

(日本銀行仮訳)

マーケット・リスクに対する所要自己資本額算出に用いる  
内部モデル・アプローチにおいてバックテストニングを  
利用するための監督上のフレームワーク

バーゼル銀行監督委員会

バーゼル

1996年1月

## .序文(イントロダクション)

当ペーパーは、バーゼル銀行監督委員会(以下「当委員会」)が、マーケット・リスクに対応する所要自己資本額算出に用いる内部モデル・アプローチに、バックテストを導入する際のフレームワークを示したものである。これは同時に発出された「マーケット・リスクを自己資本合意の対象に含めるための改定」のパートB.4(j)の細目に該当するものである。

内部モデルを利用してマーケット・リスクを測定している金融機関の多くは、モデルによって算出されたリスク量と実際のトレーディングの損益とを日々比較し、自行のリスク管理モデルの質や精度をチェックしている。これは一般に「バックテスト」と呼ばれるが、リスク管理モデルを導入し発展させてきた多くの金融機関にとって有益なものと考えられている。

リスク管理モデルの精度を評価する手段としてバックテスト手法は、今後とも発展していくであろう。バックテストにおける新たな手法は次々と開発されており、こうした問題がリスク管理に関心を持つ関係者の間において議論されている。現在のところ、バックテストを用いた比較手法は銀行によって区々であり、結果についての評価基準も統一されていない。目下、正確なリスク管理モデルと不正確なリスク管理モデルとを明確に識別する手法の確立を目標として、現在利用されている手法を改善し精緻化しようとする意欲的な試みが続けられている。

バックテストの本質は、実際のトレーディング結果を、モデルによって算出されたリスク量と比較することである。仮に、バックテストの結果、双方の計数が十分に似通ったものであれば、当該リスク管理モデルは精度面で問題がないと考えられる。一方、比較分析の結果、双方の計数にかなりの開きが生じ、モデル自体ないしはバックテストの前提条件に重大な問題が存在することが明らかになる場合もあろう。こうした2つの極端な事例の間には、バックテスト結果からだけでは一概に結論付けられないグレーな領域が存在する。

当委員会は、規制体系上バックテストを導入することによって、内部モデル・アプローチに整合的かつ様々な状況にも対応可能な形で適切なインセンティブが織り込まれたと確信する。事実、1995年4月の内部モデル・アプローチ提案に対して寄せられたコメントでは、銀行がリスク管理モデルを継続的に高度化していくとする明確なインセンティブを与え続ける必要性を強調するものが多かった。マーケット・リスクの所要自己資本額算出に用いる内部モデル・アプローチの中にバックテストをより明確な形で取り込む上で、当委員会は次の2点を考慮した。即ち、市中において最適なバックテストの手法について未だコンセンサスが得られていないこと、及び、モデルの精度を測定する手段としてバックテストを用いること自体に限界があることである。

当委員会は、当ペーパーに掲げるフレームワークが、バックテストの限界を認識しつつ、適当なインセンティブを取り込む必要性にも配慮したバランスのとれたものであると確信する。また当委員会は、リスクの計測やバックテストに係る手法が未だ発展過程にあることを認識しており、かかる分野において新たに重要な発展がみられた場合には、上記フレームワークの中に適宜取り込んでいく所存である。

以下、当ペーパーで、内部モデル・アプローチにおけるバックテストのフレームワークを説明する。このフレームワークの目的は、市中金融機関がバックテストの手法をさらに一段と精微化することを促進し、バックテストの結果を銀行監督上どのように判断するかについて明らかにするものである。次のセクションではバックテスト自体のもつ特性について考察し、さらに続くセクションでは、銀行監督の観点から見たバックテスト結果に対する解釈を示した上で、これらの点に関して当委員会で合意をみた基準を提示する。

