存構造を持たないと仮定する． $t$  期において  $X_{g,t} = x_{g,t}$  の条件の下ではデフォルト事象は独立なので、デフォルト債務者数  $D_t$  は、成功確率  $p_g(x_{g,t})$ 、試行回数  $N_{g,t-1}$  の 2 項分布に従う．この 2 項分布  $b(D_{g,t}; N_{g,t-1}, p(x_{g,t}))$  の確率関数を  $f(D_{g,t}; N_{g,t-1}, p(x_{g,t}))$  とすれば (5) 式とともに次のようになる．

$$f(D_{g,t}; N_{g,t-1}, p(x_{g,t})) = \binom{N_{g,t-1}}{D_{g,t}} p(x_{g,t})^{D_{g,t}} (1 - p(x_{g,t}))^{N_{g,t-1} - D_{g,t}} \quad (6)$$

$\mathbf{X}_t = (X_{1,t}, X_{2,t}, \dots, X_{G,t})$  をカテゴリ数  $G$  分のカテゴリ内共通ファクター  $X_{g,t}$  を要素とする多変量正規分布に従う確率変数ベクトルとして、(6) 式を  $\mathbf{X}_t$  で積分し無条件化することで次の一般的な尤度関数を得る．ここで全カテゴリのファクター負荷とデフォルト閾値をそれぞれ要素とするベクトルを  $\boldsymbol{\rho} = (\rho_1, \rho_2, \dots, \rho_G)$  と  $\boldsymbol{\theta} = (\theta_{0,1}, \theta_{0,2}, \dots, \theta_{0,G})$ 、相関係数行列が  $G$  次の正方行列  $\boldsymbol{\Sigma}$  となる  $G$  次多変量正規分布の密度関数を  $\phi_{\boldsymbol{\Sigma}}(\mathbf{x})$ 、 $\mathbf{x} = (x_1, x_2, \dots, x_G)$  とした．

$$l(\boldsymbol{\rho}, \boldsymbol{\theta}) = \sum_{t=1}^T \ln \left\{ \int_{-\infty}^{\infty} \prod_{g=1}^G (f(D_{g,t}; N_{g,t-1}, p_g(x_{g,t}))) \phi_{\boldsymbol{\Sigma}}(\mathbf{x}) d\mathbf{x} \right\} \quad (7)$$

(7) 式を最大化することで、 $\boldsymbol{\rho}$ 、 $\boldsymbol{\theta}$ 、 $\boldsymbol{\Sigma}$  を最尤推定することが理論的には可能であるが、カテゴリ数が増えるにしたがって推定パラメータ数並びに計算負荷が増大するので、Gordy & Heitfield [11] においてもこの形式での推定は行っていない．

表 1: 各相関パラメータの意味

	定義	説明
カテゴリ内共通ファクター間相関	$\rho_0^2$	カテゴリ内共通ファクター $X_{g,t}$ 間の相関係数
カテゴリ内相関	$\alpha_g^2 + \beta_g^2 = \rho_g^2$	同じカテゴリ $g$ に属する債務者の潜在変数間の相関係数
カテゴリ間相関	$\alpha_g \alpha_h = \rho_g \rho_h \rho_0^2$	別のカテゴリ $g$ と $h$ に属する債務者の潜在変数間の相関係数

#### 4. 本論のモデル

複数カテゴリのデフォルト実績データで前節の最尤法による推定を行うために、ここで潜在変数の2ファクターモデル形式を基に相関パラメータ数を削減した尤度関数を定義する。

##### 4.1. 潜在変数の2ファクターモデル

時点  $t$  におけるカテゴリ  $g$  に属する債務者  $i$  の潜在変数  $V_{g,i,t} \simeq N(0, 1)$  について次のような2ファクターモデルを考える。

$$V_{g,i,t} = \alpha_g Y_t + \beta_g Z_{g,t} + \sqrt{1 - \alpha_g^2 - \beta_g^2} \varepsilon_{i,t} \quad (8)$$

ここで  $Y_t \simeq N(0, 1)$  は全体の変動を表すグローバルファクター、 $Z_{g,t} \simeq N(0, 1)$  はカテゴリ独自の要因をあらわすカテゴリ固有ファクター、 $\varepsilon_{i,t} \simeq N(0, 1)$  は債務者  $i$  独自の要因を表す債務者固有ファクターである。 $\alpha_g, \beta_g$  はカテゴリごとにファクターの加重和のウェイトを表す0以上1以下（このとき  $\alpha_g^2 + \beta_g^2 \leq 1$ ）の実数であり、 $i$  および  $t$  によらずカテゴリ  $g$  ごとに一定であるとする。

(9) 式において、0以上1以下の実数  $\rho_0, \rho_g$  により  $\alpha_g = \rho_g \rho_0, \beta_g = \rho_g \sqrt{1 - \rho_0^2}$  と置くと、

$$V_{g,i,t} = \rho_g \left( \rho_0 Y_t + \sqrt{1 - \rho_0^2} Z_{g,t} \right) + \sqrt{1 - \rho_g^2} \varepsilon_{i,t} \quad (9)$$

となる。これは、(4) 式の1ファクターモデルにおいて、 $X_{g,t} = \left( \rho_0 Y_t + \sqrt{1 - \rho_0^2} Z_{g,t} \right)$  としたものである。すなわち、カテゴリ内共通ファクター  $X_{g,t}$  に関する相関係数行列  $\Sigma$  について、

$$\Sigma = \begin{pmatrix} 1 & \rho_0^2 & \dots & \rho_0^2 \\ \rho_0^2 & 1 & \dots & \vdots \\ \vdots & \vdots & \ddots & \rho_0^2 \\ \rho_0^2 & \dots & \rho_0^2 & 1 \end{pmatrix} \quad (10)$$

と仮定したことになる。これは言い換えれば、あらゆるカテゴリの組み合わせにおいて、カテゴリ内共通ファクター  $X_{g,t}$  間の相関を  $\rho_0^2$  としたものである。またこのモデルにおいて、カテゴリ内相関（同じカテゴリ  $g$  に属する債務者の潜在変数の相関）は  $\alpha_g^2 + \beta_g^2 = \rho_g^2$ 、カテゴリ間相関（別のカテゴリ  $g$  と  $h$  に属する債務者の潜在変数の相関）は  $\alpha_g \alpha_h = \rho_g \rho_h \rho_0^2$  となる。以上を表1並びに図1にまとめた。

この2ファクター形式のデフォルトに関する潜在変数法において、 $Y_t = y_t, Z_{g,t} = z_{g,t}$  の下でのカテゴリ  $g$  に属する債務者の条件付きデフォルト確率を  $p_g(y_t, z_{g,t})$  とすると、(5) 式

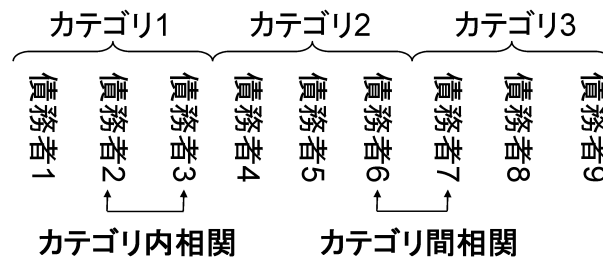


図 1: カテゴリ内相関  $\rho_g^2$  とカテゴリ間相関  $\rho_g \rho_h \rho_0^2$  の概念図

の導出と同様にして次のようになる。

$$\begin{aligned}
 p_g(y_t, z_{g,t}) &= P \left[ \varepsilon_{i,t} \leq \frac{\theta_g - \rho_g (\rho_0 Y_t + \sqrt{1 - \rho_0^2} \cdot Z_{g,t})}{\sqrt{1 - \rho_g^2}} \middle| Y_t = y_t, Z_{g,t} = z_{g,t} \right] \quad (11) \\
 &= G_2 \left( \frac{\theta_g - \rho_g (\rho_0 y_t + \sqrt{1 - \rho_0^2} \cdot z_{g,t})}{\sqrt{1 - \rho_g^2}} \right)
 \end{aligned}$$

#### 4.2. 2ファクター形式の尤度関数

以上の準備を踏まえ、最尤法でのパラメータ推定のために、カテゴリ内共通ファクター  $X_{g,t}$  が各々独立の場合（モデル1）、共有されている場合（モデル2）、相関が存在している場合（モデル3）のそれぞれで以下のように尤度関数を定める。潜在変数は全て標準正規分布に従っており、モデル3との比較では、モデル1は  $\rho_0 = 0$ 、モデル2は  $\rho_0 = 1$  を前提としている。

モデル1（カテゴリ内1ファクターモデル）:  $\rho_0 = 0$

カテゴリ  $g$  の共通ファクター  $X_{g,t}$  が他のカテゴリの共通ファクターと独立である場合対数尤度関数は次のようになる。

$$l(\rho, \theta) = \sum_{t=1}^T \sum_{g=1}^G \ln \left\{ \int_{-\infty}^{\infty} f(D_{g,t}; N_{g,t-1}, p_g(z_{g,t})) \phi(z_{g,t}) dz_{g,t} \right\} \quad (12)$$

$\phi(z)$  は標準正規分布の密度関数である。カテゴリ内共通ファクター  $X_{g,t}$  は各々独立であるので、(12) 式を最大化することは、各カテゴリにおける次の対数尤度関数、

$$l_g(\rho_g, \theta_g) = \sum_{t=1}^T \ln \left\{ \int_{-\infty}^{\infty} f(D_{g,t}; N_{g,t-1}, p_g(z_{g,t})) \phi(z_{g,t}) dz_{g,t} \right\}$$

をそれぞれ最大化することと同義である。

モデル2（グローバル1ファクターモデル）:  $\rho_0 = 1$

カテゴリ  $g$  の共通ファクター  $X_{g,t}$  が他と完全に依存し全てのカテゴリで共有される場合、つまり  $X_{1,t} = X_{2,t} = \dots = X_{G,t}$  であるとき、対数尤度関数は次のようになる。

$$l(\rho, \theta) = \sum_{t=1}^T \ln \left\{ \int_{-\infty}^{\infty} \prod_{g=1}^G (f(D_{g,t}; N_{g,t-1}, p_g(y_t))) \phi(y_t) dy_t \right\} \quad (13)$$

モデル3 (2ファクターモデル):  $0 \leq \rho_0 \leq 1$

カテゴリ  $g$  の共通ファクター  $X_{g,t}$  が他のカテゴリの共通ファクターと独立の場合、依存性がある場合を全て網羅すると、対数尤度関数は以下ようになる。

$$l(\boldsymbol{\rho}, \boldsymbol{\theta}) = \sum_{t=1}^T \ln \left\{ \int_{-\infty}^{\infty} \prod_{g=1}^G \left\{ \int_{-\infty}^{\infty} (f(D_{g,t}; N_{g,t-1}, p_g(y_t, z_{g,t}))) \phi(z_{g,t}) dz_{g,t} \right\} \phi(y_t) dy_t \right\} \quad (14)$$

