1999年度日本オペレーションズ・リサーチ学会
秋季研究発表会

金融に変革をもたらす大規模信用リスクシミュレーション

ニューメリカルテクノロジーズ株式会社　鳥居　秀行
http://www.numtech.com/　E-Mail: htorii@numtech.co.jp

近年、先進的な一部の金融機関において、数万次元規模のモンテカルロシミュレーションを使ったリスク管理・経営リソース配分が行われるようになってきた。この分野では、非線形特性を持ち互いに相関のある膨大な数の変動要因をモデル化する。通常、この種のテクノロジーは経済学的見地から解説されるが、その工学的側面も十分に興味深い。しかし、熾烈な競争下にあり企業機密を意識する産業研究の常として、その詳細は公表されていない。そこで本稿では、世界的にも数少ない信用リスク管理シミュレーションシステム製品（CreditBrowser）を持ち、上位金融機関をユーザーとする弊社の経験をもとに、その工学技術について概説する。

1．信用リスク計量化問題の背景

大規模な金融機関が破綻すれば、資金決済を通じて他の金融機関に影響が連鎖的に拡大、最終的には一国経済や国際金融システムそのものが危機に陥りかねない。この現象は金融専門家の間でシステミックリスクと通称されている。その防止を目的として、金融機関が抱える資産負債ポートフォリオのリスク総量を数値化し、これに応じて一定のペナルティを与えるための政府規制が1980年代末から主要各国で実施された（BIS自己資本比率規制）。

ところが、中南米危機、北欧危機、欧州通貨危機、アジア危機、ロシア危機、米国ヘッジファンドの破綻など、1990年代になってもこの種の危機やその一歩手前の事態が何度も発生した。そこで、政府規制は強化・精緻化の方向にあり、民間も精力的に研究に取り組んでいる。今日では国際的な金融機関の多くが数値シミュレーションを利用した確率論的リスク管理（VaR: Value at Risk）を日常的に行っている。

現在、このリスク管理の分野で主要テーマになっているのがポートフォリオベースの信用リスク計量化（信用 VaR）である。特にわが国においては1990年代を通じ不良債権問題は大きな懸念材料であった。ようやく終息に向かいつつあるとはいえ、依然として信用リスクこそ金融機関が抱える最大のリスクであることに変わりはない。

2．シミュレーションを要する理由

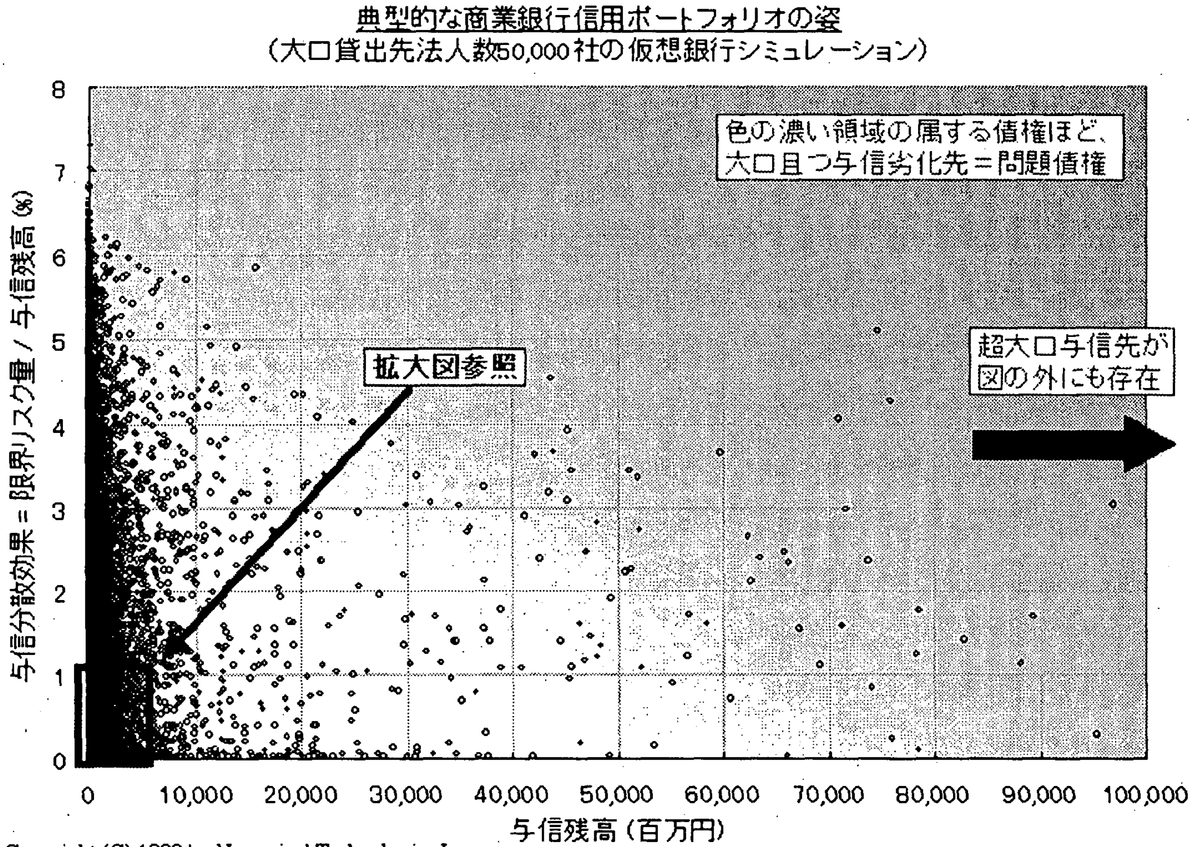
さて、経済学的諸問題を別にしても、信用リスク計量化には次のような工学的壁があって非常に難しい。

