ベータ分布デフォルト率モデル

ベータ分布デフォルト率モデル
-A Stochastic Default Rate Model

王　京穂

要旨　複数の債権から構成されるポートフォリオの信用リスクを計測するには、デフォルト相関とデフォルト連鎖を表現する必要がある。本研究では、デフォルト率が確率変数で、ベータ分布に従うモデルを利用して、デフォルト相関とデフォルト連鎖を整合的に表現することを試みる。

キーワード：信用リスク、デフォルト相関、デフォルト連鎖、デフォルト率の変動、ベータ分布、ベータ二項分布

1、研究背景

企業のデフォルト過程を表現するモデルとしては、Structural Approach（構造モデル）と Reduced-form approach（外生変数モデル）がある。構造モデルは、デフォルトの原因を企業資産価値の変動に求め、その資産価値の変動プロセスからデフォルトの構造を究明するものである。その代表的なものは Merton 型モデルである。外生変数モデルにおいては、企業外部から、企業内部の変化を完全に察知できないものとし、デフォルトを偶然におきる事故として扱う。外生変数モデルにおける外生変数は、瞬間のハザードレート、あるいは年間デフォルト率などのものをさす。

構造モデルは、原理的に優れるもので、デフォルトの発生、株価、社債価格と資産価格を整合的なプラットフォームの上に捉えることができる。一方、実際のデフォルト事例を分析すると、少なくとも企業の外部からみて、デフォルトイベントは意外性を伴って発生することも多い。すなわち、デフォルトイベントには、構造モデルだけでは表現しきれないものが存在するのである。この視点から、少なくとも今の時点においては、デフォルトを事故として捉える外生変数モデルの方が、デフォルトイベントをより適切に表現できるものである。

外生変数モデルでは、デフォルト相関を表現するには、多くの選択肢がある。例えば、デフォルト率は外生的に与えられるものとして、デフォルト相関の関係を Merton 型モデルに求めるハイブリッドモデルがある（CreditMetrics）。また、デフォルト率の変動を何らかのリスクファクターに駆動され、確率的に変動するモデルがある。この場合、デフォルト率の変動およびリスクファクターとの関係に関する設定は重要になる。しかし、これに関する研究はあまり多くなかった。豊沢［1997］は分布を特定せず、リスクファクターに駆動されるデフォルト率の分散を中心に展開した。CreditPlus はデフォルト率がガンマ分布に従うものとした。扱いやすさから、デフォルト率とリスクファクターとの関係をプロビット変換に求めるモデルは多い。

デフォルト時刻を無視すれば、対象期間中におけるデフォルト数は基本的に二項分布に従う。ただ、デフォルト率は変動するので、この二項分布はいわゆる一般二項分布（Generalized binomial distribution）になる。Ahahira, Kashima, Takahashi［1997］によると、二項分布の母数をベータ分布に従うモデル（Beta-Binomial model）を用いて一般二項分布を高い精度で表現できる。本研究は、この Beta-Binomial model をデフォルト過程への応用を試みる。全体的な構成は以下の通りである。

2、ハザードモデルとデフォルト率

債務者総数を n とし、それぞれを番号 (i) で特定される。債務者 (i) の状態を次の状態関数で表現する。

$$
(1)\qquad
\begin{aligned}
&z_i(0)=0\\
&z_i(t)=\begin{cases}0 & non-default\\ 1 & default\end{cases}
\end{aligned}
$$

期間 $(t, t+\Delta t)$ に、債務者 (i) がデフォルトする確率を $\lambda_i(t)\cdot\Delta t$ で表す。ここの $\lambda_i(t)$ はハザードレートで、事象の発生可能性を表すパラメータである。$\lambda_i(t)$ と $z_i(t)$ との関係は、次の通りである。

$$\Pr(z_i(t+\Delta t)=1|z_i(t)=0)=\lambda_i(t)\cdot\Delta t=1-e^{-\lambda_i(t)\Delta t}$$

(2) $$\Pr(z_i(t+\Delta t)=0|z_i(t)=0)=1-\lambda_i(t)\Delta t=e^{-\lambda_i(t)\Delta t}$$

$$\Pr(h_i(t+\Delta t)=1|h_i(t)=1)=1$$

上記定義より、将来 t 時点において、企業の状態分布は次の通りになる。

(3) $$\Pr(non-default)\equiv\Pr(z_i(t)=0)$$
$$\Pr(default)\equiv\Pr(z_i(t)=1)$$

(4) $$\Pr(z_i(t)=0)=\lim_{\Delta t\to 0} e\prod_{k=0,t/\Delta t} e^{-\lambda_i(k\Delta t)\Delta t}=e^{-\int_0^t \lambda_i(s)ds}$$
$$\Pr(z_i(t)=1)=1-e^{-\int_0^t \lambda_i(s)ds}$$