#### ・バックテストのフレームワークについて

当委員会の定めたバックテストのフレームワークは、マーケット・リスク

を管理するモデルを内部的に利用している銀行の多くが採用しているフレームワークに基づいている。こうしたバックテスト・プログラムは、通常、銀行の日々のバリュー・アット・リスク値と、実際に生じた日々の損益(いわゆる「日々のトレーディング損益」とを定期的に比較するものである。殆どの場合、バリュー・アット・リスク値の方がトレーディング損益に対し大きくなるはずであるが、トレーディング損益がバリュー・アット・リスク値を超過した日数の観測日数全体に対する割合は、バリュー・アット・リスク値を算出する際の信頼水準に依存する。バリュー・アット・リスク値をトレーディング損益と比較することは、単純に言えば、例えば、リスク値がトレーディング損益を上回っている日数を数えることである。つまり、実際に生じたトレーディング損益がリスク値の範囲内に収まった日数の観測日数全体に対する割合と、当初予想された信頼水準とを比較することによって、銀行のリスク管理モデルの精度を検証することができる。この検証に当たって統計的なテスト手法を利用することもできるが、時には上述した簡便な手法が用いられることもある。

当ペーパーに規定する監督上のバックテストのフレームワークは、上記の手順を全て含むものであり、各手順が銀行にとって過度な負担とならないよう実現可能性及び、統合的な解釈に配慮したものである。なお、バリュー・アット・リスクにおけるリスク量は、一定の保有期間中の市場変動の結果、特定の信頼水準の下で、一定のポジションから発生する想定可能な最大損失額を意味する。

(当規制において)適用されるバックテストは、リスク管理モデルによって算出されたリスク量の範囲にトレーディング結果が収まっている日数の全体に占める割合が、信頼水準 99%と整合性がとれているかどうか比べるものである。即ち、信頼水準 99%の下での銀行のリスク管理モデルが実際にトレーディング損失の 99%をカバーしているかどうかをみようとするものである。もっとも、こうした 99%という極めて高い信頼水準を設定することは、信頼水準が低い場合に比べ、ただでさえ難しいモデルの精度の検証をさらに難しくするのではないかと、といった議論を展開することも可能であろう。しかしながら当委員会は、バックテスト

イングの結果と「自己資本合意の改定」に明記された信頼水準との整合性を図ることが重要である、との結論に達した。

バックテストにどのようなリスク量とトレーディング結果を用いることが適切かを決定するに当たっては、バリュー・アット・リスクによるリスク計測手法が、通常、固定されたポートフォリオの瞬間的な市場価格の変動に対する感応度を用いて計算していることに起因する問題も考慮しなければならない。即ち、当日の取引が終了した時点におけるトレーディング・ポジションをリスク管理モデルに入力することにより、この固定されたポートフォリオの価値が、特定の保有期間における価格や金利の変動により最大限どの程度変動しうるかを予想する結果となっている。

この手法は理論的には極めて分かり易いものであるが、実務上はバックテストを複雑にしている。例えば、保有期間中にポートフォリオの構成が変化することにより実際のトレーディング損益が「濁る (contaminated)」ことが避けられないため、厳密にはバリュー・アット・リスク値とトレーディング結果を単純に比較することはできない、としばしば指摘される。これは、ポートフォリオの構成が変化することによって生じるトレーディング損益 (の「振れ」) や、手数料収入などをトレーディング損益に含めてしまうことは、そうした損益がバリュー・アット・リスクを算出する上で前提とされている固定されたポートフォリオに内在するリスクとは全く無関係であることに照らし適切ではないという考え方に基づいている。

こうした議論は、保有期間がより長期に亘る場合の価格変動を捉えようとするバリュー・アット・リスク手法において説得力をもつ。おそらく、保有期間 10 日間、信頼水準 99% の下で内部モデルによって算出されたリスク量を、10 日間のトレーディング損益と見比べることは、余り意味のある作業とはいえないであろう。とりわけ、トレーディング業務を積極的に行っている金融機関では、10 日の間に当初のポジションに比べ、ポートフォリオの構成が大幅に変化していることは、よく起