(9) 式より、モデル1のファクターモデル形式は、 $V_{g,i,t} = \rho_g Z_{g,t} + \sqrt{1 - \rho_g^2} \varepsilon_{i,t}$  となり、同一カテゴリ内の債務者同士でのみ相関する1ファクター形式となる。モデル2は、 $V_{g,i,t} = \rho_g Y_t + \sqrt{1 - \rho_g^2} \varepsilon_{i,t}$  となり、属するカテゴリにかかわらず等しく相関する1ファクター形式である。モデル3は(9)式のようにカテゴリ内、カテゴリ間の相関を考慮できる2ファクター形式となっている。また  $\rho_0$  の値によってモデル1, モデル2を包含している。

以上に基づいて、対数尤度最大化計算をMATLABにより実装しパラメータ推定を可能とした。この詳細については付録C参照。

#### 4.3. 既存研究との差異

Gordy & Heitfield [11] では、(7) 式の形式での推定を提案したが、モデル3の形式での計算は実際には行っていない。カテゴリ分の次元の多変量正規分布関数の積分を要すること、各共通ファクター間の相関行列の推定が必要となることなどから計算の負担が重いことが主な理由であると考えられる。本論ではカテゴリ内共通ファクター  $X_{g,t}$  間の相関  $\rho_0^2$  を一定とする前提((10)式)を加えた上で、(14)式の形式で推定計算を可能とした。

Demey et al. [4] では(14)式と類似の形式のモデルでパラメータ数の削減の上推定を試みている。本論ではカテゴリ内共通ファクター  $X_{g,t}$  間の相関を  $\rho_0^2$  で一定としているのに対して、彼らのモデルでは  $X_{g,t}$  間の相関は一定とならない。これはカテゴリ間のアセット相関を一定値としている((9)式において、 $\alpha_g = \alpha_h$ ) 点の本論のモデルと相違しているためである。本論はGordy & Heitfield [11] を踏襲して、カテゴリ内共通ファクター間に依存性が存在する場合と、独立の場合、完全に共有される場合で各々推定を行い、最適なモデルを選択することを目的の一つとしている。(14)式の形式により、モデル1とモデル2はそれぞれモデル3の特殊形である  $\rho_0 = 0$  ( $\alpha_g = 0, \beta_g = \rho_g$ ),  $\rho_0 = 1$  ( $\alpha_g = \rho_g, \beta_g = 0$ ) の場合と位置づけることができ、モデル選択の議論は  $\rho_0$  の推定問題に帰着される<sup>8</sup>。

#### 4.4. カテゴリ選択とカテゴリ区分

カテゴリについては、業種、地域、企業規模、信用度指標など様々な選択肢が考えられる。本論では推定モデルの汎用性・一般性を保ちながら議論を進めるため、カテゴリを具体的に何とすべきかにはあえて立ち入らない。分散投資を効果的に行うという実務上の本来の目的から考えれば、カテゴリ間相関ができるだけ低く ( $\rho_0$  が最も小さく) なるようなカテゴリを選択し、これに従いカテゴリ内への集中を抑えた信用ポートフォリオとすることで、最大損失発生の可能性を低くするということとなる。

またカテゴリ選択に加えて、カテゴリ区分(どの水準でカテゴリを区分けするか)も重要な問題である。区分についても同様にカテゴリ間相関ができるだけ低くなるような区分を探し出すことを通して、分散効果を最大化していくことになる。

<sup>8</sup>また、Demey et al. [4] のモデルでは全てのファクター負荷を実数とするため二つの推定対象パラメータ間に大小関係の制約をつける必要があるが、その制約は現実での理由付けが困難である。一方(9)式で表される本論のモデルではそのような制約は不要である。



このように、本論のようなカテゴリを考慮した相関係数の推定を様々なデータに適用することによって、分散投資の観点から有用な情報が得られると考えられる。しかしこのような実証に基づく検討のためには、十分なデフォルト履歴情報とカテゴリ情報を同時に揃えることが必要であり、データ整備については少なからず労力を要するであろう。

## 5. モンテカルロ法による推定誤差の評価

本節では、人工的に作られた仮想のデフォルト数・債務者数データで最尤法により推定計算を行う試行を多数回行い、パラメータ推定値の分布と人工データの元となった真のパラメータ値を比較することで、モデル1~3による推定がどのような誤差・バイアスをもたらすのかを検証する。

### 5.1. 実験方法

カテゴリ内共通ファクター負荷  $\rho$ 、デフォルト閾値  $\theta$ 、そしてカテゴリ内共通ファクター  $X_{g,t}$  間の相関  $\rho_0^2$  を所与のものとして、(11)式に沿って確率変数部分に乱数を当てはめデフォルト数をシミュレートすることで、カテゴリ内の債務者数  $N_{g,t-1}$  からカテゴリ毎のデフォルト数  $D_{g,t}$  のサンプルを  $T$  期分ランダムに人工的に作った。この乱数によるデフォルト数シミュレート計算の詳細については付録Dを参照のこと。

このデータを用い、モデル1~3の対数尤度関数(12)式、(13)式、(14)式による最尤法でパラメータ推定を行う<sup>9</sup>。推定に誤差がなければ所与とした真の値  $\rho, \theta, \rho_0$  と同じ値が推定値として常に得られるはずであるが、実際にはばらつきが発生する。推定値の真の値からの偏差を推定誤差として、ランダムに作ったデータによる推定を多数回繰り返してこれを得ることで、この推定方法のもたらす誤差の特性を見る。

### 5.2. 実験結果と考察

対象は3カテゴリ。カテゴリ1, 2, 3の真のカテゴリ内ファクター負荷はそれぞれ  $\rho_1 = 0.15, \rho_2 = 0.10, \rho_3 = 0.05$ 、真の閾値は  $\theta_g = -3.3$ 、対象期間は  $T = 60$ 、債務者数は  $N_{g,t-1} = 2^{16}$  として1000回の推定値計算のシミュレーションを行った<sup>10</sup>。

最初に表2が、それぞれのモデルの前提に沿って、モデル1, 2, 3において  $\rho_0 = 0, 1, \sqrt{0.5}$  としたときの1000回試行の要約である。このときカテゴリ内共通ファクター  $X_{g,t}$  間の相関  $\rho_0^2$  はそれぞれ0, 1, 0.5となる。表中には  $\rho, \theta, \rho_0$  の推定値の平均・標準偏差と、推定のパフォーマンスを計る目的で真の値からの根平均二乗誤差を出した<sup>11</sup>。推定値の平均はいずれのモデルにおいても真の値に非常に近く、当モデルによる推定計算は基本的に適切に行われている。

表2の各モデルの前提どおりの  $\rho_0$  における推定では、モデル間で比較して推定値の平均、誤差ともに目立った差はない。このことから、カテゴリ内共通ファクター  $X_{g,t}$  間の相関がモデルの前提どおりである限りにおいて、モデル間の優劣は存在しない。また閾値について

<sup>9</sup>実際の最適化計算において計算が終了しない試行が稀であるが発生した(主にモデル1)。本論中の集計ではこのような試行は含まれていない。Gordy & Heitfield [11]にも同様の観測があり、特に1カテゴリだけではデフォルトの情報が少ないことが原因であると考えられるが、この原因については今後の課題とする。

<sup>10</sup>Gordy & Heitfield [11]では、ある限定された前提条件上ではあるが、 $T$ が短い期間でしか取れない場合に、誤差が大きくなるうえに、推定値に真の値からの下方バイアスが発生する可能性があることを示唆している。

<sup>11</sup>この推定実験が6節のような実データでの推定と違うのは、推定すべき「真の値」があらかじめ正確にわかっている点である。推定モデルのパフォーマンスを計るのに、推定値の「真の値」からの乖離を計る以上によい方法はないと考え根平均二乗誤差を用いている。実データでの推定のように「真の値」が不明である場合には、6節で用いているAICなどの尤度関連指標がモデルのもっともらしさの指標として有用であろう。

表 2: ファクター負荷推定のモンテカルロシミュレーション 1. カテゴリ内共通ファクター間相関がモデルの前提と同じ場合

モデル1	$\rho_0$	$\rho_1$	$\rho_2$	$\rho_3$	$\theta_1$	$\theta_2$	$\theta_3$	
真の値		0	0.15	0.10	0.05	-3.3	-3.3	-3.3
平均	-	0.1479	0.0976	0.0486	-3.3008	-3.3005	-3.3002	
標準偏差	-	0.01537	0.01202	0.00918	0.02252	0.01493	0.00906	
根平均二乗誤差	-	0.01552	0.01226	0.00929	0.02253	0.01494	0.00906	
推定値0の割合	-	0%	0%	0%	0%	0%	0%	
モデル2	$\rho_0$	$\rho_1$	$\rho_2$	$\rho_3$	$\theta_1$	$\theta_2$	$\theta_3$	
真の値		1	0.15	0.10	0.05	-3.3	-3.3	-3.3
平均	-	0.1491	0.0998	0.0497	-3.3006	-3.3004	-3.3000	
標準偏差	-	0.01536	0.01172	0.00832	0.02143	0.01489	0.00924	
根平均二乗誤差	-	0.01538	0.01172	0.00832	0.02144	0.01490	0.00924	
推定値0の割合	-	0%	0%	0%	0%	0%	0%	
モデル3	$\rho_0$	$\rho_1$	$\rho_2$	$\rho_3$	$\theta_1$	$\theta_2$	$\theta_3$	
真の値	0.7071	0.15	0.10	0.05	-3.3	-3.3	-3.3	
平均	0.7086	0.1475	0.0977	0.0485	-3.3007	-3.3003	-3.3004	
標準偏差	0.07731	0.01608	0.01200	0.00948	0.02169	0.01428	0.00880	
根平均二乗誤差	0.07733	0.01628	0.01223	0.00960	0.02171	0.01428	0.00880	
推定値0の割合	0%	0%	0%	0%	0%	0%	0%	

は、いずれのモデルにおいても、真の値に非常に近い推定値が得られており、モデル選択は閾値の推定値に大きく影響しないと言える。

次に表3は、モデル1, 2において前のモデル3と同様の  $\rho_0 = \sqrt{0.5}$  とした時の1000回試行の要約である。表2の分析と違いカテゴリ内共通ファクター  $X_{g,t}$  間の相関  $\rho_0^2$  がモデル1, 2の前提に反している。このとき、モデル2の推定値の平均が真の値より小さく下方バイアスが発生し、誤差も他のモデルに比べて大きくなっている。対してモデル1では、モデル3と比較しても推定値の平均、誤差ともに遜色が見られない。モデル1では、カテゴリ内共通ファクター  $X_{g,t}$  間の相関が前提に反し0でない場合であっても、カテゴリ内相関のみであれば比較的うまく推定できている。

モデル2では全てのカテゴリで同一（完全相関）のカテゴリ内共通ファクター  $X_{g,t}$  があることを前提としている。この前提に反し真の値としての  $\rho_0$  が1より小さくカテゴリ内共通ファクター間の相関が部分的である場合には、完全相関するそれぞれのカテゴリのカテゴリ内共通ファクターの変動で各カテゴリのデフォルト実績を同時に説明しづらくなり、結果カテゴリ内共通ファクターへの依存度が低い、つまり  $\rho_g$  が低いという推定結果になるものと考えられる。一方でモデル1ではもともと無相関のカテゴリ内共通ファクターを想定しているため、 $\rho_0$  やカテゴリ間相関が実際どうであったとしても、あるカテゴリのデフォルト実績のみを単独で用いて推定したカテゴリ内相関  $\rho_g$  の推定値には影響せず（表1）、この実験のように実際にカテゴリ内共通ファクター間の相関が部分的であったとしても  $\rho_g$  推定値にバイアスがかからず、結果がモデル3に結果が近くなるものと考えられる。

最後に表4は、債務者数を  $N_{g,t-1} = 2^{13}$  とした場合の1000回試行の要約である。真の閾値は変わっていないので、債務者数が減少すればデフォルト債務者の発生数は少なくなる。結果に見られるように、このようなデータを用いた推定では推定値の平均が真の値より下がる下方バイアスが発生し、推定値の誤差が大きくなる傾向がある<sup>12</sup>。特にモデル1につい

<sup>12</sup>債務者数が少ない場合、デフォルト率が低い場合のこのような問題は Gordy & Heitfield [11], Demey et al.