ここで、デフォルト時刻を考慮せず、最終状態 $z_i(t)$ のみを観察していれば、これは1つのベルヌーイ試行である。このベルヌーイ試行におけるデフォルト事象の発生確率のパラメータは次のようなものになる。

(5) $$p_i=1-e^{-\int_0^t \lambda_i(s)ds}$$

これから、この p_i をデフォルト率と呼ぶ。$z_i(t)$ の平均・分散は次のように計算される。

(6) $$E[z_i(t)]=p_i$$
$$V[z_i(t)]=p_i\cdot(1-p_i)$$

n社のプールにおける t 時点までの（累積）デフォルト数を $Z(t)$ とする。$Z(t)$ は次のように構成される。

(7) $$Z(t)=\sum_{i=1,n} z_i(t)$$

このプールにおける t までの累積デフォルト率は次のように計算される。

(8) $$R(t)=\frac{Z(t)}{n}$$

それぞれのデフォルトイベント間に関連がなければ、

(9) $$E[Z(t)]=\sum_i E[z_i(t)]$$
$$V[Z(t)]=\sum V[z_i(t)]$$

(10) $$E[R(t)]=\frac{1}{n}\sum_{i=1,n} E[z_i(t)]$$
$$V[R(t)]=\frac{1}{n^2}\sum_{i=1,n} V[z_i(t)]$$

となる。とくに、プール全体が同じ p を持つ場合（同じ格付け）、$Z(t)$ は二項分布 B(n,p) に従う。このとき、

(11) $$E[R(t)]=p$$
$$V[R(t)]=\frac{1}{n}\cdot p\cdot(1-p)$$

となる。

3、デフォルト相関とその表現

式(11)の結果と実績デフォルト率の時系列データを比べると、明らかに大きなずれが見られる。図1は、実績デフォルト率の時系列と式(11)に従う二項分布の1つのサンプルの比較である（二項分布をシミュレートするときの n は時系列計測の対象企業数を考慮して設定)。図からわかるように、実際のデフォルト率の分散がはるかに大きい。すなわち、デフォルト事象間に何らかのプラスの相関が存在する。この相関の構造は、次のように考えられる。企業活動は経済景気の波に大きく依存するので、企業はもちろん、そのデフォルトの可能性（デフォルト率）も景気の波から共通影響を受ける。その結果、同じ格付の企業でもそのデフォルト率が景気に連動して変動する。すなわち、景気の波に影響され、デフォルトが多く発生したり、あまりしな

表1．実績デフォルト率の「分散不足」現象

	全プール	投資適格	投機的債券
平均	0.014358	0.000985	0.043408
分散	0.009836	0.001018	0.027212
#値	0.003072	0.000701	0.005882

－　データソース：S&P　1981-2006)
－＃は二項分布と見なした場合の分散推計値（サンプリング数調整)

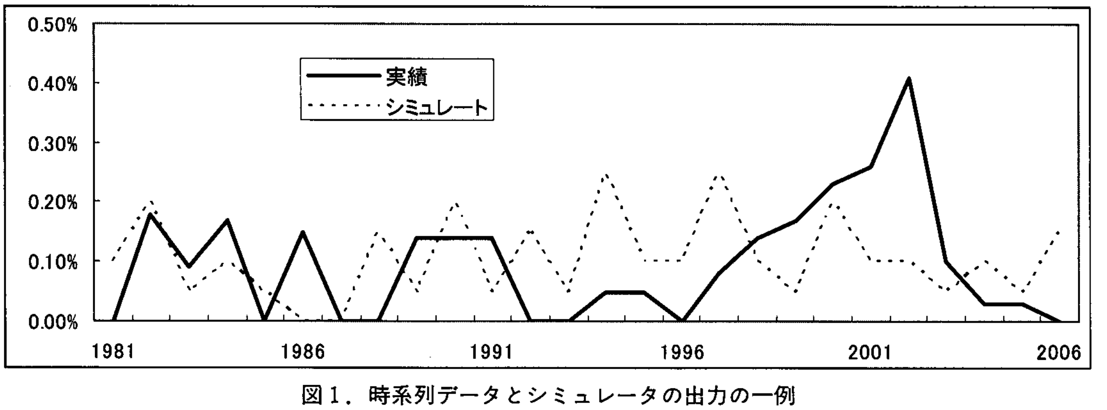


図1．時系列データとシミュレータの出力の一例
（シミュレータは対象企業数を考慮)