ることである。こうした理由から、当ペーパーに規定されるバックテストिंगのフレームワークでは、保有期間1日間の下で算出されるリスク量を基準として用いることとする。当ペーパーで特に明記された条件を除き、バックテストिंग手法は銀行が内部管理上利用しているものに基づくことになる。

リスク量の算出に当たり保有期間1日間としていることとの平仄をとる上で、1日間のトレーディング損益をバックテストिंगに利用することが適当である。しかしながら、トレーディング損益を日次ベースでみることも、上述したトレーディング損益の「濁り」の問題は解消されない。言い換えれば、日々のトレーディング損益は日中トレーディングの影響を受けるだけでなく、新規契約の結果計上される手数料収入を含む可能性があるため、これを日々のリスク量と比較することは適切ではない、との指摘である。

日中取引はトレーディング損益のボラティリティを高める傾向があり、日中取引によりトレーディング損益がリスク量を超過する事態が生じる可能性がある。もっとも、こうした事態はリスク量を算出する手法自体に問題があることを意味しない。むしろ、こうした問題はリスク管理手段としてのバリュー・アット・リスク手法が想定している範囲を超えたものである。これに対し、手数料収入をトレーディング損益に含めることは、同収入が年金のように安定したキャッシュ・フローが期待できるといった特性をもつことから、バックテストिंग結果を逆の方向に歪めることになる。通常、リスク量を算出するに当たり手数料収入は考慮されていないことが多く、バックテストिंगを行う上での「トレーディング損益」の定義に手数料収入を含めることによって、リスク管理モデル自体の問題がみえにくくなる可能性がある。

こうした問題に対する考え方としては、一方で、たとえリスク量を算出する前提条件に限界があったとしても、実際のトレーディング損益は、銀行がリスク管理を行う上で最も重要かつ意味のある計数であり、同結果をリスク量と見比べるべき、とする見方がある。これに対し、手数料収入の問題は、多少乱暴な形であるにせよ、

バックテストを実施する前にトレーディング損益から手数料収入の平均値を除くことにより、十分解決可能である、とする見方もある。後者については、より精緻な手法として、日々の収入を手数料、スプレッド、価格変動、日中取引など各種の要因に分解した上で分析することも考えられよう。

バックテストを、バリュー・アット・リスク・モデルの精度をみる純粋に統計的なテスト手法と捉えるならば、「濁りを除去した」テストを実施できるように日々のトレーディング損益を定義することが明らかに適当である。この基準を満たすためには、銀行は、前営業日終了時のポジションを固定させた上で、ポートフォリオの価値がどのように変動したかという方法でバックテストを実施できるような能力を備えなければならない。

一方、(ポートフォリオを固定せず)実際のトレーディング結果を用いたバックテスト手法も有用である。何故なら、こうした手法により、リスク管理モデルが正確に計算を行っていたとしてもトレーディングによって損益が振れるようなケースを発見することが可能だからである。

こうした理由から、当委員会は、銀行が、(ポートフォリオを固定させた)仮想(hypothetical)のトレーディング損益と、実際のトレーディング損益をそれぞれ用いてバックテストを実施できることが望ましいと考えている。こうしたバックテスト手法に関する2つの異なるアプローチのうち、何れを重視するかは各国監督当局により扱いが異なることとなるが、どちらのアプローチにも意味があることは自明である。両アプローチを併用すれば、モデルによって算出されたリスク量とトレーディング損益との関係について、より一層理解を深めることが可能となる。

バックテスト・プログラムの仕様を決定するための次のステップは、バックテスト内容とバックテストの実施頻度である。当委員会が採用するフレームワークは、モデルによって算出されたリスク量をトレーディング損失と比較する最も単純な方法であるが、トレーディング損失がリスク管理モデルが想定し