表 3: ファクター負荷推定のモンテカルロシミュレーション 2 . カテゴリ内共通ファクター間相関を全て  $\rho_0 = \sqrt{0.5}$  とした場合

モデル1	$\rho_0$	$\rho_1$	$\rho_2$	$\rho_3$	$\theta_1$	$\theta_2$	$\theta_3$
真の値	0.7071	0.15	0.10	0.05	-3.3	-3.3	-3.3
平均	-	0.1474	0.0982	0.0481	-3.3028	-3.3010	-3.3010
標準偏差	-	0.01554	0.01161	0.00960	0.02166	0.01464	0.00911
根平均二乗誤差	-	0.01576	0.01176	0.00979	0.02183	0.01467	0.00917
推定値0の割合	-	0%	0%	0%	0%	0%	0%
モデル2	$\rho_0$	$\rho_1$	$\rho_2$	$\rho_3$	$\theta_1$	$\theta_2$	$\theta_3$
真の値	0.7071	0.15	0.10	0.05	-3.3	-3.3	-3.3
平均	-	0.1474	0.0757	0.0307	-3.2936	-3.2916	-3.2966
標準偏差	-	0.02281	0.01853	0.01083	0.02697	0.01514	0.00936
根平均二乗誤差	-	0.02296	0.03060	0.02214	0.02772	0.01730	0.00996
推定値0の割合	-	0%	0%	0%	0%	0%	0%
モデル3	$\rho_0$	$\rho_1$	$\rho_2$	$\rho_3$	$\theta_1$	$\theta_2$	$\theta_3$
真の値	0.7071	0.15	0.10	0.05	-3.3	-3.3	-3.3
平均	0.7086	0.1475	0.0977	0.0485	-3.3007	-3.3003	-3.3004
標準偏差	0.07731	0.01608	0.01200	0.00948	0.02169	0.01428	0.00880
根平均二乗誤差	0.07733	0.01628	0.01223	0.00960	0.02171	0.01428	0.00880
推定値0の割合	0%	0%	0%	0%	0%	0%	0%

て、カテゴリ 3 の  $\rho_3$  の推定値がゼロ<sup>13</sup>となる試行が 2 割以上発生しており、推定値の平均も真の値を下回る度合いが大きく、モデル 3 よりも誤差の大きい結果となっている。モデル 3 は最も良い結果である。

デフォルト数がゼロ以上の整数値しか取らないので、債務者数が少ない場合はデフォルト率がゼロでなくとも、デフォルト実績にゼロが出現する可能性が高くなる。このようなデフォルト数が一定値をとりやすいデータにおける  $\rho_g$  推定値はゼロに近くなると考えられる。これは (5) 式においてデフォルト確率の不確実性 (カテゴリ内共通ファクター  $X_{g,t}$  への依存度) は  $\rho_g$  が小さいほど小さくなり、 $\rho_g$  がゼロとなればデフォルト確率が (カテゴリ内共通ファクターの値にかかわらず) 一定値になることで分かる。モデル 1 のほうがモデル 3 よりこの問題が深刻なのは、モデル 3 の場合は他カテゴリの実績もあわせれば全てのカテゴリでデフォルト実績がゼロとなる可能性は小さくなるので、 $\rho_0$  については推定が適切に行われる可能性が高く、これが各カテゴリの  $\rho_g$  の推定値に影響を及ぼすのに対して、モデル 1 は  $\rho_g$  推定値が単一カテゴリのみのデフォルト実績のみに基づいており、デフォルト数実績が比較的ゼロを取りやすいためではないかと考えられる。

以上の 3 種類の実験から、推定誤差 (推定値ゼロの発生率)、 $\rho_g$  推定値の平均の真の値からの乖離度合いでみれば、モデル 3 (2 ファクターモデル) が最も推定パフォーマンスが良く、モデル 1 (カテゴリ内 1 ファクターモデル) がそれにやや劣り、モデル 2 (グローバル 1 ファクターモデル) は大幅に劣っているといえることができる。特に債務者数が少なくデフォルトサンプルが十分でない場合には、カテゴリ間相関  $\rho_g \rho_h \rho_g^2$  に関心がなく、カテゴリ内相関  $\rho_g^2$  のみを推定すれば十分である場合であっても、他のカテゴリの情報を同時に使うモデル 3 を用いるのが最も推定誤差を小さくする推定方法である。

[4] でも指摘されている。  
<sup>13</sup>計算上は  $10^{-5}$ 。

表 4: ファクター負荷推定のモンテカルロシミュレーション 3. 債務者数を  $2^{16}$  から  $2^{13}$  と減らした場合

モデル1	$\rho_0$	$\rho_1$	$\rho_2$	$\rho_3$	$\theta_1$	$\theta_2$	$\theta_3$
真の値	0.7071	0.15	0.10	0.05	-3.3	-3.3	-3.3
平均	-	0.1458	0.0917	0.0393	-3.3000	-3.3005	-3.2994
標準偏差	-	0.02858	0.03210	0.03441	0.02743	0.02233	0.01996
根平均二乗誤差	-	0.02890	0.03316	0.03602	0.02743	0.02234	0.01997
推定値0の割合	-	0%	1%	21%	0%	0%	0%
モデル2	$\rho_0$	$\rho_1$	$\rho_2$	$\rho_3$	$\theta_1$	$\theta_2$	$\theta_3$
真の値	0.7071	0.15	0.10	0.05	-3.3	-3.3	-3.3
平均	-	0.1408	0.0671	0.0296	-3.2999	-3.3004	-3.3007
標準偏差	-	0.03095	0.03453	0.02398	0.02718	0.02247	0.02017
根平均二乗誤差	-	0.03229	0.04768	0.03147	0.02718	0.02248	0.02018
推定値0の割合	-	1%	6%	17%	0%	0%	0%
モデル3	$\rho_0$	$\rho_1$	$\rho_2$	$\rho_3$	$\theta_1$	$\theta_2$	$\theta_3$
真の値	0.7071	0.15	0.10	0.05	-3.3	-3.3	-3.3
平均	0.7241	0.1464	0.0933	0.0452	-3.2998	-3.3003	-3.3002
標準偏差	0.23456	0.02805	0.03080	0.02966	0.02701	0.02318	0.01875
根平均二乗誤差	0.23517	0.02828	0.03153	0.03005	0.02701	0.02318	0.01875
推定値0の割合	0%	0%	1%	12%	0%	0%	0%

## 6. 実際のデータによるパラメータ推定

前節の分析結果を踏まえて、本論のアセット相関の推定モデルを実際のデータに適用する<sup>14</sup>。この推定法は、中小企業向けや住宅ローンなど個人向けローン等、金額的に分散した市場性の低い信用プールに対して適しており、このような信用プールに対する相関構造並びに分散効果の実証が可能となる。

### 6.1. 推定計算の概要

対象データのうち、デフォルト数  $D_{g,t}$  として、東京商工リサーチの全国企業倒産の推移データ<sup>15</sup>（負債総額 1,000 万円以上倒産集計）から月次の資本金別企業倒産数データ<sup>16</sup>を、債務者数  $N_{g,t-1}$  として国税庁調べの税務統計による普通法人数<sup>17</sup>の資本金別法人数の年次データを月次に線形補完したものをを用いた<sup>18</sup>。

本論では、カテゴリ区分として資本金を用いて推定を行う。倒産数、普通法人数それぞれのデータセットで取得可能な資本金区分を互いに揃え、カテゴリ 1: 資本金 1 億円以上、カテゴリ 2: 資本金 5 千万円以上 1 億円未満、カテゴリ 3: 資本金 1 千万円以上 5 千万円未満、カテゴリ 4: 資本金 500 万円以上 1000 万円未満の 4 カテゴリとした<sup>19</sup> <sup>20</sup>。分析対象のデータセッ

<sup>14</sup>本論では中小企業の資本金別倒産データを用いるが、カテゴリごとの債務者数とデフォルト数の時系列データがあれば一般的にモデルは適用可能である。

<sup>15</sup>(株) 東京商工リサーチ [25] 並びに同社ホームページ <http://www.tsr-net.co.jp/topics/zenkoku/> より。

<sup>16</sup>損失を蒙るのは必ずしも借り手の倒産に限らず、リスケジュールや債務免除なども対象に含まれるべきであるが、データの制約もあり、本論では損失を蒙るイベントは倒産（企業の法的処理）と同義という前提。

<sup>17</sup>国税庁ホームページ <http://www.nta.go.jp/category/toukei/tokei.htm> より。

<sup>18</sup>これらのデータを用いたデフォルト率（倒産発生率）は、中小企業白書 [3] や東京商工リサーチ [25] などで発表されており一定の指標性が認められている。しかし厳密には、別々の情報元から得たデータでデフォルト率を得ており、調査対象企業の不一致、調査時点の不一致、などの問題がありうる。

<sup>19</sup>カテゴリ区分は結果に直接的に影響するために本来は慎重な吟味が必要であるが、今回の区分はデータ制約上の理由によって採用されている。

<sup>20</sup>資本金 500 万円未満の倒産発生数もデータとしては存在するものの、東京商工リサーチ集計の企業倒産の定義が負債総額 1 千万円以上としているため使用しなかった。

トを1982年7月から2002年7月までの、のべ241ヶ月分とした<sup>21</sup>。図2が法人数の、図3が倒産数の推移である。

計算方法としては、上記データの120ヶ月分を1セットとし( $T = 120$ )、対象期間を1ヶ月ずつずらしながら、計122セットについて、モデル1~3それぞれでファクター負荷並びにデフォルト閾値 $\rho, \theta, \rho_0$ の推定を行った。最適化計算において、推定対象の $\rho_g$ の初期値によって得られる最適解がわずかにわずかに違う場合があった。その違いは論旨に影響を与えない水準ではない。しかし尤度関数の性質が明らかでないこともあり、局所解をできるだけ避けるために $\rho_g$ の初期値を $\rho_g = 0.05, 0.10, 0.15$ 、 $t - 1$ 期の推定結果と4種類変えて計算を行い、目的関数の値が最小のものを推定量とした<sup>22</sup>。