かったり、というような相関関係が形成される。

このデフォルト相関は、確率デフォルト率モデルを利用して表現することができる。確率デフォルト率モデルとは、デフォルト率pが、

$$(11)\quad \begin{aligned} &E[\mathrm{p}]=\bar{p}\\ &V[\mathrm{p}]=V[p] \end{aligned}$$

の分布に従うものとし、pの元においては、デフォルトイベント間に関連が無いとするものである。確率デフォルト率モデルの元に、デフォルト数は、

$$(12)\quad \begin{aligned} &E[\mathrm{Z}]=\textstyle\sum_i E[z_i]=\sum_i E_p[E[z_i|\mathrm{p}]]=\sum_i \bar{p}=n\cdot\bar{p}\\ &E[\mathrm{Z}^2]=E[(\textstyle\sum_i z_i)^2]=E[\sum_i z_i\cdot z_i+\sum_{i\neq j} z_i\cdot z_j]\\ &\qquad =n\cdot\bar{p}+n\cdot(n-1)\cdot(\bar{p}^2+V[p])\\ &V[\mathrm{Z}]=n\cdot\bar{p}\cdot(1-\bar{p})+n\cdot(n-1)\cdot V[p] \end{aligned}$$

と計算される。この場合のプールの実績デフォルト率は、

$$(13)\quad \begin{aligned} &E[\mathrm{R}]=\bar{p}\\ &V[\mathrm{R}]=\frac{\bar{p}\cdot(1-\bar{p})}{n}+\frac{n-1}{n}\cdot V[p] \end{aligned}$$

となる。

ここの分散$V[R]$の構成を見ると、pが定数の場合と比べ、$V[p]$の項が存在し、それがnが増加しても減少しない構造になっている。この項はデフォルト相関の結果である。また、nが大きい場合、この関係を利用して、$V[p]$を推定することができる。

なお、図1ケースにおいて、$V[p]=0.00074^2$と推定される。この数字は後に利用される。

4、pが従う分布

pが従う分布として、プロビット分布を採用することが多い。プロビット分布では、正規分布に従うリスクファクター（$x_1, x_2, \cdots\cdots$）を利用して、pは、

$$(14)\quad p=\frac{1}{1+\exp(a_0+a_1\cdot x_1+a_2\cdot x_2+\cdots\cdots+c\cdot\varepsilon)}$$

と書ける。ここの$\{a_i, c\}$はp（業種・格付別）に対応する係数で、これらの係数によって、共通性も表現される（εは共通的ではない変動を表す）。

CreditPlusはpがガンマ分布に従うものとしたが、pがベータ分布に従うと主張する意見はかなり多い。類似的な問題として、毎年の男児出生数を二項分布で表現すると、上記の「分散不足」と同様な問題が発生する。Fisher[1970]が男児の出生率をベータ分布に従うモデルを展開して、分散不足問題を解決した。Ahahira, Kashima, Takahashi[1997]はベータ二項分布の幅広い表現力を証明した。ここでは、pがベータ分布に従うものとする。

区間$(0,1)$上で定義された連続型確率変数pの確率密度関数が，aおよびbを0より大きな定数として、

$$(15)\quad f(p)=\frac{1}{BT(a,b)}p^{a-1}(1-p)^{b-1}\quad 0\leq x\leq 1$$

$$BT(a,b)=\int_0^1 t^{a-1}(1-t)^{b-1}dt$$

で与えられるとき，pの分布をパラメータ（a,b）ベータ分布（beta distribution）という。なお、BT（a,b）はベータ関数（beta function）と呼ばれ、次の性質がある。

$$(16)\quad BT(a,b)=\frac{\Gamma(a)\cdot\Gamma(b)}{\Gamma(a+b)}=BT(b,a)$$

ガンマ関数$\Gamma(\ )$は、次のようなものである。

$$\Gamma(a)=\int_0^\infty t^{a-1}e^{-t}dt$$

pの平均と分散は次のようになる。

$$(17)\quad \begin{aligned} &E[p]=\frac{a}{a+b}\\ &V[p]=\frac{a\cdot b}{(a+b)^2\cdot(a+b+1)} \end{aligned}$$