たリスク量を超過した回数 (exceptions、以下「超過回数」)を単に数えることである。例えば、信頼水準 99%、保有期間 1 日のバリュー・アット・リスク・モデルについては、平均的に、営業日数 200 日のうち 2 日を除く 198 日において、トレーディング損益が事前に算出されたリスク量の範囲内に収まっていなければならない。

バックテストの実施頻度に関しては、バックテストを可能な限り多くの観測データに基づいて実施すべきであるという要請と、定期的を実施すべきであるという要請とのバランスを図る必要がある。当ペーパーが採用するバックテストのフレームワークでは、規制上規定されたバックテストおよび超過回数の分析を、四半期毎に直近 12 か月のデータに基づいて実施すべきものとする。

バックテスト・プログラムの導入時期は、内部モデル・アプローチによる自己資本規制が発効する時点、即ち、遅くとも 1997 年末からとする。これは、1998 年末までに、バックテスト・プログラムによる上記超過回数の分析が開始されていることを意味する。もっとも、これは、各国監督当局が、上記日付以前にバックテストの結果を徴求すること、とくに各国裁量により、内部モデルを承認する過程の一部としてバックテストのデータを利用することを排除するものではない。

直近 12 か月のデータを利用することは、約 250 営業日分の観測データをバックテストの目的で利用することを意味する。各国監督当局は、(250 営業日のうち)銀行のモデルにおいて生じた超過回数を規制上の対応措置をとる際の基準として用いることになる。多くの場合には、規制上の対応措置がとられないことになろう。また、場合によっては、監督当局は内部モデルについて問題がないか調べるため、銀行と意見交換を開始することになろう。さらに、最も深刻な事態として、監督当局が、銀行の所要自己資本の増額を命じたり、自己資本規制上、内部モデルの利用を禁止することもある。

バックテストにおいて、超過回数の発生状況を主たる基準として用いるこ



との利点は、そのアプローチの単純明快さにある。統計的な観点からみると、銀行のモデルの精度を超過回数に基づいてチェックすることによって、比較的少ない前提条件を仮定することで足りることになる。とくに重要な前提条件は、日々のバックテストの結果（トレーディング損益がバリュー・アット・リスク値を超過したかどうか）が、それぞれ独立であることである。

当然のことながら、当委員会は、こうしたテスト手法が正確なモデルと不正確なモデルとを完全に峻別できないことを承知している。統計の専門家にとっては、何らかの形で問題を含むモデルを全て識別する傍ら、他の正確なモデルを誤って排除することがないような形でバックテスト手法を調整することは、不可能であることを意味しよう。こうした限界は、当ペーパーで説明されるフレームワークを構築する上で非常に重要なポイントであるばかりか、各国の監督当局が銀行のバックテスト・プログラムの結果を評価する上でとりわけ関心を抱く点である。しかしながら、当委員会は、こうした限界の存在によって、バックテストの利用を否定すべきものとは考えていない。仮にバックテスト自体に限界があり不完全性が拭えないとしても、監督上の基準を明確なフレームワークの上に構築することは、まったく裁量的な基準やインセンティブのない基準を導入することに比べ、はるかに望ましい。

#### ・バックテスト結果を評価するための監督上のフレームワーク

##### (a) 3ゾーン・アプローチについて

当委員会は、バックテストの統計上の限界を念頭に置きつつ、バックテスト結果を評価するための監督上のフレームワークを導入することとしたが、バックテストから得られる情報の程度に応じ様々な監督上の対応が可能となるよう配慮した。バックテストの結果については、超過回数により3つのゾーンに色分けし、段階的な対応措置が定められている。グリーン・ゾーンは、バックテストの結果、銀行のモデル自体の質や精度に問題がないと考えられる

領域である。イエロー・ゾーンは、バックテストの結果、問題の存在が示唆されるが決定的ではないという場合である。レッド・ゾーンは、バックテストの結果、まず間違いなく銀行のリスク管理モデルに問題があると考えられる領域である。