## 6.2. 推定結果

図4、図5、図6が、それぞれモデル1、2、3におけるカテゴリ内共通ファクター負荷 $\rho_g$ の推定値である。図7が、カテゴリ内共通ファクター間相関 $\rho_0^2$ を表す $\rho_0$ の推定値である。前述のとおり、カテゴリ内相関係数はいずれのモデルにおいても $\rho_g^2$ 、カテゴリ間相関係数は、モデル1(カテゴリ内1ファクターモデル)の場合前提上0、モデル2(グローバル1ファクターモデル)は $\rho_g \rho_h$ 、モデル3(2ファクターモデル)は $\rho_g \rho_h \rho_0^2$ となる。

5節のシミュレーションにおいて、推定値の真の値からの標準偏差(推定誤差)が0.01~0.03のオーダーであった一方で、この現実のデータによる $\rho_g$ 推定値は時期により変動あるもののほとんどにおいて0.1程度のオーダーとなっている。5節と本節では全く同じ推定計算を用いているので、本節の推定が多数セットで行われており、このオーダーの差が常に発生していることを鑑みれば、本節における推定対象において $\rho_g$ の真の値がゼロであるという蓋然性は低い<sup>23</sup>。いずれのモデルによる推定においても $\rho_g$ の真の値はゼロではなく、現実にカテゴリ内相関は存在していると考えられる。

一方で $\rho_0$ 推定値は十分大きな値が出ており、カテゴリ内共通ファクター $X_{g,t}$ 間の相関は高いと言える。加えて対象期間の当初(82年~92年)は1に近い値が得られているが、30セット目付近(85年~94年)から徐々に下がり始め、直近(92年~02年)では0.6程度となっており、時期によって大きな差がある。特に $\rho_0$ が下がった最近の時期においては、カテゴリ間相関がカテゴリ内相関に対してより小さくなっており、分散効果がより強くなったといえる。この結果をより積極的に解釈しようとするならば、かつては(資本金の差で表される)企業規模が違っていてもその信用力の変化は違いが少なかったが、バブル期以降の最近ではカテゴリ毎の要因同士の相関が弱まり信用力の変動が企業規模によってより強く特徴づけられるようになった、ということもできよう。

いずれのモデルにおいても $\rho_g$ 推定値は80セット目付近(88年~98年)をピークにしており、その後下降した後また徐々に上がって直近に至っている。30セット目付近(85年~94年)以降からカテゴリ1~カテゴリ4の $\rho_g$ 推定値の序列は安定しており、資本金区分で資本金が大きいほどカテゴリ内相関が高くなっている。中小企業のアセット相関は高々1~3%程度であり<sup>24</sup>、また(資本金規模が企業規模を代理しうると考えれば)企業規模が小さいほど

<sup>21</sup>負債総額、資本金区分については名目ベースによる区切りであるため、厳密に言えば、比較的長い期間においては貨幣価値の変動による実質的な集計対象のずれがある可能性がある。

<sup>22</sup> $\theta_g$ については、定義上は任意の実数となりうるものの、観測上のデフォルト率からモデル上推定された平均のデフォルト率が大きくかけ離れることは考えにくく、既存研究でもそのような報告がされている。計算の高速化のため、初期値を $\hat{\theta}_g = \Phi^{-1}(\sum D_{g,t} / \sum N_{g,t-1})$ 、この $\pm 0.5$ に収まることを制約条件とした。

<sup>23</sup>より厳密な議論のためには、5節のシミュレーション実験を真の値をゼロとして行い、その誤差分布と本節の推定値(多数セット分)との乖離度に関する分析を行うことが考えられる。

<sup>24</sup>アセット相関の絶対的な水準については今後様々なデータによる実証研究を通して議論されるべきであると

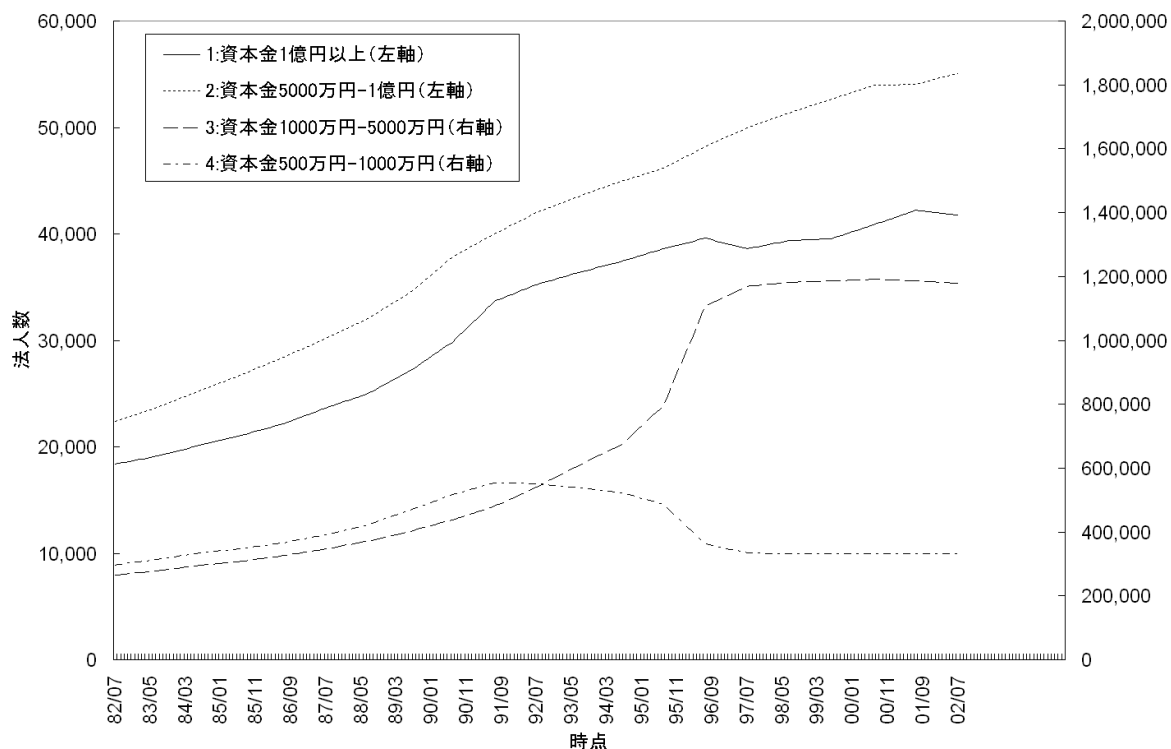


図 2: 法人数データ

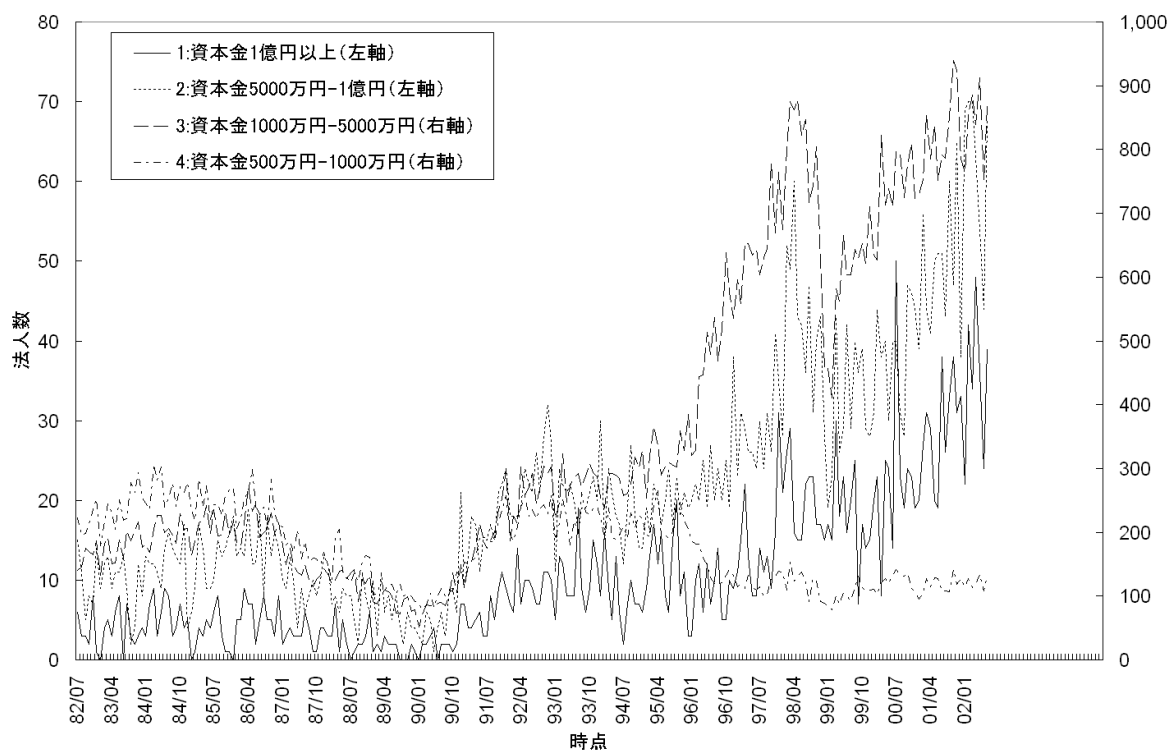


図 3: 倒産数データ

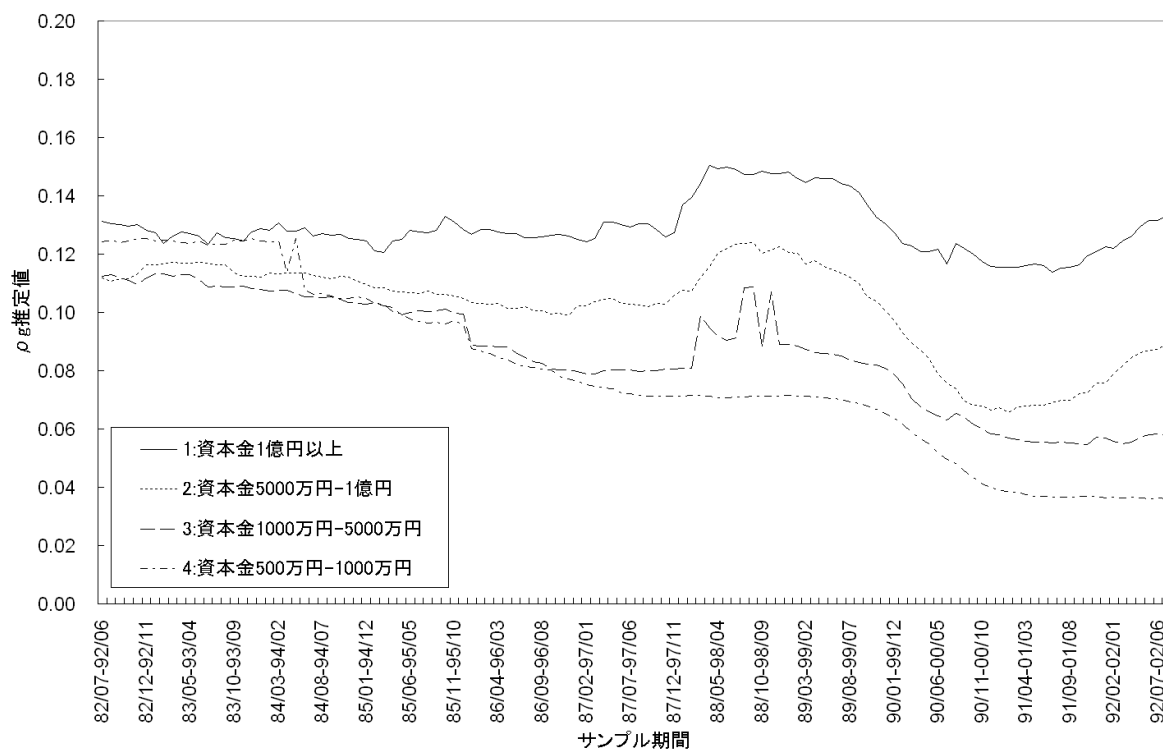


図 4:  $\rho_g$  推定値 (モデル1)