5、ベータ分布デフォルト率モデル

二項分布$B(n,p)$の二項確率pがベータ分布に従うとき、この二項分布はベータ二項分布になる。Fisherのモデルにおける男児の出身数も、ここでのデフォルト数もベータ二項分布に従う。

ベータ二項分布に従う確率変数をXとし、

$$(19)\quad \begin{aligned} \Pr[X=x]&=\int_0^1 {}_nC_x p^x(1-p)^{n-x}dp\\ &={}_nC_x\frac{1}{BT(a,b)}\int_0^1 p^{x+a-1}(1-p)^{n-x+b-1}dp\\ &={}_nC_x\frac{BT(a+x,b+n-x)}{BT(a,b)} \end{aligned}$$

となる。ここでは、

$$(20)\quad \frac{\Gamma(a+x)}{\Gamma(a)}=(a+x-1)(a+x-)\cdots\cdots(a+1)a$$

である。次のように$(a)_x$を定義する。

$$(a)_x\equiv(a+x-1)(a+x-)\cdots\cdots(a+1)a$$

整理すると、

$$(21)\quad \Pr[X=x]={}_nC_x\frac{(a)_x(b)_{n-x}}{(a+b)_n}$$

を得る。この分布をパラメータ(n,a,b)のベータ二項分布（beta-binomial distribution）という。確率の和が1になることから、もちろん、

$$(22)\quad \textstyle\sum_{x=0}^{n} {}_nC_x\cdot(a)_x\cdot(b)_{n-x}=(a+b)_n$$

である。

Xに関して、

$$E[X]=n\cdot\frac{a}{a+b}=n\cdot\bar{p}$$

(23) $$V[X]=n\cdot\frac{ab(a+b+n)}{(a+b)^2(a+b+1)}$$

$$=n\cdot\bar{p}\cdot(1-\bar{p})+n\cdot(n-1)\cdot V[p]$$

となる。

このモデルをベータ分布デフォルト率モデルと呼ぶ。

6、ベータ分布採用の正当性

ここで、実績デフォルト数の時系列を用いて、デフォルト率の分布にベータ分布を採用する有意性を確認する。よく利用されるプロビット分布との比較も同時に行う。

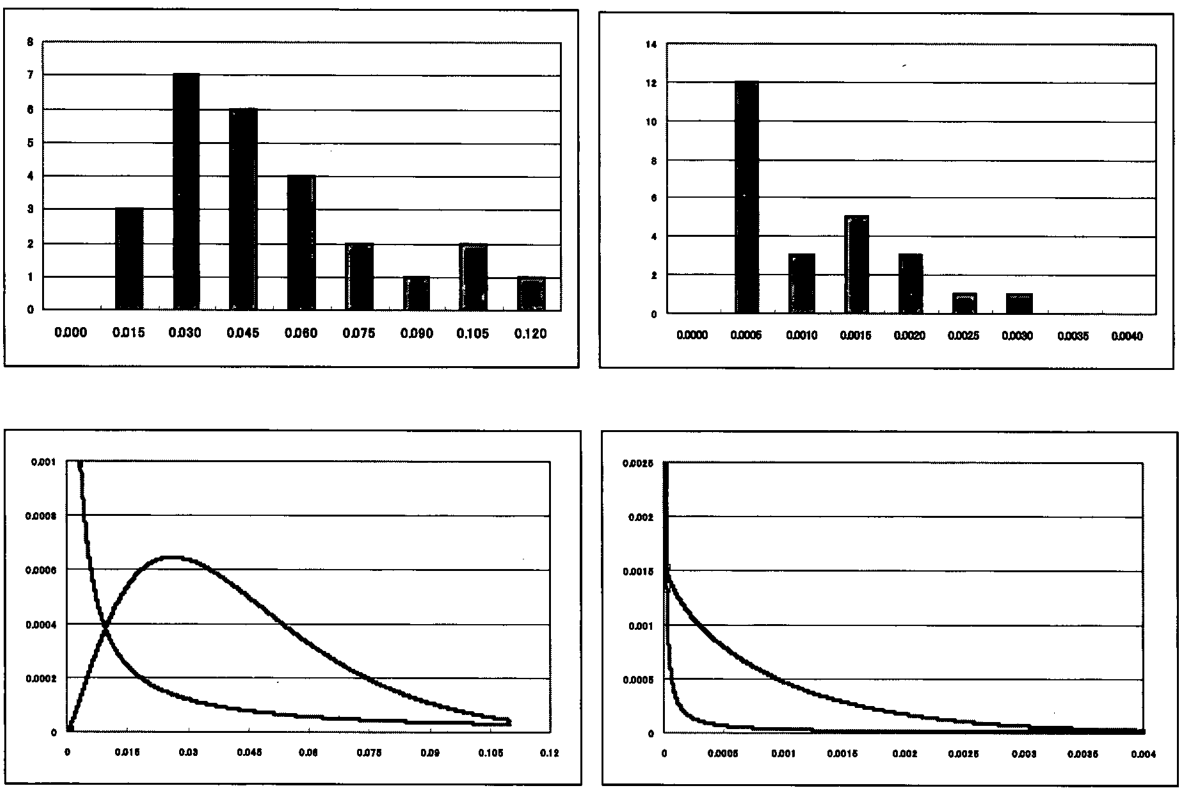
実績デフォルトの度数分布とそれに基づいて推定されたベータ分布とプロビット分布の確率密度関数との対比は図2の通りである。図からわかるように、プロビット分布の確率密度関数のピークは殆どの場合、0の近くに存在するが、ベータ分布のそれが、実績値の最頻値の近くに位置する。高格付け（デフォルト率が非常に小さい）場合、両者間の違いがあまりないが、格付けが下がるにつれて、ベータ分布の方がデータを正確に表現していることがわかる。