当委員会は、こうしたゾーンの定義に関する基準として、バックテスト・プログラムにおいて生じた超過回数を用いることについて合意した。その内容は以下に記すとおりである。しかしながら、こうした定義を正しく理解するためには、銀行のリスク管理モデルの精度に関して様々な仮定をおいた上で、一定の超過回数が発生する確率を調べることが有用である。

#### (b)ゾーンを決定する際に統計面から考察すべき事項

3つのゾーンについて簡単に説明したが、ゾーン間の境界線は次の2つの統計上のエラーのバランスを図るべく決定された。すなわち、(1)バックテストの結果、正確なリスク管理モデルが不正確なものと判定されてしまう可能性と、(2)不正確なモデルが、バックテストによって不正確であると判断されない可能性である。

表1は、モデルがどの程度の正確性を有しているかについて幾つかの条件を設定した上で、250個の独立した観測データから特定の超過回数が発生する確率を示したものである(2項確率を提示)。具体的には、表1の左半分は正確なモデルの確率(即ち、真に99%の信頼水準を満たしている場合)を示したものである。この前提条件の下、「exact」と記された列は、例えば、標本全体の6.7%において超過回数が丁度5回となると期待されることを示したものである。

表1の右半分は、幾つかの不正確なモデルの精度(信頼水準)を示したものであり、真の信頼水準が98%、97%、96%、95%のモデルを左から順番に並べたものである。従って、実際の信頼水準97%のモデルにおける「exact」の欄をみると、このようなモデルの場合、標本全体の10.9%において超過回数が5回となること

が期待されることが示されている。

また、表 1 は重要なエラーが発生する確率についても示している。モデルが 99% の信頼水準を実際に満たしている（望ましい捕捉水準）との前提条件の下、モデルの精度に関する最低基準として超過回数の特定値を決めた場合に、正確なモデルを誤って排除することとなるエラー（「タイプ 1」のエラー）が生じる確率を示している。例えば、仮にモデルを排除する基準を超過回数「1」と定めた場合、正確なモデルであっても、実際に超過回数が発生せず排除されないのは 8.1% の場合に過ぎず、正確なモデルは 91.9% の確率で誤って排除されることとなる（正確なモデルでも 91.9% の確率で超過回数 1 以上となる）。モデルを排除する最低基準として超過回数を引き上げるにつれ、こうしたエラーが発生する可能性は減少していく。

表 1 は、モデルの実際の信頼水準が 99% ではないと仮定し、モデルの精度に関する最低基準として特定の超過回数を定めた場合に、一定の確率で、仮定した信頼水準の（不正確な）モデルを誤って容認することとなるエラー（「タイプ 2」のエラー）が発生する確率についても示している。例えば、仮にモデルの実際の信頼水準が 97% であり、モデルを排除する最低基準として超過回数全 7 回と定めた場合、表 1 によればこの（不正確な）モデルは 37.5% の確率で誤って容認されることになる。

表 1 における情報をみる上で重要な点としては、同表に掲げられているモデルの信頼水準にさほど違いがない（97% と 99% とはあまり変わらない）ようにみえる場合であっても、算出されるリスクの大きさはかなり異なることである。即ち、銀行の算出したリスク量が正確なモデルによって算出されるものに比べ相当小さくなっていても、トレーディング損益の 97% がカバーされているという状況が起こりうるのである。例えば、トレーディング損益が正規分布であるとの前提の下、信頼水準 97% をカバーすることは 1.88 標準偏差（ ）に相当するリスク量を指すが、信頼水準 99% の場合は 2.33 標準偏差に相当するリスク量となり、その結果、リスク量としては 25% 程度増加することとなる。従って監督当局としては、信頼水準 99% のモデルと、実際には信頼水準 97% しかカバーされていないモデルとを峻別

することは極めて重要であると考えている。

### (c) 各ゾーンの定義

表1の結果は、バックテストの統計的な限界について示すものでもある。とくに、モデルを排除する超過回数の最低基準を設定する上で、誤って正確なモデルを排除する確率が低く、同時に、不正確なモデルを間違えて容認してしまう確率が低いような、特定の値は存在しない。こうした理由から、当委員会は単一の最低基準値に拠るアプローチを採用しなかった。