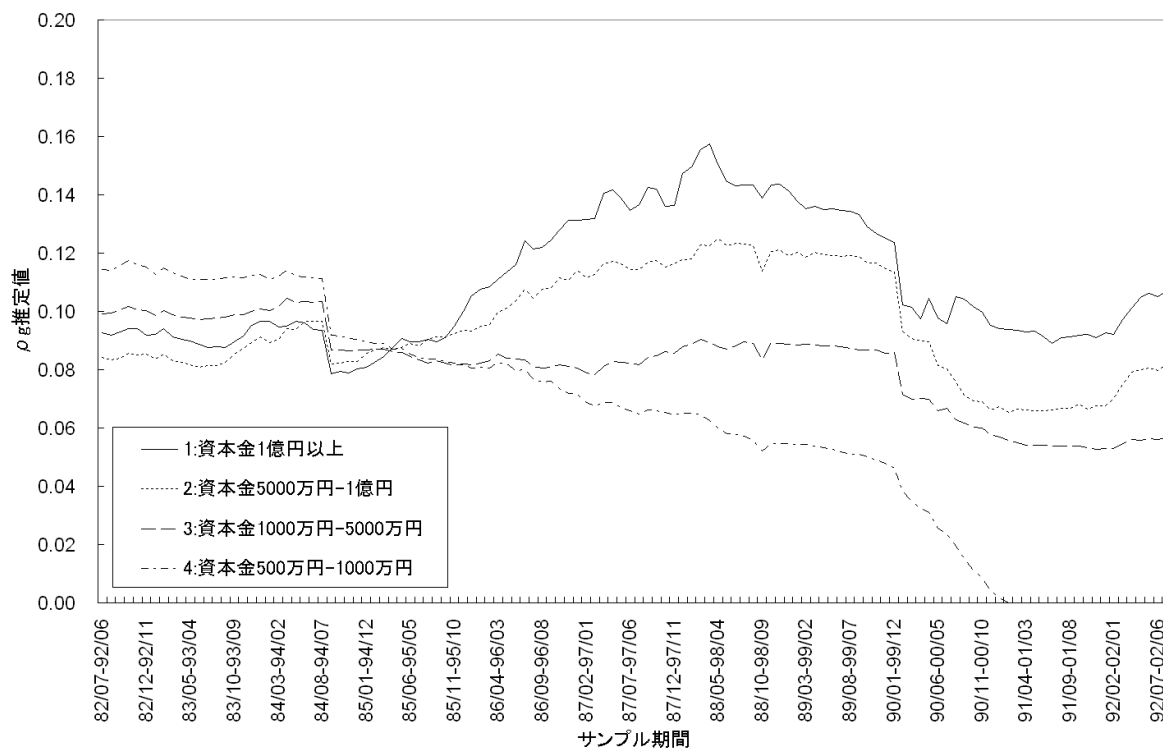


図 5:  $\rho_g$  推定値 (モデル2)

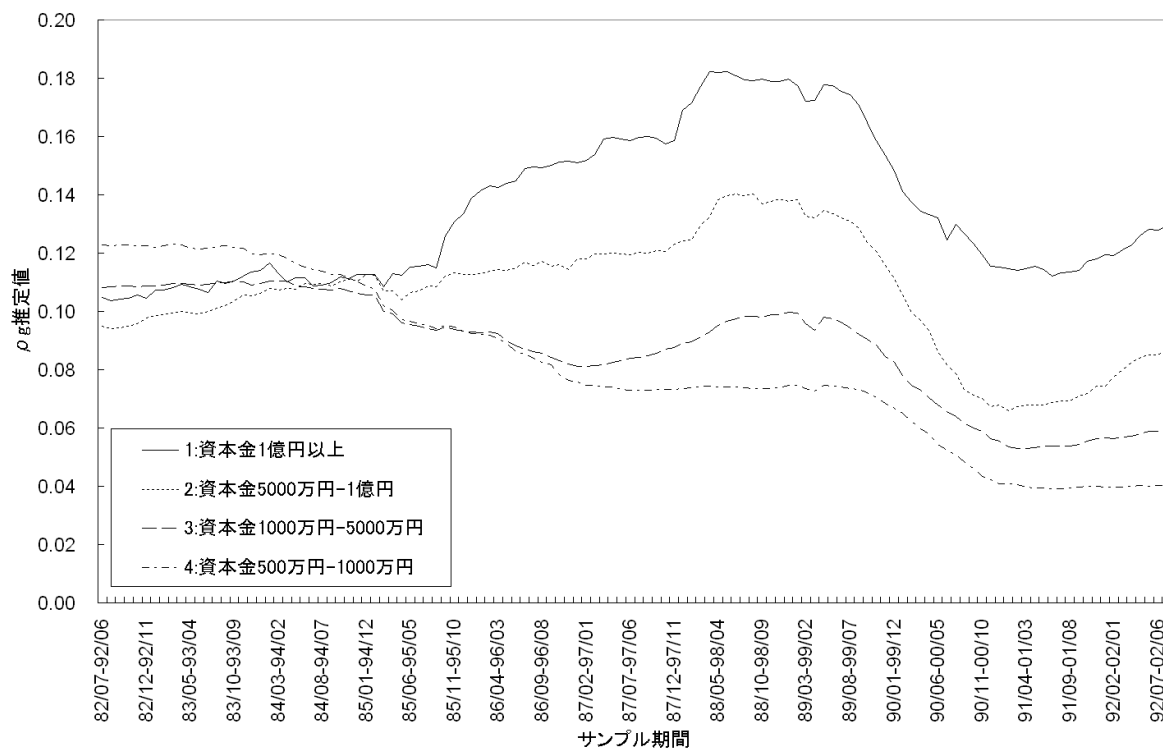


図 6:  $\rho_g$  推定値 (モデル3)

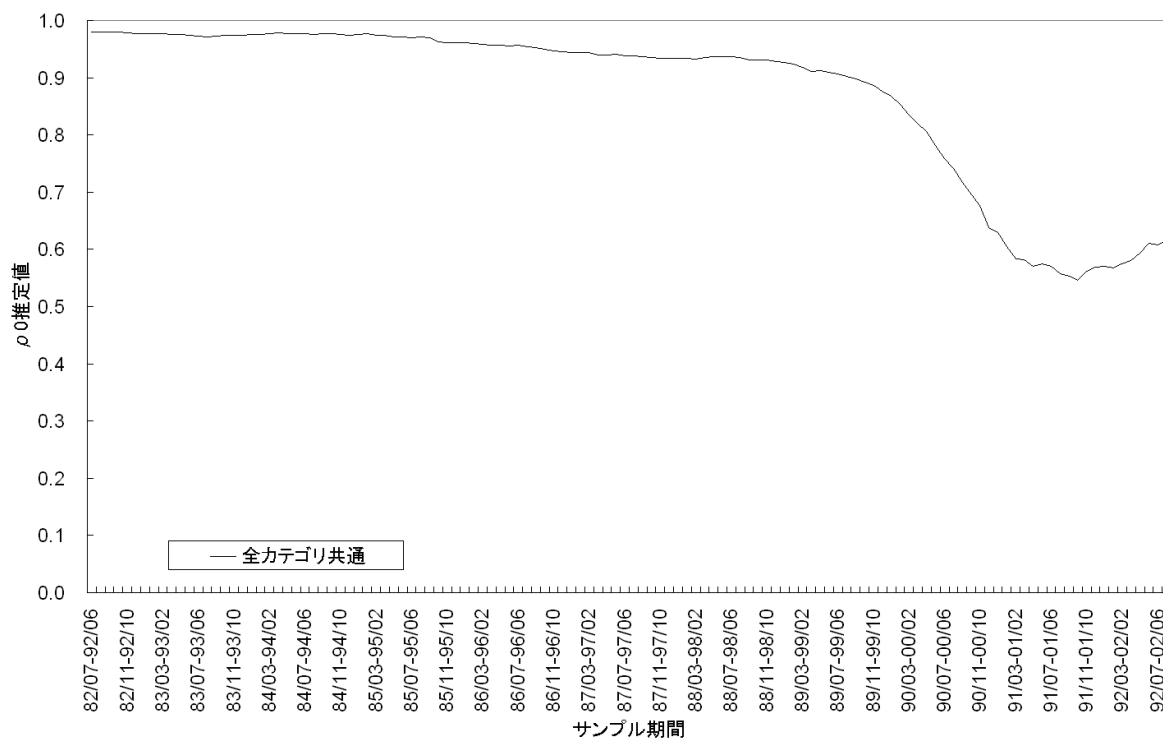


図 7:  $\rho_0$  推定値 (モデル3)



アセット相関は低くなるとした Dietsch & Petey [7] と整合的な結果である。

$\rho_g$  推定値の全体的な水準はモデル3, 1, 2の順で高くなっている。特にモデル3では, カテゴリ1, 2の推定値が他より高い期間がある。前節の実験において, カテゴリ内共通ファクター  $X_{g,t}$  間に部分的な相関がある場合には, 特に債務者数の少ない場合に推定値の下方バイアスが起こるが, その度合いはモデル3, 1, 2の順で小さかった。このモデルのバイアスが本推定計算においても発生していると考えられる。 $\rho_0$  推定値が低くなった95セット目付近(90年~00年)以降モデル2のカテゴリ4は推定値がほぼゼロとなっている。前節の実験においてカテゴリ内共通ファクター間に相関がある場合はモデル2の推定誤差は大きくなり推定値に下方バイアスが大きく出る結果となっており(表3), 本推定計算においても同様の理由によりゼロとなっていると考えられる。

モデル1とモデル3の推定結果が近い部分がある。特に $\rho_0$ の推定値が低くなった95セット目付近(90年~00年)以降はカテゴリによらずモデル1とモデル3の $\rho_g$ 推定値が近くなっているが, これは実際の $\rho_0$ がモデル1の前提 $\rho_0 = 0$ により近く, そもそもモデル1とモデル3の差が小さいことが理由であると考えられる。加えて, 債務者数の少ないカテゴリ1, 2における推定値がモデル3より低い一方で, 債務者数の多いカテゴリ4でモデル1と3の $\rho_g$ 推定値が期間を通じて非常に近くなっている。これはカテゴリ間相関の推定値がカテゴリ内相関に影響しないために(表1)債務者数の十分な場合はモデル1と3の $\rho_g$ 推定値の差は小さいという5節の実験結果を反映していると言える。