3の $V[p]=0.00074^2$ を利用したベータ二項分布のサンプルは図3の通りである。これによって、分散不足問題が解決されることが分かる。

格付のUp-Grade、Down-Gradeのイベントの発生率も基本的にデフォルト率と類似的性質を持つ。すなわち、Up-Grade率・Down-Grade率も分布に従い、何らかの相関構造を持つのである。このUp-Grade率、Down-Grade率もベータ分布を使って表現すると、当てはめが大変よいことがわかる（図4）。

7、業種相関の表現

ここの業種相関とは、業種別のデフォルト率の相関関係をさす。プロビットモデルにおける業種相関の表現は、共通リスクファクターに対する係数によって構成される。変



左上： 実績デフォルト率（Speculative Grade 債券）

右上： 実績デフォルト率（Investment Grade 債券）

左下： ベータ分布とプロビット分布の確率密度関数（Speculative Grade 債券）

右下： ベータ分布とプロビット分布の確率密度関数（Investment Grade 債券）

図2．実績デフォルト率とベータ分布・プロビット分布
（データソース：S&P 社1981-2006）

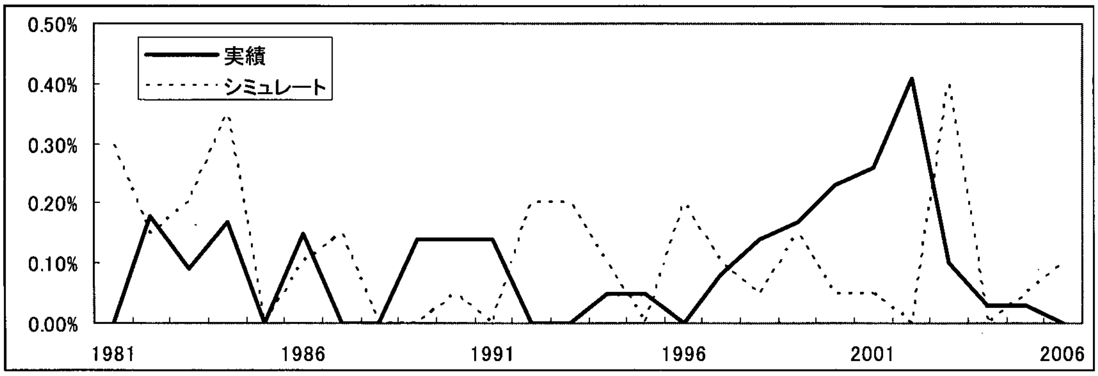
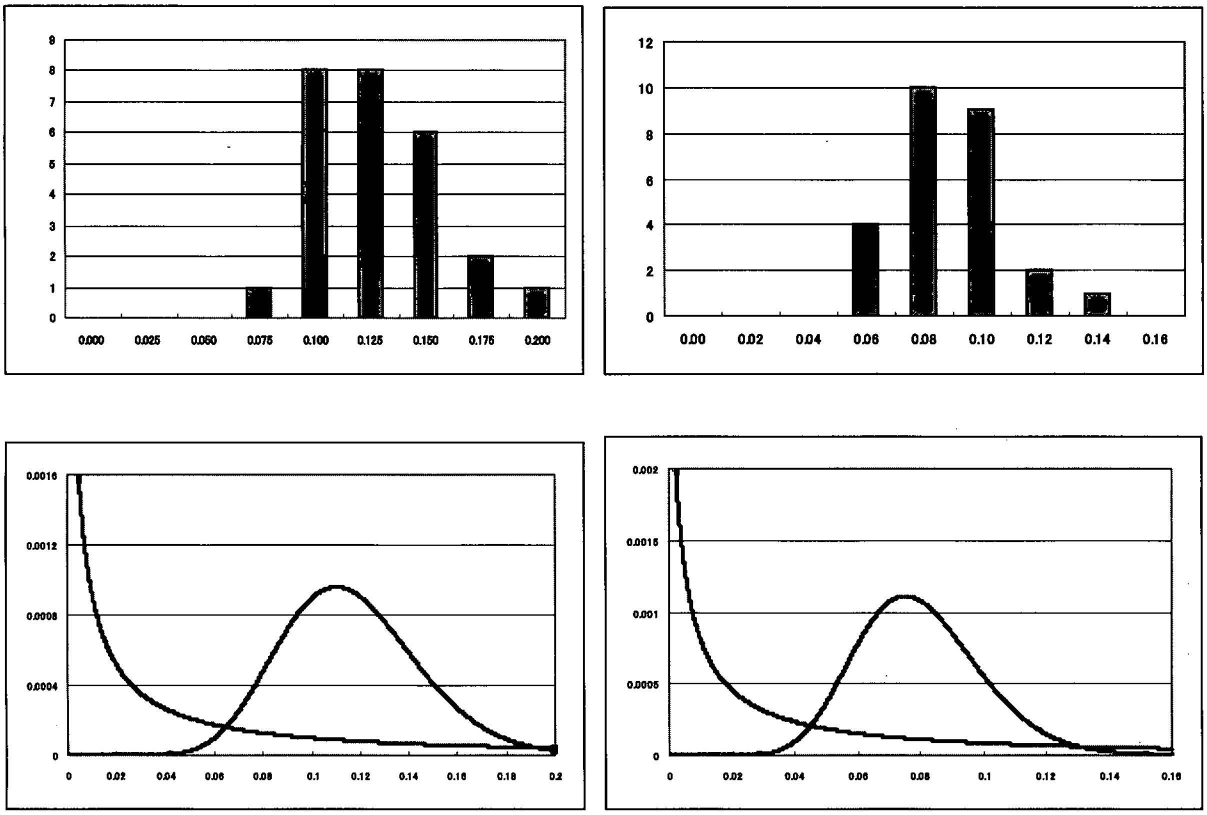


図３．時系列データとベータ分布を導入したシミュレータの出力の一例



左上：　実績 Down-Grade 率　　　　　右上：　実績 Up-Grade 率