こうした限界を踏まえつつ、当委員会はトレーディング損益を3つのカテゴリーに分類した。第1のカテゴリーは、バックテストの結果が正確なモデルと整合的であり、不正確なモデルを間違えて容認してしまう確率が低い領域（グリーン・ゾーン）である。その裏返しとして、バックテスト結果からみる限り正確なモデルによって算出されたとは考えにくく、間違えて正確なモデルを排除してしまう確率が殆どない領域（レッド・ゾーン）がある。この両者の中間には、バックテスト結果に照らしモデルが正確であるとも不正確であるともいえる領域（イエロー・ゾーン）が存在し、同領域については、銀行に対し内部モデルに関する追加的な情報を求めた上で、監督当局が具体的な行動をとることが望ましいと考えられる。

表2は、250の観測データに基づき、当委員会が合意に達した上記各ゾーンの境界と、各バックテストの結果に対して監督当局が原則として採る措置を定めたものである。標本数が250以外の場合の境界線については、表1にあるように、真に99%の信頼水準をもつモデルにおける2項確率を計算することによって求められる。イエロー・ゾーンは、超過回数が一定値以下となる確率が95%以上のところを最低線とし、それ以上の領域を指す。表2は、こうした一定の超過回数それぞれに該当する累積確率を示したものであり、250の観測データの下、モデルの真の信頼水準が99%である場合、95.88%の確率で超過回数が5回以下となることが分かる。従って、イエロー・ゾーンは超過回数5回以上となる。

同様に、レッド・ゾーンは、超過回数が一定値以上となる確率が 99.99%以上のところを最低線とする。表 2 では、信頼水準 99%のモデルについて 250 の観測データをみると、同ゾーンは超過回数 10 回以上ということになる。

(d) グリーン・ゾーン

グリーン・ゾーンは余り説明を要しないであろう。信頼水準 99%のモデルでは、250 の観測データについて、大抵の場合、4 回までの超過回数が観察される可能性が非常に高いことから、バックテストの結果がこの領域に入ることを問題視する必要はなからう。これは、表 1 に示されているように、このゾーンにおける結果を容認しても、不正確なモデルを容認する確率が非常に低いことによって補強されよう。

(e) イエロー・ゾーン

イエロー・ゾーンは超過回数 5 ~ 9 回の範囲を指す。バックテストの結果が同ゾーンに入った場合、モデルは正確であるとも不正確であるとも評価できるが、一般には表 1 からモデルは不正確であることの蓋然性が高いといえよう。さらに、表 1 の結果をみると、超過回数が 5 回から 9 回に増えるにつれてモデルが不正確である可能性が強まることが指摘できる。

当委員会は、イエロー・ゾーンにおいて、一般に超過回数と所要自己資本額の増加の程度を結び付けることについて合意した。表 2 は、イエロー・ゾーンにおけるバックテスト結果から内部モデルによる所要自己資本額の算出において適用されるマルチプリケーション・ファクターの増加分を監督当局が決定するに当たり、当委員会で合意されたガイドラインを示したものである。

こうしたガイドラインは、内部モデル・アプローチに組み込まれたインセンティブの構造を維持する上で必要なものであり、とくに、超過回数が増えるにつれて監督上のペナルティーが増すことが重要である。表 1 は、一般に超過回数 9 回は 5 回に比べ問題があるという考え方を支持するものであり、表 2 に示されている超過回