- 確率変動を勘案すべきファクターが非常に多い。大手金融機関を例にとれば、主要貸出先に絞っても数万社の企業を個別の確率変数として扱う必要がある。
- 各社に紐付けされた数十万取引件別が存在、個別の取引件別はさらに数百万個のキャッシュフローに分解される。すなわち大規模なデータ処理が必要とされる。
- 確率変動によって起こる事象は非線形の特性を持ち、裾野の長い、あるいは一定しない確率密度関数となっている。このため、正規分布を仮定した簡易な統計処理を用いることができない。

この問題に対して正面から取り組んだ場合の計算量の膨大さは容易に想像できよう。そこで、現在のように技術高度化が進む以前は、簡易に解析式を組み立ててリスク量を求める以外に方法がなかった。これは、保険数学(アクチュアリー、個別の保険契約者の生存関数を記述し、早期の契約者死亡を保険会社サイドの損失であると定義、ポアソン分布、ワイブル分布を用いた大域関数を使い、解析的にリスク量を求める)の応用であり、契約者の死亡と企業の倒産を同一視してモデル化する。しかし、残念ながらこうした簡易手法では、現実に直面する以下の問題を解決できず、結果的にリスク量を著しく過小評価することが今日では判明している。

- 各統計サンプルは対等ではなく、少数ながら大きな与信量を持つサンプルが全体に対して大きな影響を及ぼす。ゆえに大数の法則を仮定できない。
- 個々の確率変数は単純な相関関係にあるわけではない。企業は個別の業績変動以外にも、連鎖倒産の形で他企業からの影響を被りうる。特定の産業や地域金融機関のモデル化においては大きな障害となる。
- 確率変動は時系列変化を伴う。多期間のモデル化を行わねばならないが、経路途中でデフォルト(倒産)した場合は吸収状態としてその先の未来を扱わないように処理する必要がある。

不均等分布をイメージしやすくするために、現実の銀行を模した仮想ポートフォリオを用い、信用リスクシミュレーションモデルで計算した結果を示す(下図)。



Copyright (C) 1999 by Numerical Technologies Inc.

この図では、X 軸が各統計サンプルの大きさ(与信額)、Y 軸が当該サンプルのポートフォリオ全体へのリスク寄与度、各ドットが一つの貸出先企業を示している。仮に図の各サンプルがほぼ同じ大きさ、すなわち X 軸方向に固まって分布していれば保険数学的手法を適用できる可能性がある。しかし、現実には特定の貸出先に大幅な与信集中があり、しかもそれが当該金融機関の問題債権となっているケースが多い。これでは、解析的に結果を予測することは困難である。

3. 計算量の克服

こうして、ポートフォリオベースの信用リスク計量化においては、数値シミュレーションがその主要手段として台頭しつつある。しかし、このナンバークランチング的な手法が実現するまでには1990年代に至る工学および数学の成果を待たねばならなかった。