左下：　ベータ分布とプロビット分布の Down-Grade 率の確率密度関数

右下：　ベータ分布とプロビット分布の Up-Grade 率の確率密度関数

図４．Down-Grade 率・Up-Grade 率とベータ分布
（データソース：S&P 社1981-2006）

換された変数が正規分布と仮定すれば、諸分析やシミュレーションにおいて利便性が高い。デフォルト率がベータ分布に従う場合、相関は、線形相関ではなく、順位相関が適切である。すなわち、業種ごとのデフォルト率はそれぞれのベータ分布に従うが、デフォルト率間の関係を順位相関で定義する。

業種によって、経済環境への感応度が異なるので、複雑な相関構造が予想されるが、順位相関は大小関係のみを捉え、モデルの正当性と安定性に寄与すると思われる。特にデフォルト率はオーダーのかなり異なる数字の集まりで、高格付けの場合、下方は大きく圧縮されている。このようなデータに対して、線形関係を前提にしたモデルより、順位相関のモデルの方が適切である。

統計的に、順位相関は線形相関よりロバストで、異常値などに影響されないといわれる。この特徴から、順位相関の導入により、実績デフォルトデータの少ない高格付デー

タに対する分析に何らかの前進をもたらすことも期待できると思う。なお、S&P社のデータに対する分析からは、安定性の点から、順位相関と線形相関の間にはっきりした傾向が見られなかった。

順位相関採用のデメリットは、数式による展開が制約され、シミュレーションにより依存することである。しかし、Copulaなどの手法を利用すれば、数値計算やシミュレーションを実現させることは容易である。