数に対応した段階的なペナルティーに反映されている。

このように設定されたマルチプリケーション・ファクターの増加分は、同ファクターを増加させることによってモデルの精度を 99%相当にまで引き上げなければならない、という考え方に依拠している。例えば、250 の標本数のうち 5 つの超過回数は、98%の信頼水準を示唆するものである。従って、信頼水準 98%相当のモデルを 99%相当のモデルに変換するに足りるだけ、スケーリング・ファクターの水準を増加させる必要がある。言うまでもなく、こういった計算を厳密に行うためには、必ずしも全ての事態に当てはまるとはいえないような統計上の前提条件を追加的に設定することが必要である。例えば、トレーディング損益が正規分布を示すとの仮定の下では、信頼水準 99%のモデルによるリスク量と、同 98%のモデルによるリスク量の比率は約 1.14 であり、マルチプリケーション・ファクターが 3 の場合には、同水準を 0.40 程度引き上げる必要がある。もし実際の分布が正規分布ではなく「ファット・テイル」(fat tail)となっている場合、信頼水準 99%の基準を満たすには、さらにマルチプリケーション・ファクターの増加幅を引き上げる必要がある。こうしたファット・テイルに対する懸念は、表 2 におけるプラス・ファクターの増加幅を決める上での重要な要素の一つであった。

しかしながら、こうした段階的なプラス・ファクターの適用は純粹に自動的なものではないことを強調しておく必要がある。表 1 の結果からは、イエロー・ゾーンに入る結果が常にモデルが不正確であることを意味しない。当委員会は単に不運に見舞われた銀行にペナルティーを課すことは考えていない。もっとも、インセンティブが適切な形で作用するには、バックテストの結果がイエロー・ゾーンに入った場合、原則としてマルチプリケーション・ファクターが加算されなければならない。ただし、当該行が、係る措置が妥当なものでないことを示すことができる場合には加算が免除される。

言い換えれば、こうした状況においては、当該モデルに問題があるということについて監督当局が立証責任を負うべきではなく、銀行が自行のモデルの健全性を示

すべきこととするのが妥当である。この場合、銀行のモデルを評価する上で必要な追加的情報としては様々なものが考えられよう。

例えば、銀行のトレーディング業務の部門毎にバックテストの結果をみることは非常に重要となる。定期的なバックテストを実施している多くの銀行は、トレーディング・ポートフォリオ全体を、リスク・ファクターや商品別のトレーディング・ユニットに分解している。このように分解することにより、全リスク量として統合された段階において表面化した問題を、その源泉である特定のトレーディング・ユニットや特定のリスク管理モデルといったところまで遡って分析することが可能となる。

銀行は、バックテストの結果生じた超過回数のそれぞれについて原因を分析した文書を残す必要がある。こうした文書はイエロー・ゾーンにおけるバックテスト結果に対する監督上の適切な措置を決定する上で重要となる。また、銀行は、99%以外の信頼水準の下でバックテストを実施したり、当ペーパーでは規定されていないような手法でバックテストを行うこともあり得る。当然のことながら、こうした情報はリスク管理モデルを評価する上で非常に有用なものとなる。

実務上、バックテストにおける超過回数については、様々な説明が可能であり、中にはモデルの基本的な欠陥に起因すると看做することができるものもあれば、モデルの精度が不十分であることを示すものもある。また、単に不運だけであったり、日中取引で大損した場合もあろう。銀行のモデルから生じた超過回数をこうしたカテゴリーに分類した上で分析することは非常に有用であろう。

#### モデルに基本的な欠陥がみられる場合

- 1) 銀行のリスク管理モデルが、システム上、全ポジションのリスクを把握していない(例えば、海外拠点からのポジションの報告が不正確である等)
- 2) リスク管理モデルにおいてボラティリティや相関が正しく計算されていない(例えば、コンピュータの計算上 225 で割り算をすべきところを 250 で割って

いる等)

#### モデルに改善の余地がある場合

- 3) リスク管理モデルが、ある商品のリスクを十分正確に評価していない(例えば、期間区分が少なすぎたり、スプレッドを考慮していない等)