何よりも最初に直面したのは計算量の壁である。金融においては、次表のように多変量の確率変数を扱うニーズは多い。

金融モデルの登場年代と次元数

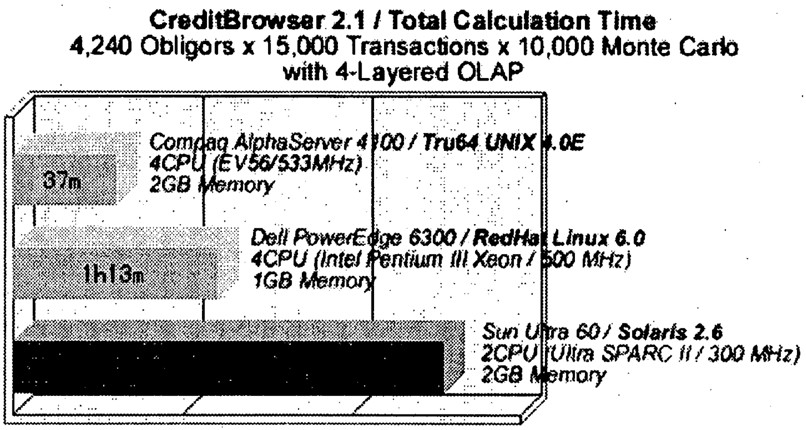
利用目的	変動要素数	登場年代	主な解法
単変量オプションモデル	1	～1980	解析式またはツリー(二分木)
多変量オプションモデル	2～3	1980年代	ツリーまたはモンテカルロ
イールドカーブモデル	3	1980年代末	主成分分析＋解析式
MBS, MBS デリバティブ	～30	1980年代後半	準乱数モンテカルロなど
為替、金利 VaR	300～700	1994～	擬似乱数モンテカルロ
株式 VaR	3,000～	1997～	擬似乱数モンテカルロ
信用 VaR	10,000～	1997～	擬似乱数モンテカルロ

高次元になれば加速度的に計算量が膨らむ。表に示す通り、数百次元超のシミュレーションが実用に供されるようになったのは比較的最近の話である。

その最大の理由はコンピュータの性能向上にある。これは異分野ではあるが、個別の天体運動シミュレーションに基づく惑星形成や銀河の衝突過程研究が同時期に花開いたのとほぼ同じ背景と考えてもよいだろう。現在、かつてのスーパーコンピュータの性能は、ありふれて存在する汎用ワークステーションで利用できる。弊社の CreditBrowser を例にすれば、邦銀大手行の主要法人貸出先数に相当する 15 万件、5 万次元×10,000 回の信用リスクシミュレーションを約 2 時間で計算できる(Compaq AlphaServer 4100 上での計測値)。ハイエンドの計算サーバーを用意すれば、技術的には日本の全法人企業を入力してシミュレーションすることも不可能ではない。もちろん高度な並列処理や数値計算プログラミングといったコンピュータサイエンスの寄与も大きいとはいえ、近年のハードウェアの進歩が最大の寄与をしている。

別の側面として、ハードウェア価格の低下も大きな要因となっている。もともと大手金融機関にとって億円級のハードウェア投資は日常的話題であり、CRAY スーパーコンピュータの主要ユーザーの一角でさえあった。しかし、高額のハードウェアは一般に扱いづらい。

その点、最近の Linux や PC サーバーの台頭は、金融研究者に対して多大なる貢献をしている。実際、ある金融機関では CreditBrowser を Linux 機（4CPU の PC サーバー）上で使用しているが、その計算性能は大型の商用 UNIX サーバー中位機種に相当する。小規模金融機関が対象ならば24時間以内に10万回級のシミュレーションを十分遂行しうる。つまり、大規模並列（MPP）を要する分野でない限り、今やコスト的に商用 UNIX サーバーの十分の一に過ぎない並列 PC サーバーで必要な性能に到達しうるのである。