表2．相関$\rho(z_1, z_2)$と$V[p]$

ρ	p＝1％	p＝5％
0.01	0.000994987	0.000216852
0.10	0.009949874	0.002168525
0.20	0.019899749	0.00433705
0.30	0.029849623	0.006505575
0.40	0.039799497	0.008674099
0.50	0.049749372	0.010842624
0.60	0.059699246	0.013011149
0.70	0.069649121	0.015179674
0.80	0.079598995	0.017348199
0.90	0.089548869	0.019516724
0.99	0.098503756	0.021468396

8、グループ企業のデフォルト連鎖

企業が外部の経済環境を介して共通影響をうけるという構造は、確率デフォルト率モデルによって表現される。一方、親子会社、グループ会社のように、外部経済環境を介さなくても、直接に緊密な関連を持つ企業（群）がある。このような企業（群）におけるデフォルトイベント間の関連関係をここでデフォルト連鎖という。ここでは、このデフォルト連鎖問題に対して、ベータ分布デフォルト率モデルの応用を試みる。

まず、同じ格付けの2社(1,2)を考える。ここでは、pの変動が外部の経済環境によるものではなく、2社間の関係によって引き起こされるものであるとする。この場合、

(24)　$V[z_1+z_2]=V[z_1]+V[z_2]+2\rho\sqrt{V[z_1]\cdot V[z_2]}$

と計算される。式の中のρは状態関数（z_1, z_2）間の相関である。ρは(0,1)間の値をとり、$\rho=0$は互いに独立、$\rho=1$は2つの状態関数の完全一致を表現する。このρで2社間の連鎖度合いが表現できる。

一方、式(12)より、ベータ分布デフォルト率モデルの元に、

(25)　$V[z_1+z_2]=2\cdot V[z]+2\cdot V[p]$

である。整理して、

(26)　$V[p]=\rho\cdot\bar{p}(1-\bar{p})$

となる。式(26)から、$V[p]$を適切に設定することによって、このモデルはあらゆる強さのデフォルト連鎖を表現できることがわかる。すなわち、$V[p]$によってデフォルト連鎖の強さが表現される。

当然、デフォルト連鎖は、普通のデフォルト相関より、もっと強い関連性を持つので、ここで要求される$V[p]$は、式(12)での$V[p]$よりはるかに大きい。

図5、図6は、ρに対応する$V[p]$から計算されるベータ分布の確率密度関数である。図からわかるように、ρが大きい（$V[p]$が大きい）場合、確率密度関数は0と1の近辺で非常に大きな値をとる（縦軸は対数）。すなわち、高い相関の実現のメカニズムは、同時に$p=0$か$p=1$をとることにある。ベータ分布デフォルト率モデルは、$V[p]$を大きくするだけで、このメカニズムを実現してくれる性質を持っているのである。

格付けが異なる2社間の問題は少し複雑であり、以下のように展開できる。

1）それぞれのデフォルト率はベータ分布に従う
2）デフォルト率分布の平均は格付ごとに決まる
3）デフォルト率分布の分散は、2社間の連鎖の強さから計算する
4）2つのデフォルト率は（順位）相関1で変動する