図8, 図9, 図10は, それぞれモデル1, 2, 3におけるデフォルトに関する潜在変数法における閾値 $\theta_g$ ((1)式)の推定値である。120ヶ月の平均デフォルト率から計算される閾値(図11)とそれぞれの推定値を比較すると, 非常に近い値となっている。モデル3は最もその差が小さく, モデル1もそれに近い。一方モデル2については閾値の小さいカテゴリの推定値に非連続な部分が観察されるが, 前節の実験での結果と同様, 閾値の推定値は相関の推定値に比較するとモデルにかかわらず安定しており, モデル選択の影響を受ける度合いは小さいといえる。

### 6.3. モデル選択

ここでモデル選択の基準として各モデルの最大対数尤度におけるAIC(赤池情報量規準)を計算した<sup>25</sup>。結果が図12である。全ての期間を通してモデル3が最小AIC推定値となっている。つまり, この分析対象データにおいては, パラメータをひとつ増やした点を考慮してもモデル3(2ファクターモデル)が最適である。図7とあわせてみると,  $\rho_0$ の推定値が1に近い時期においてはモデル2のほうがモデル1よりAICが小さく,  $\rho_0$ の推定値が1より低い時期にはモデル1のほうがモデル2よりAICが小さくなっている。モデル1と2はモデル3のそれぞれ $\rho_0 = 0$ ,  $\rho_0 = 1$ の場合に対応しており, 実際の $\rho_0$ がモデル1, 2の仮定するところと近い場合には当てはまりがモデル3に近くなっている。

モデル3の欠点は尤度計算((14)式)において積分が2重になっており計算負荷が非常に高いことである。カテゴリ内相関 $\rho_g^2$ のみが関心事である場合, もし $\rho_0$ が1に近いということが既知であればモデル2(グローバル1ファクターモデル)が簡便な代用となりうる。しかし本節で見たように, 現実においては $\rho_0$ は0でも1でもなく, また時期によって変動し

考えるが, 本論での分析対象が最も分散の効いていると思われる全国の企業を網羅したものであるため, 現実の(何らかの偏りのある)評価対象に比して低い推定値が得られていると考えるのが自然であろう。

<sup>25</sup>AICは対数尤度関数である(12)式, (13)式, (14)式による推定計算で得られた最大対数尤度と各モデルの自由パラメータ数(モデル1と2は8, モデル3は9)から $-2 \times (\text{最大対数尤度} - \text{自由パラメータ数})$ として得られる。最尤法におけるAICでのモデル選択については坂元他[21]等。

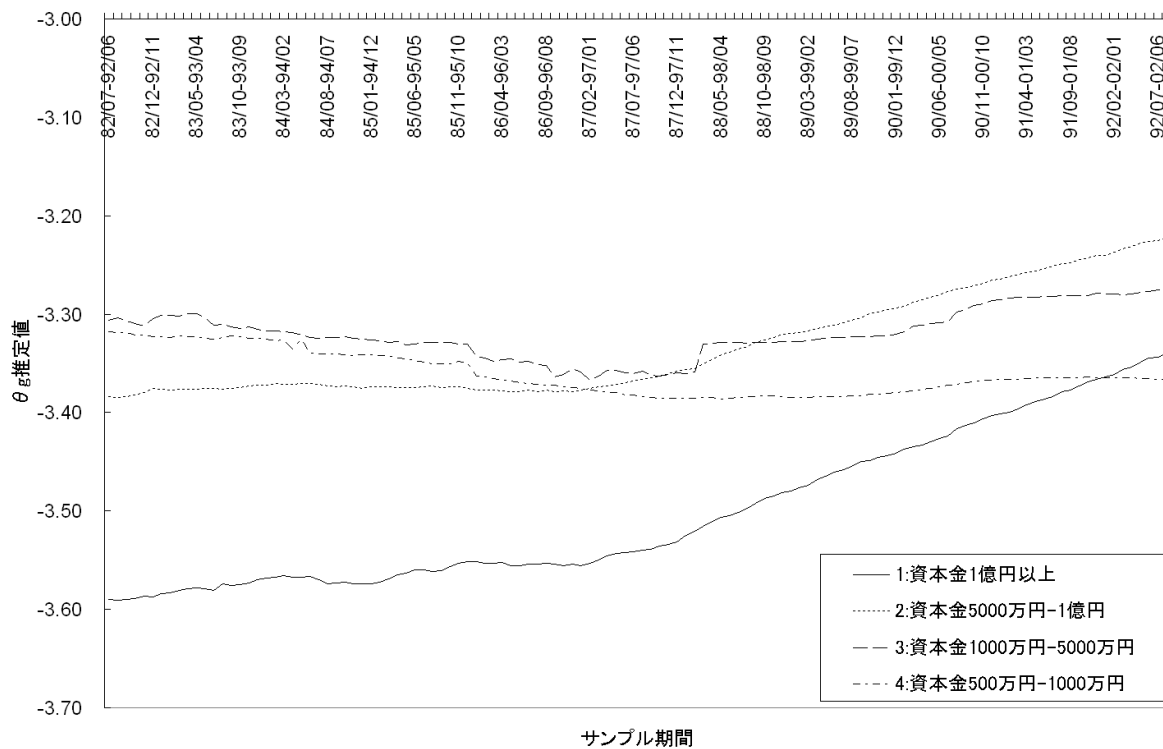


図 8:  $\theta_g$  推定値 (モデル1)

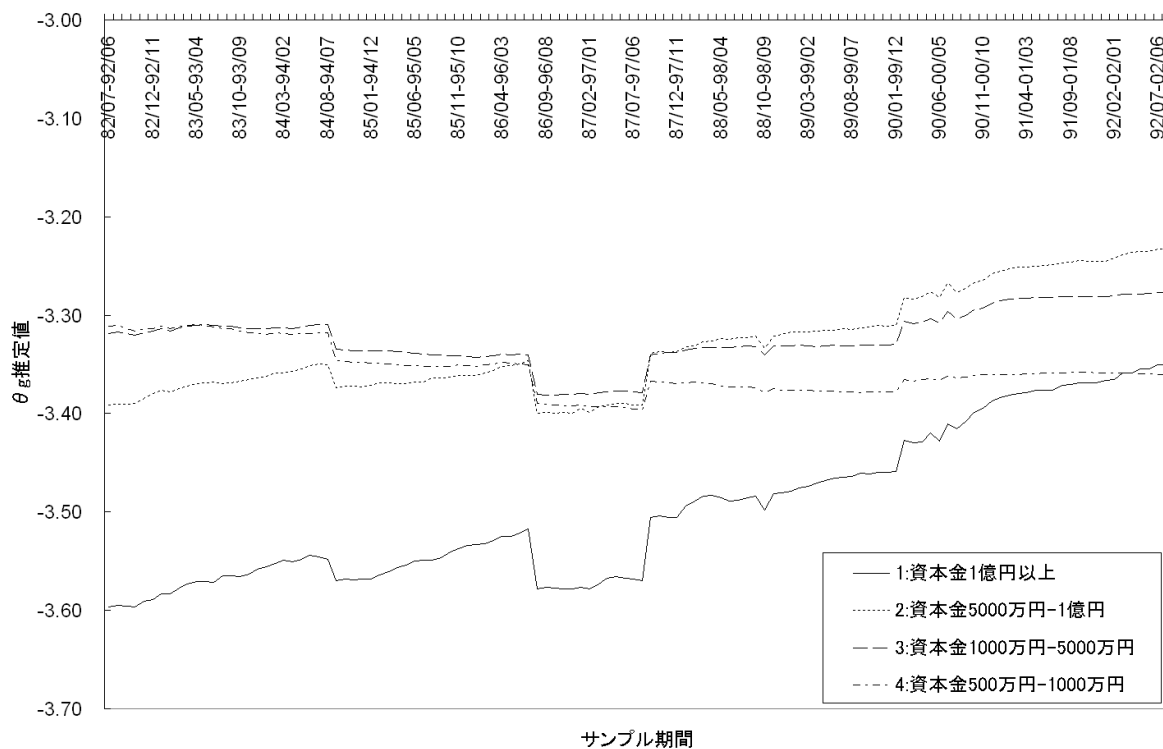


図 9:  $\theta_g$  推定値 (モデル2)

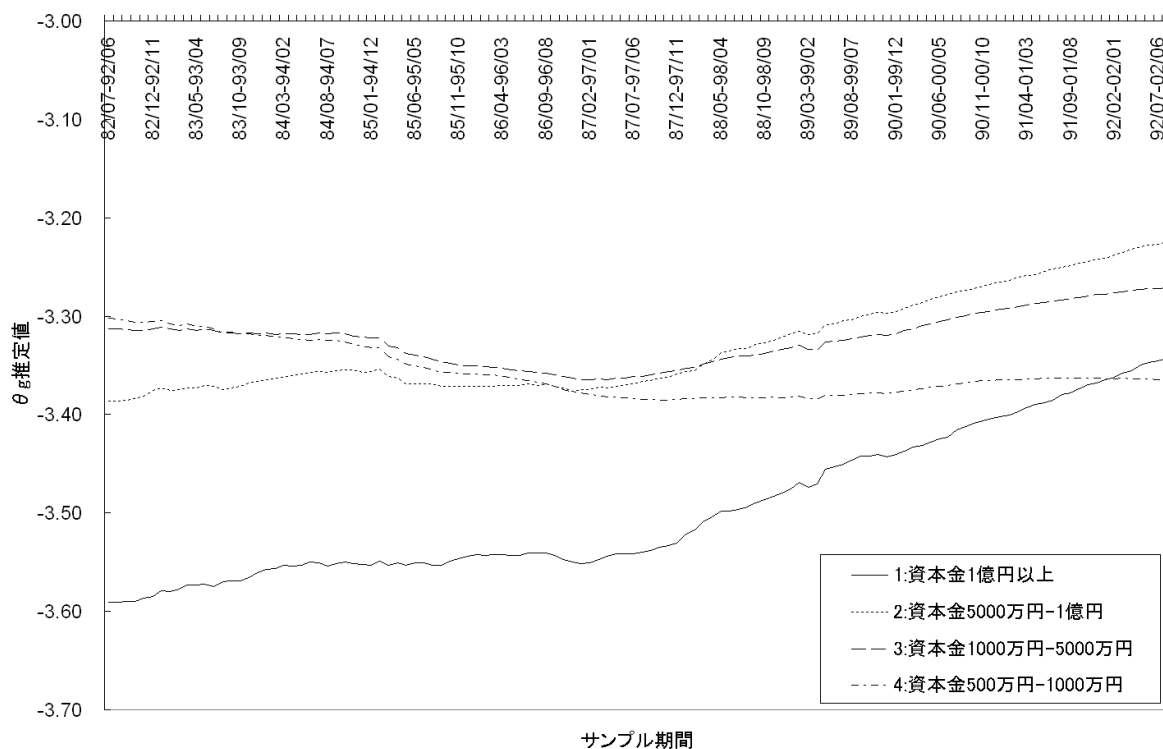


図 10:  $\theta_g$  推定値 (モデル3)