4．擬似乱数技術

数学からの寄与としては、高次元シミュレーションに必要な高次元均等分布・超長周期乱数が手に入ったことを指摘できる。

モンテカルロシミュレーションの利用者にとって大きな悩みは計算量である。原始的なモンテカルロ法（Crude Monte Carlo）の理論性能（＝計算誤差）は、シミュレーション回数を n とすれば、次の数式に支配されており、シミュレーション回数を増やしても容易に精度は向上しない。

$$O\left(\frac{1}{\sqrt{n}}\right)$$

1980 年代末、まだ計算機の速度が遅かった時代には準乱数法（quasi-random または LDS）、準乱数モンテカルロがよく用いられた。疑似乱数の代わりに規則性を持った数値列＝準乱数を使い、より望ましい確率分布に近づけようとする準乱数モンテカルロを用いれば、シミュレーション回数を n、次元数を d として、

$$O\left(\frac{(\log n)^d}{n}\right)$$

となり、単変量または低次元における、特に期待値や分散を求める用途において大きな効果を発揮する。例えば単変量のオプションプライシングには少ない繰り返し計算で絶大な精度向上の効果があり、1980 年代後半の低速なコンピュータでも十分に実用に供したので、MBS／ABS ブームに乗って主要金融プレーヤーの間で流行した（大半は Faure 列あるいは Sobol 列の改変版）。

ところが、準乱数は高次元下においては先の数式の次元数 d が効き、著しく精度が劣化する。加えて、分布のテールにおける準乱数の挙動はよく解明されていない。VaR 計算では分布の期待値近傍ではなくテールが重要である。順位数計算などには使い難い。そこで 1990 年代に入って金融界では疑似乱数（pseudo-random）が見直されるに至る。

ここで問題になったのが、モデルに要求される次元数が増大するとともに乱数周期を使い果たす危険性である。モンテカルロシミュレーションは乱数を大量に使用する。信用リスクモデルでは1企業に1個の乱数を割り当てる単純なモデルであっても10,000回、50,000社のシミュレーションを行えば 5×10^8 個の乱数が必要となる。これはシステム組み込みのrand関数（線形合同法が多い）の乱数周期に匹敵する。モンテカルロシミュレーションを安全に行うには必要な乱数個数の3乗以上の乱数周期が必要との経験則に従えば、rand関数はもちろんのこと、大半の乱数生成法は不合格となってしまう。加えて、限界リスクのように金融シミュレーションでは条件付き確率分布の計算が必要であり、高次元における均等分布性能も重要となってくる。

そこで、特殊なハードウェアに頼り、再現性のあるシミュレーションが出来ないという大きな欠点（リスク管理モデルを設計する上ではモデル監査のクリアが難しく致命的）のある物理乱数を別にすれば、より良い疑似乱数を探さなければならない。その先鞭となったのが、慶應義塾大学理工学部（現在、九州大学）の松本助教授・西村拓士による疑似乱数Mersenne Twisterである。

Mersenne Twisterは乱数周期 $2^{19937}-1$ という超長周期、高次元均等分布性を持ちながら非常に高速なアルゴリズムであり、信用リスクシミュレーションにも少なからぬ寄与をしている。

5．モーメントマッチング

高次元シミュレーションのもう一つの鍵は、モンテカルロ法の発生誤差を小さくし、少ない試行回数でも収束性を改善する技術である。

疑似乱数をそのまま用いるCrude Monte Carloでは収束性が悪い。しかし、数学・統計学に詳しい者ばかりではない金融実務の世界では、ごく最近まであたり前のようにCrude Monte Carloが使われてきた。特に初期のVaRモデルの中には著名米銀の販売製品の中にすら、「良い乱数がとれるまでひたすら試行を繰り返す」という気の長い方法を採用したものさえある。これでは高次元シミュレーションなど望むべくもない。理解が進むにつれて必然的に、生成する疑似乱数を操作し、より望ましい確率分布に近づけようとする一連のモーメントマッチング手法が金融界でも普及した。特に、対称変量法 と2次サンプリング法 （quadratic resampling）が単独あるいは併用されることが多い。

小さなサンプルデータを使い、各種基本統計量の変化を見ながらその効果を検証してみよう。次の表に示す各種統計量を満たすような3変量の多変量相関乱数シナリオを、モンテカルロシミュレーションを使って100個求めたいとする。

<u>目的の多変量正規乱数系列</u>

基本統計量

	系列1	系列2	系列3
平均（1次モーメント）	0.01000	0.02000	0.03000
標準偏差（2次モーメント）	0.29833	0.24083	0.21213
歪度（3次モーメント）	0.00000	0.00000	0.00000
尖度（4次モーメント）	0.00000	0.00000	0.00000

相関係数行列

	系列1	系列2	系列3
系列1	1.00000	0.54843	0.50011
系列2	0.54843	1.00000	0.68817
系列3	0.50011	0.68817	1.00000