3）の分散の計算は、格付けごとに、式(26)の$V[p]$の最

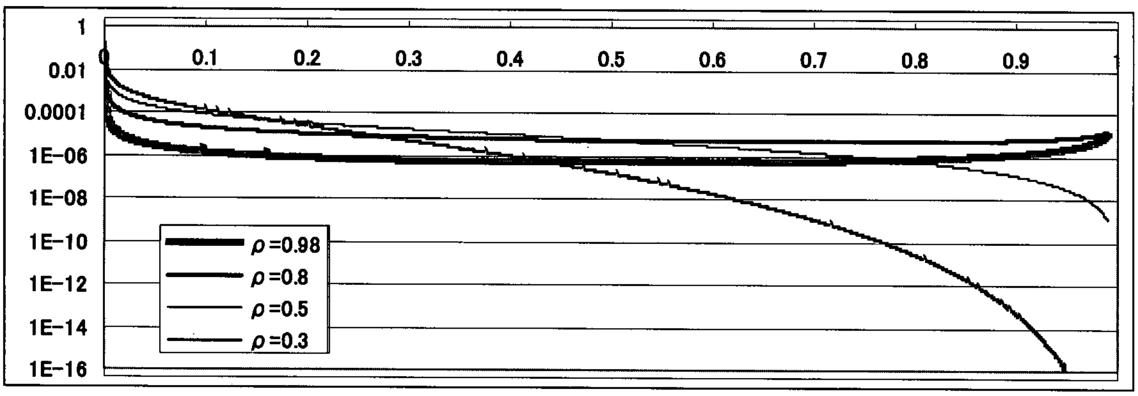


図5．相関$\rho(z_1, z_2)$とベータ分布の確率密度関数（p＝1％）

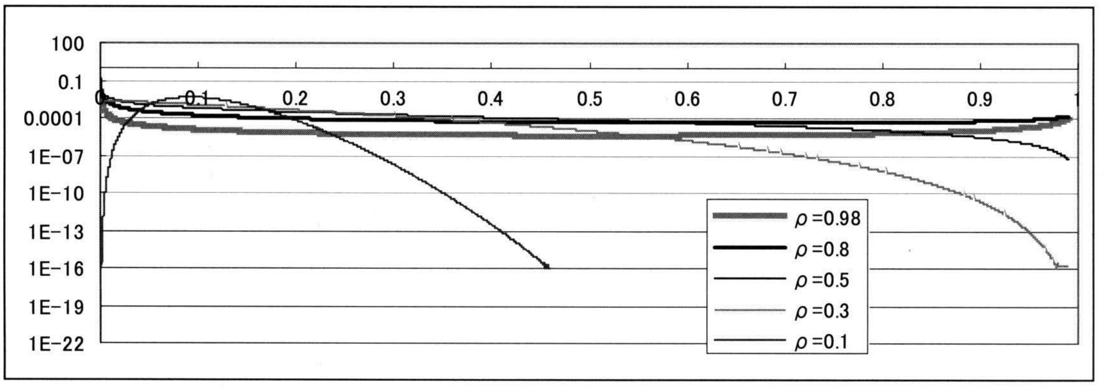


図6．相関 $\rho(z_1, z_2)$ とベータ分布の確率密度関数（p＝10%）

大値 $\overline{p}(1-\overline{p})$ を利用して決めることができる。すなわち、$V[p]$ を連鎖の強さのパラメータとして利用し、格付けごとに（$0, \overline{p}(1-\overline{p})$）間の値を取ればよい。

デフォルト率の変動も用いてデフォルト連鎖を表現するのは難しいと言われている。その原因の1つはおそらくデフォルト率の分布の操作の困難さにあると思われる。ベータ分布デフォルト率モデルにおいては、この操作が容易である。$V[p]$ を設定するだけで、広いレンジのデフォルト連鎖が表現できる。

実際問題として、業種相関とデフォルト連鎖を同時に表現しなければならない。業種相関とデフォルト連鎖の間に関連性がないとすれば、この表現は2つのベータ分布によって簡単に実現される。デフォルト相関（業種相関）の表現は、相関関係を持つ複数のベータ分布によってできる。デフォルト連鎖の表現は、相関を持たない企業グループ数分のベータ分布によって実現する。いずれにしても、モンテカルロシミュレーションのような方法で計算することになる。

9、まとめ

デフォルト相関とデフォルト連鎖は、デフォルトモデルの2つの重要なパーツである。本研究は、この相関と連鎖を表現するモデルとして、ベータ分布デフォルト率モデルを提唱した。ベータ分布デフォルト率モデルは、デフォルト率の変動をより正確に表現し、デフォルト相関とデフォルト連鎖を整合的に捉えることができる。また、ベータ分布の導入による順位相関の採用によって、モデルのロバスト性の向上も期待できる。

なお、相関を持つ複数のベータ分布の生成方法は明示的ではないが、Copula 的な手法を利用すれば、容易にできる。

参考文献

[1] David Shimko, “Credit Risk—Models and Management”, Risk Books, 2004

[2] 刈屋武昭「信用リスク分析の基礎」　東洋経済新報社　1000

[3] 豊沢泰寿“複数の資産で構成されるポートフォリオのクレジットリスクの定量化のフレームワーク”、証券アナリストジャーナル第10号、1997

[4] M. Akahira, H. Kashima, K. Takahashi, “A generalized binomial distribution determined by a two-state Markov chain and a distribution by the Bayesian approach”, Statistical Papers 38, 1997

[5] STANDARD&POORS, “Annual 2006 Global Corporate Default Study And Rating Transitions”, STANDARD&POORS, 2007

[6] Philip J. Schonbucher、望月衛訳「クレジット・デリバティブ」　東洋経済新報社　2003

[7] Fisher, R.A、遠藤健児・鍋谷清治訳　「研究者のための統計的方法」　森北出版　1970