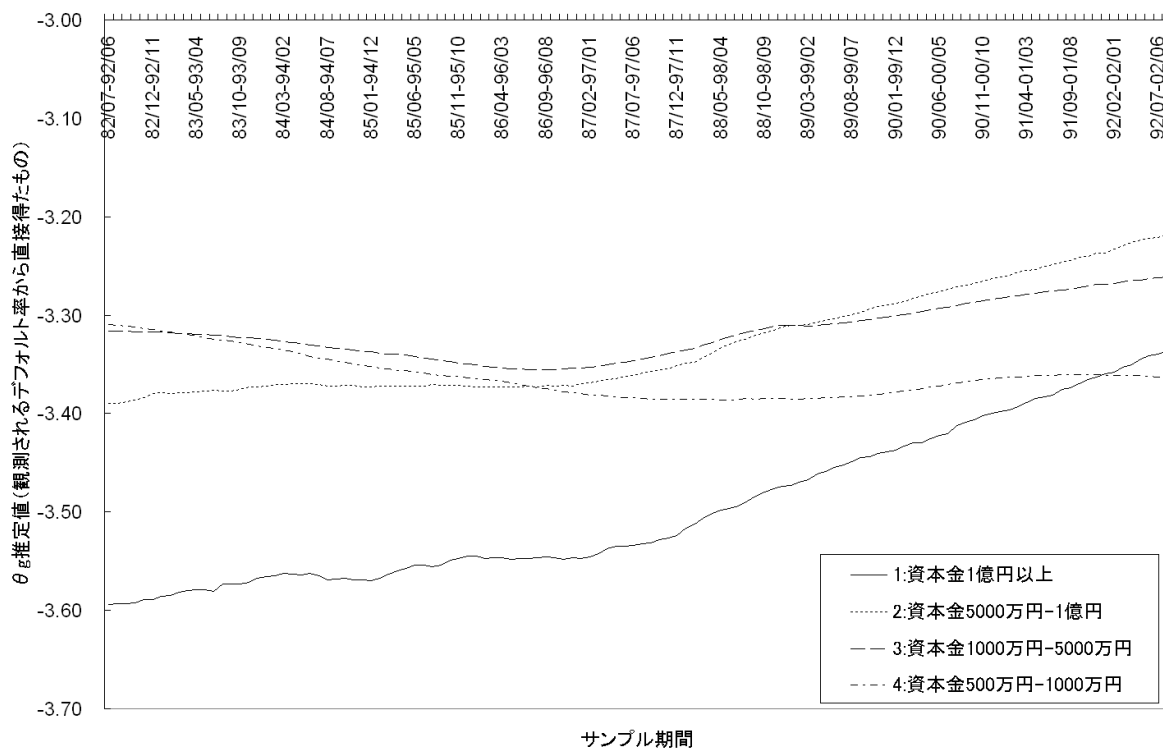


図 11:  $\theta_g$  推定値 (観測されるデフォルト率から得られた閾値:  $\bar{\theta}_g = \Phi^{-1}(\sum D_{g,t} / \sum N_{g,t-1})$ )

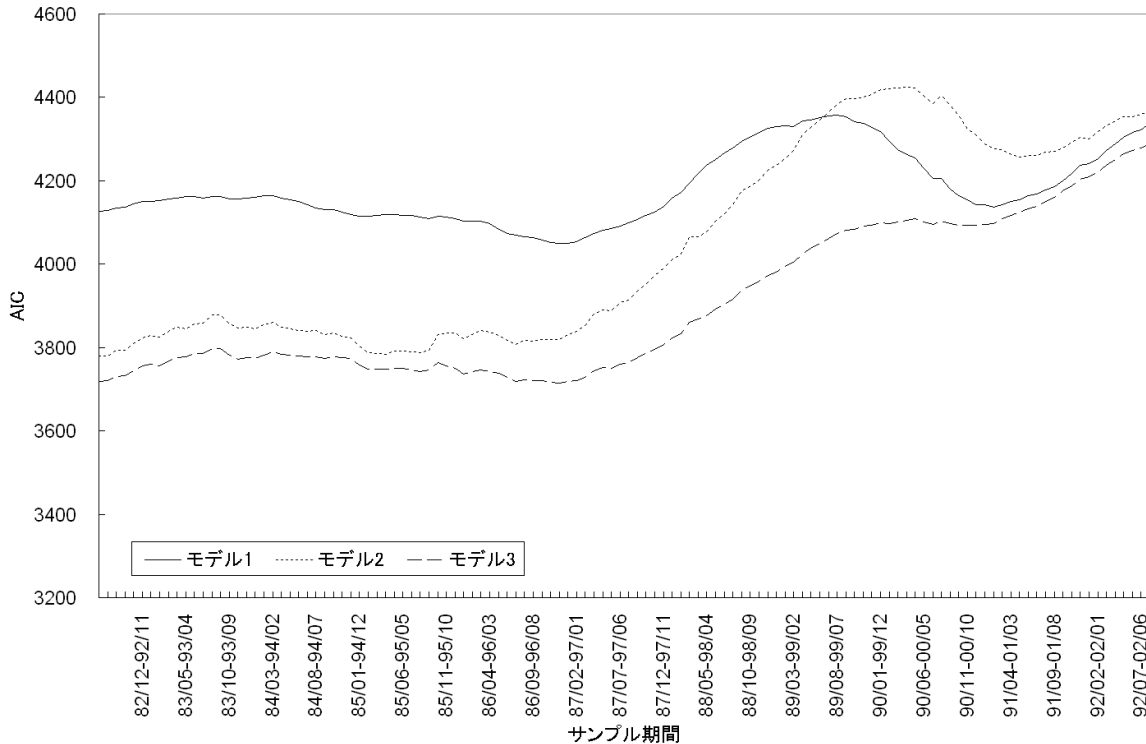


図 12:  $AIC = -2 \times (\text{最大対数尤度} - \text{自由パラメータ数})$ 。モデル 3 が AIC を最小とする最適なモデルである

ていると考えられる。前節の分析通り、このような場合モデル 2 は大きな推定誤差をもたらすため、単純なモデル 2 のモデル 3 の代用としての利用には問題がある。

カテゴリ間相関の推定値がカテゴリ内相関に影響しないために（表 1）、カテゴリ内共通ファクター  $X_{g,t}$  間に相関が存在している場合においても、 $\rho_0$  が比較的小さく債務者・デフォルト数が十分にある場合には、カテゴリ内相関  $\rho_g^2$  のみの推定に限定すればモデル 1（カテゴリ内 1 ファクターモデル）はモデル 3 の比較的良好な代用となると考えられる。ただし、債務者数が少ない、デフォルト率が低いなどのデフォルトに関する情報の少ないカテゴリ内相関の推定には誤差が大きくなるリスクが増えることに留意すべきである。

モデル 1 とモデル 3 では推定できる相関係数行列が違っている。モデル 1 では  $\rho_0$ 、すなわちカテゴリ間相関が推定できず、モデル 3 の場合（式 10）と比較するとカテゴリ間をまたぐ債務者間のアセット相関係数はゼロとみなされる。本節の実データによる推定で  $\rho_0$  はゼロでないということは明らかになっている。結論としては、カテゴリ間相関の推定のためにはモデル 3 の適用が不可欠という結論ということになる。

## 7. 結論

本論では、潜在変数の 2 ファクターモデルの形式を基礎として、デフォルトに関するアセット相関パラメータの最尤推定の方法を提示した。このモデルは Gordy & Heitfield [11] の提案した 1 ファクターモデルで各カテゴリ内の共通ファクター間に相関がある形式のモデルに、それらの相関係数が一定であるという仮定を付け加えることによって、簡素化した形と位置づけられる。

まず、所与のパラメータから人工的にデフォルト発生数を作り出し、これをサンプルとして用いて推定計算を多数回行うことで、各モデルの推定誤差の水準を確認し、またバイアスの存在を調べた。その際、本モデル（モデル3、2ファクターモデル）と、相関がゼロでカテゴリ内共通ファクター間相関もゼロとなる場合（モデル1、カテゴリ内1ファクターモデル）、相関が1で完全にカテゴリ内共通ファクターが同一の場合（モデル2、グローバル1ファクターモデル）の3つのモデルの比較を行った。実験の結果、モデル2では、モデルの仮定に反して共通ファクター間の相関が低いときにファクター負荷推定値の下方バイアスが起こる。モデル1では共通ファクター間の相関が高いときにもそのようなバイアスは起こりにくい。債務者数が少ない場合にはそのような下方バイアスが起こることがわかった。

次に、日本の企業の資本金区分別倒産数のカテゴリ別時系列データ約20年分のデータを用いパラメータ推定を行った。このようなデフォルト実績を用いた日本企業のアセット相関の推定には既存研究が存在せず、本論が初めての試みであると考えられる。推定の結果、推定値はゼロでなく、カテゴリ内のアセット相関は時点により変動していることが明らかになった。加えてカテゴリ内共通ファクター間の相関も100%近くから25%程度まで大きく変動していることがわかった。このためモデル1やモデル2の仮定は実際に満たされておらず、適用には少なからず問題があることが示された。

結論としては、現実のデータを踏まえた本論の分析によれば、カテゴリ内共通ファクター間の相関の可能性を織り込んだモデル3が最適である。これは、カテゴリ間相関の推定を通して、他のカテゴリの情報も同時に用いて推定を行うモデル3が、特にデフォルト数が少ないなど情報が不足している場合に、誤差や推定値のバイアスの小さいより頑強な推定結果をもたらすためである。モデル2はカテゴリ内共通ファクター間の相関が高い場合にしか適切な推定ができないので、適用は問題が多い。モデル1は最も単純であり、一つのカテゴリの時系列デフォルト数、債務者数データを揃えればその債務者カテゴリ内の相関は推定可能であるため、実用的には扱いやすい。しかしこのような限定された時系列データによる限定された債務者集団データ単独による相関の推定は、他のカテゴリの情報を広範に利用するモデル3に比較して、カテゴリ内相関の推定値が過少評価されるなどのバイアスが存在する可能性がある。

今後の課題としては、シミュレーション実験をより多くの種類の前提条件（特に対象期間、債務者数、閾値）下で行うことがある。特にデフォルト実績に関する情報の不足する、債務者数が少ない場合、デフォルト率が低い場合等の推定誤差の増大について、モデルごとに傾向があるかをより詳しくみることは、実務への適用の側面からも意義が高いと考えられる。またそのような検討が進み、かつ推定計算の能力が向上すれば、相関パラメータ数の増加や正規分布ファクターモデルで前提となっている正規コピュラ以外の依存構造を用いた推定など、実証を伴った推定モデルの改良も課題とされるようになるであろう。

## 謝辞

本論をまとめるにあたり、東京工業大学大学院イノベーションマネジメント研究科の中川秀敏助教授から貴重なコメントをいただいた。この場を借りて謝意を表したい。本研究は、部分的に独立行政法人日本学術振興会科学研究費補助金 (A) No.16201033 の助成を受けて行っている。なお本論の内容は個人的な研究成果であり、著者の現在並びに過去の勤務先のいかなる意見や手法を示すものでもない。