原始的なモンテカルロ法を用いた場合、わずか100回程度の計算試行では、収束性能は次の程度でしかない。

原始的なモンテカルロ法（試行回数100回）

基本統計量

	系列1	系列2	系列3
平均（1次モーメント）	0.05210	0.02710	0.03757
標準偏差（2次モーメント）	0.29159	0.22906	0.19863
歪度（3次モーメント）	-0.06096	0.39699	0.41137
尖度（4次モーメント）	-0.30666	0.17643	0.55905

相関係数行列

	系列1	系列2	系列3
系列1	1.00000	0.47414	0.45856
系列2	0.47414	1.00000	0.71127
系列3	0.45856	0.71127	1.00000

対称変量法を使えば、平均、歪度などの奇数次モーメントをゼロに調整できる。対称変量法に若干の修正を加えれば、次のように平均値のドリフトも完全に調整できる。

対称変量法（試行回数100回）

基本統計量

	系列1	系列2	系列3
平均（1次モーメント）	0.01000	0.02000	0.03000
標準偏差(2次モーメント)	0.28964	0.23389	0.21311
歪度（3次モーメント）	0.00000	0.00000	0.00000
尖度（4次モーメント）	-0.13148	-0.21078	-0.04597

相関係数行列

	系列1	系列2	系列3
系列1	1.00000	0.58045	0.52405
系列2	0.58045	1.00000	0.74021
系列3	0.52405	0.74021	1.00000

2次サンプリング法を使えば、次のように平均、標準偏差、相関係数行列など、2次モーメント以下の統計量を試行回数に関わらず自由に調整することができる。

2次サンプリング法（試行回数100回）

基本統計量

	系列1	系列2	系列3
平均（1次モーメント）	0.01000	0.02000	0.03000
標準偏差（2次モーメント）	0.29833	0.24083	0.21213
歪度（3次モーメント）	0.00470	0.40025	0.57564
尖度（4次モーメント）	0.43710	1.71197	0.72973

相関係数行列

	系列1	系列2	系列3
系列1	1.00000	0.54843	0.50011
系列2	0.54843	1.00000	0.68817
系列3	0.50011	0.68817	1.00000

さらに、対称変量法と2次サンプリング法を併用すれば平均（1次モーメント）、標準偏差と相関係数（2次モーメント）、歪度（skewness、3次モーメント）、およびより高次の奇数モーメントを、シミュレーション回数に関係なく入力データと一致させることができる。

対称変量法と2次サンプリング法の同時使用（試行回数100回）

基本統計量

	系列1	系列2	系列3
平均（1次モーメント）	0.01000	0.02000	0.03000
標準偏差（2次モーメント）	0.29833	0.24083	0.21213
歪度（3次モーメント）	0.00000	0.00000	0.00000
尖度（4次モーメント）	-0.73001	1.19494	0.15166

相関係数行列

	系列1	系列2	系列3
系列1	1.00000	0.54843	0.50011
系列2	0.54843	1.00000	0.68817
系列3	0.50011	0.68817	1.00000

このようにモーメントマッチングを行えば劇的な改善効果がある。しかし、急速な収束性改善が見られたからといって、真の解に必ず漸近することとはもちろん別問題であり、ある程度の計算試行は必要である。時々、「わずか100回のモンテカルロシミュレーションで完璧な答えが出た」といった誤った報告を見かけることがあるが、これは典型的Type I Errorである。とはいえ、こうしてようやく金融分野におけるモンテカルロシミュレーションも実用域に達した。

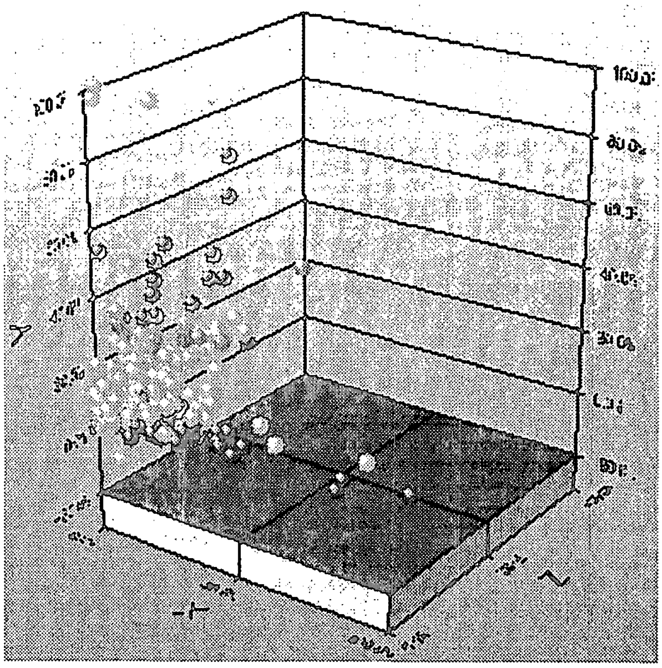
なお、弊社ではMersenne Twisterによる多変量乱数、対称変量法、2次サンプリング手法などをExcelから使用できるアドインソフトウェアNtRandを無償提供している。ここにあげたサンプルもNtRandを使用して作成した。詳しくはhttp://www.numtech.co.jp/documents/19981222/をご参照頂きたい。