## 参考文献

- [1] L. Boegelein, A. Hamerle, R. Rauhmeier and H. Scheule: Modelling default rate dynamics in the CreditRisk+ framework. *Risk*, October (2002), 24–28.
- [2] P.S. Calem and J.R. Follain: The asset-correlation parameter in Basel II: for mortgages on single-family residences. *Working Paper* (Board of Governors of the Federal Reserve System, 2003).
- [3] 中小企業庁編: 中小企業白書 2003 年版 (中小企業庁, 2003).
- [4] P. Demey, J. Jouramin, C. Roget and T. Roncalli: Maximum likelihood estimation of default correlations. *Risk*, November (2004), 104–108.
- [5] S. Duchemin, M-P. Laurent and M. Schmit: Asset return correlation and Basel II: The case of automotive lease portfolios. *Working Paper* (CEB 03/007, 2003).
- [6] K. Dullmann, and H. Scheule: Determinants of the asset correlations of German corporations and implications for regulatory capital. *Presentation paper* (10th annual meeting of German Finance Association, October 2003).
- [7] M. Dietsch and J. Perey: Should SME exposures be treated as retail or corporate exposures? A comparative analysis of default probabilities and asset correlations in French and German SMEs. *Journal of Banking and Finance*, **28** (2004), 773–788.
- [8] C.C. Finger: Conditional approaches for CreditMetrics portfolio distributions. *CreditMetrics Monitor* (RiskMetrics Group, April 1999), 14–33.
- [9] R. Frey and A. McNeil: Dependent defaults in models of portfolio credit risk. *Working Paper* (University of Leipzig Department of Mathematics and ETHZ Department of Mathematics, 2003).
- [10] M. Gordy: A Comparative anatomy of credit risk models. *Journal of Banking and Finance*, **24** (2000), 119–149.
- [11] M. Gordy and E. Heitfield: Estimating default correlations from short panels of credit rating performance data. *Working Paper* (Federal Reserve Board, 2002).
- [12] G.M. Gupton, C.C. Finger, and M. Bhatia: *CreditMetrics Technical Document* (RiskMetrics Group, 1997).
- [13] A. Hamerle, T. Liebig, and D. Rosch: Benchmarking asset correlations. *Risk*, November (2003), 77–81.
- [14] 北野利幸: デフォルト相関に関する  $t$  分布ファクターモデル –CDO 評価への応用–. *ジャーナル・ジャーナル 金融工学と証券市場の計量分析 2006* (東洋経済新報社, 2006), 83–117.
- [15] D. Lucas: Default correlation and credit analysis. *Journal of Fixed Income*, **5** (1995), 76–87.
- [16] A. Lucas, P. Klaassen, P. Spreij and S. Straetmans: An analytic approach to credit risk of large corporate bond and loan portfolios. *Journal of Banking and Finance*, **25** (2001), 1635–1664.
- [17] The MathWorks: *Optimization Toolbox User's Guide* (The MathWorks Inc, 2004).
- [18] R. Merton: On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance*, **29** (1974), 449–470.

- [19] M.J. Miranda and P.L. Fackler: *Applied Computational Economics and Finance* (MIT Press, 2002).
- [20] K. Nagpal and R. Bahar: Measuring default correlation. *Risk* March (2001), 129–132.
- [21] 坂元慶行, 石黒真木夫, 北川源四郎: 情報科学講座 A・5・4 情報量統計学 (共立出版, 1983).
- [22] M. Schmit: Credit risk in the leasing industry. *Journal of Banking and Finance*, **28** (2004), 811–833.
- [23] P. Schönbucher: Factor models for portfolio credit risk. *Working paper* (Department of Statistics, Bonn University, 2000).
- [24] A. Servigny and O. Renault: Correlation evidence. *Risk*, July (2003), 90–94.
- [25] 東京商工リサーチ編: 倒産月報 (東京商工リサーチ, 2004).
- [26] O.A. Vasicek: Probability of loss on loan portfolio. *Working paper* (KMV Corporation, 1987).
- [27] C. Zhou: An analysis of default correlations and multiple defaults. *Review of Financial Studies*, **14** (2001), 555–576.

北野利幸

東京工業大学大学院社会理工学研究科

〒152-8552 東京都目黒区大岡山 2-12-1

## A. デフォルト相関とアセット相関の定義

## デフォルト相関係数

$$\rho_{ij}^b = \frac{p_{ij} - p_i p_j}{\sqrt{p_i(1-p_i)}\sqrt{p_j(1-p_j)}}$$

$p_{ij}$  は債務者  $i, j$  の同時デフォルト確率,  $p_i, p_j$  は債務者  $i, j$  の (周辺) デフォルト確率 .  
アセット相関係数

$$\rho_{ij}^a = \frac{Cov(V_i, V_j)}{\sqrt{Var(V_i)}\sqrt{Var(V_j)}}$$

$V_i, V_j$  は債務者  $i, j$  の潜在変数 .

## B. モーメント法によるアセット相関の推定法

以下は Gordy [10] による推定方法を本論に添った形で整理したものである<sup>26</sup>. 式 (5) の条件付きデフォルト確率に対応する  $p(X)$  ( $X$  は 1 ファクターモデルにおける共通ファクターである, 標準正規分布に従う確率変数) の分布の 1 次並びに 2 次モーメントが観測されたデフォルト率分布のそれらと合うようにパラメータ推定される .

サンプルとしてデフォルト率の観測値と母集団の数の複数期分の観測値の集合をそれぞれ  $\hat{p}, \hat{n}$  とする . まず, 1 次モーメントである無条件デフォルト率については, デフォルト率観測値の平均  $\bar{p}$  であるとして推定される . すなわち  $E[p(X)] = \bar{p} = E[\hat{p}]$  である . 2 次モーメントである条件付きデフォルト率のボラティリティ  $Var[p(X)]$  については, 2 変量標準正規分布関数を  $\Phi_2$ , 単変量標準正規分布関数を  $\Phi$  として  $\bar{p}$  を用いて次が成り立つ .

$$Var[p(X)] = \Phi_2(\Phi^{-1}(\bar{p}), \Phi^{-1}(\bar{p}), \rho^2) - \bar{p}^2 \quad (15)$$

$\rho^2$  は 2 変量正規分布の相関係数である . 一方で  $Var[p(X)]$  は観測値の分散  $Var[\hat{p}]$  と母集団の数の観測値の集合  $\hat{n}$  を用いると次のように得られるので, (15) 式と (16) 式から数値計算によってアセット相関係数  $\rho^2$  を推定する .

$$Var[p(X)] = \frac{Var[\hat{p}] - E[1/\hat{n}]\bar{p}(1-\bar{p})}{1 - E[1/\hat{n}]} \quad (16)$$

## C. 対数尤度最大化計算の実装

実際に対数尤度最大化の計算においては, 尤度関数中の 2 項分布は必要に応じて正規分布で近似<sup>27</sup>されている . また積分計算にはガウス求積法を用いている<sup>28</sup>. 対数尤度関数を  $-1$  倍

<sup>26</sup>証明は Gordy [10] 付録 B, 付録 C 参照 . ここでの記述は主に Dullmann & Scheule [6] を参考にした .

<sup>27</sup> $N_{g,t-1} \cdot p_g(x_{g,t})$  と  $N_{g,t-1} \cdot (1 - p_g(x_{g,t}))$  が充分大きい場合 (20 以上) には, 次の 2 項分布の正規分布による近似式を用いた (スターリングの公式) .

$$f(D_{g,t}; N_{g,t-1}, p_g(x_{g,t})) \simeq \frac{1}{\sqrt{N_{g,t-1} p_g(x_{g,t}) (1 - p_g(x_{g,t}))}} \cdot \phi \left( \frac{D_{g,t} - N_{g,t-1} p_g(x_{g,t})}{\sqrt{N_{g,t-1} p_g(x_{g,t}) (1 - p_g(x_{g,t}))}} \right)$$

<sup>28</sup>ガウス求積法の MATLAB プログラムは Miranda and Fackler [19] にある `qwnnorm` を用いた . ノード数 (積分区間の分割数) は計算速度との兼ね合いから本論の実験を通して  $2^5 = 32$  分割 としている .



して最小化問題とし，各ファクター負荷には0以上1以下の制約<sup>29</sup>があるため，MATLABの制約付き最小化の関数 `fmincon` を用いて推定計算を実装した<sup>30</sup>．

#### D. デフォルト数シミュレーション

以下は(11)式に沿った忠実なデフォルト数のシミュレーションとして理解しやすいようにしたひとつの説明である．実際の計算手順上はバリエーションがあり得るものの，3つの確率変数  $Y_t$ ， $Z_{g,t}$ ， $\varepsilon_{i,t}$  の実現値に同じ乱数列を使っている限り作成されるサンプルデータは全く変わらない．

1. 標準正規乱数の実現値をひとつ得て， $y_t$  とする．
2. 標準正規乱数の実現値をカテゴリ  $G$  個得て，それぞれ  $z_{g,t}$  とする．
3. 標準正規乱数の実現値をカテゴリ  $g$  に属する債務者数分 ( $N_{g,t-1}$ ) 得て，確率変数である  $\varepsilon_{i,t}$  の実現値  $\varepsilon_{i,t}$  とする．
4.  $\varepsilon_{i,t} \leq \frac{\theta_g - \rho_g (\rho_0 y_t + \sqrt{1 - \rho_0^2} z_{g,t})}{\sqrt{1 - \rho_g^2}}$  であれば債務者  $i$  はデフォルトとして，カテゴリ  $g$  のデフォルト債務者数を数える．これが  $D_{g,t}$  となる．
5. 全てのカテゴリ  $g$  について3と4を繰り返す．
6. 全期間  $T$  期分のデータになるまで  $t$  を進めながら1~5を繰り返す．

<sup>29</sup>実際には計算の高速化のため  $10^{-5} \leq \rho \leq 0.5$  の制約をおいた．

<sup>30</sup>`fmincon` では複数種類の最適化アルゴリズムを使い分けることができるが，本問題では逐次2次計画法 (SQP) を用いている．これはラグランジュ関数のヘシアンを各反復ごとに近似しながら準ニュートン法的に解いていく方法である．詳細についてはMATLAB ユーザーズガイド [17] 参照．

## ABSTRACT

**ESTIMATION OF DEFAULT DEPENDENCY USING HISTORICAL  
DEFAULT DATA – MAXIMUM LIKELIHOOD ESTIMATION OF ASSET  
CORRELATION WITH 2-FACTOR MODEL –**

Toshiyuki Kitano  
*Tokyo Institute of Technology*

In the thesis, I discuss the estimation of default dependency using historical default data. 2-factor model is used to represent category-separated variability of latent variables as proxy for obligor credit. For simplification, one numerical calculation for maximum likelihood estimation of factor loading and in turn asset correlation is employed. By Monte Carlo analysis of the estimator's robustness and by real estimation with historical default data, it is confirmed that the model requires less computing power than existing methods, and generates more robust estimator than other category-blinded models that deal with real data. This shows that it is important for the robust estimation of default dependency to use multiple-category default data and estimate inter-category correlation, even when only correlation within one category is required for estimation.