6．金融モデルの新展開

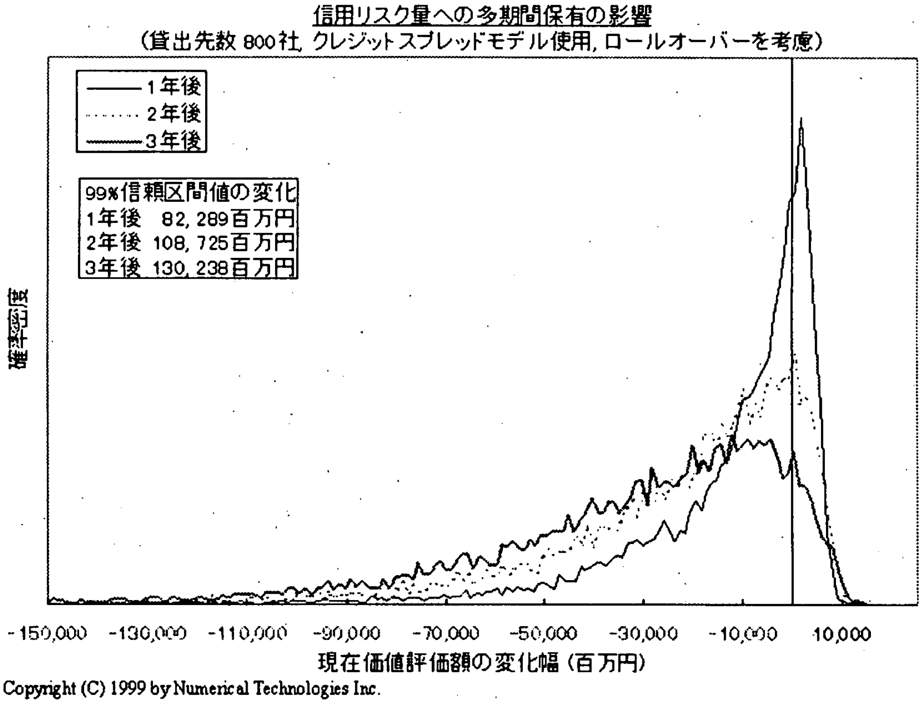
さて、乱数技術に纏わる数学的複雑さ、工学的には計算量さえ克服してしまえば、シミュレーション技術の特性として後は比較的容易に応用ができる。金融の世界でもこれは例外ではなく、多彩な成果が花開きつつある。

信用リスク計量化の分野での典型的事例には、与信残高、ポートフォリオリスク寄与度、収益スプレッドによる2次元分析、あるいは3次元分析がある（下図）。

リスクリターンの3軸分析



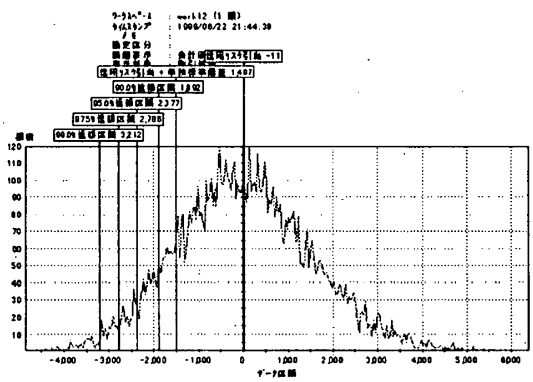
吸収型マルコフ鎖を仮定した多期間シミュレーションによる時系列分析もその一つである。



また、金融モデルはますます多くの事象を取り込み、連続的な市場変動ばかりでなく、企業倒産などイベント事象も織り込んだ長期保有シミュレーションまで可能となっている。

金融取引のプロであるディーラーが使うオプションモデルが仮定する資産価値変動の確率密度分布は次の図のようなものだ。これは、BIS 自己資本比率規制の中で「高度なリスク管理手法」とみなされている市場 VaR モデルでさえ同じである。

一般型株式 VaR モデル（7,000 資産）

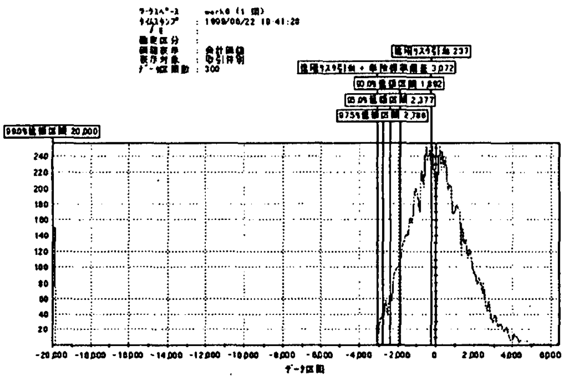


しかし、このような確率密度分布は、仮定する保有期間が長くなればなるほど実態から乖離する。その理由の一つはイベント事象を無視してモデルが構築されているためである。金融モデルの大半は時系列データの連続性を仮定して理論構築されており、企業のデフォルトといったイベント事象でデータが不連続になった時は、モデル側が対応できない。そうした場合、当該時系列データそのものを異常値や欠損値として除くのが通例である。逆に連続性のあるデータに対しては GARCH 手法などを使って必要以上とも思える詳しさで研究されている。

実践経験豊富なディーラーは、金融モデルが「平常時の市場」でしか成立せず、しばしば起こる「市場急変」に対応できないことをよく知っており、適切な対応をすることができる。しかし、人間の判断が入らない VaR モデルはそうはいかない。このイベント要因による確率変動の乱れは「ファットテール問題」と呼ばれ、しばしば問題視される。

この「ファットテール問題」に対する一つの回答となりうるのが、イベント事象のモデル化である。シミュレーション技術を駆使することにより、デフォルトイベントをモデルに取り込むことは既に可能だ。たとえば、先の図のモデルに修正を加えて CreditBrowser から出力したのが次の図である。

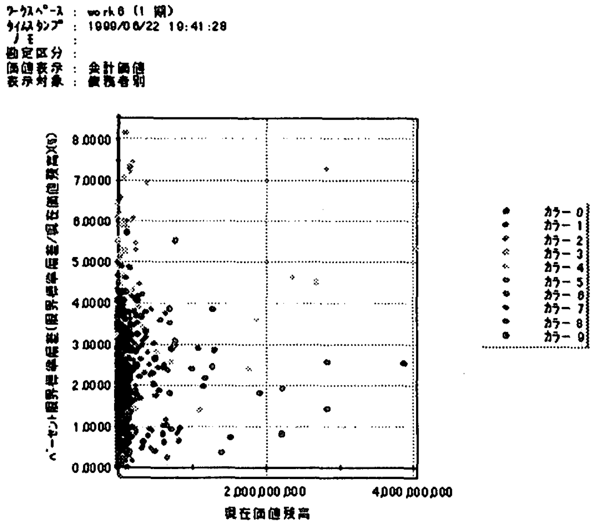
マートン型株式 VaR モデル（7,000 資産）



確率密度分布の左端（損失側）にピーク値があるが、これがデフォルトイベントを示す。値幅制限などの意図的市場操作を捨象して考えれば、市場シミュレーションとしてこちらの方がより実態に近いのは言うまでもない。

別の応用としては、リスク寄与度から見て全体ポートフォリオのリスク分散効果を損ねている資産を抽出する、資産最適化の用途にも供しうる。

株式のリスク寄与度分析



7．おわりに

従来は不可能と思われた規模の計算が可能となった時、科学研究のブレークスルーが起こることは稀ではない。数値実験の成果として流体力学、有限要素法解析、高エネルギー物理学... そしてまた金融も例外ではない。

地域モデルから始まった気象モデルが、やがて全地球モデルとなり、ついには地球を区分するメッシュの細かさを競うようになったように、国家ブロック全体をメッシュで区分するような類の社会科学シミュレーションの実現を弊社では夢見ている。