#### 不運、若しくは、モデルが予想していないような形で市場が変動した

- 4) 稀な事態(著しく低い確率の下で発生する事象)
- 5) モデルの予測範囲を超えて市場が大きく変動した(即ち、ボラティリティが想定していた以上に上昇する等)
- 6) モデルが予測した形で市場が連動しなかった(即ち、モデルが仮定した相関が大きく崩れる等)

#### 日中取引

- 7) 銀行のポジションに大きな変化が生じ(そして多大な損失が発生し)たり、日を跨いで(バリュー・アット・リスク値が計算された第1日目の終了後から、トレーディング損益が得られる翌営業日終了までの間で)何らかの大きな損益が発生した

一般に、リスク管理モデルの基本的な欠陥に関する問題は、最も深刻なものである。特定のトレーディング・ユニットにおいてこうしたカテゴリーに含まれるような超過回数が発生した場合、プラス・ファクターが適用されて然るべきである。こうした場合には、さらにモデルの抜本的な見直しや修正を行う必要があり、監督当局は、こうした対応が確実に行われるよう適切な措置をとらなければならない。

第2のカテゴリーの問題(モデルの精度に係る欠陥)は、大抵のリスク管理モデルについて多かれ少なかれ起こりうるものである。完璧な正確性を備えたモデルを構築することは不可能であり、全てのモデルには何らかの推計による部分が内在する。しかしながら、ある銀行のモデルが他の銀行のモデルに比べて、こうした問題が発生し易い傾向があると認められる場合には、監督当局は、プラス・ファクターを適用するとともに、そうした銀行に対し改善を促すような何か別のインセンティ



ブを考える必要がある。

第3のカテゴリの問題(モデルが予想しないような市場の大変動が生じるケース)も、バリュー・アット・リスク・モデルについて少なくとも一定の範囲内で発生しうることが十分予想される。特に、どんなに正確なモデルであってもトレーディング結果の100%を捕捉することは期待できない。モデルによって捕捉できない事態は必然的かつ不規則に1%は発生する。また、場合によっては、事前に推計したボラティリティや相関係数が適合しにくくなるような形で市場が変化することもある。いかなるバリュー・アット・リスク・モデルもこうした問題を抱えている。言い換えれば、将来の市場変動から生じ得るリスクを把握するために過去の市場変動に依存するといった、アプローチ固有の問題である。

最後に、バックテストを行うために用いられるトレーディング損益の定義によっては、日中取引や保有ポジションから生じ得る以外の異例のトレーディング損益の結果、超過回数が発生することもある。こうした事由によって超過回数が発生したとしても、銀行のバリュー・アット・リスク・モデルに必ずしも問題があるとは言えないが、こうしたことも監督当局の関心事となりうるものであり、プラス・ファクターが適用されることも考えられる。

トレーディング損益がモデルによって算出されたリスク値をどの程度上回ったかも有益な情報である。その他のすべての条件が等しい場合、トレーディング損益がモデルによって予測されたリスク値をはるかに超過した場合は、ほんの僅かしか超過していない場合に比べ問題は深刻であろう。

こうした点については、監督当局が銀行の所要自己資本を増額するか否かを決定する際、リスク管理体制に関する定性的基準の遵守状況などと同様に検討の対象とすることが想定されている。銀行から提供される追加的な情報に基づいて、監督当局は適切な措置を決定することとなる。

一般に、イエロー・ゾーンに入ったトレーディング結果について自己資本の積み増しを徴求することは、イエロー・ゾーン突入の原因が銀行のモデルを修正するこ

とによって回避可能であると監督当局が認識している場合には適切な対応策となる。これは、殆ど全てのモデルが予測できないような場合、具体的には、市場のボラティリティが予想外な水準まで一時的に高騰してしまった場合の対応とは全く対照的である。こうした出来事は、一種のストレス状態ではあるが、必ずしも銀行のリスク管理モデルを修正すべき問題とはいえない。また、モデルの根本に係る深刻な問題については、監督当局は、所要自己資本額を算出するに当たってモデルの使用を引続き認めるかどうかも含めて考える必要がある。

(f) レッド・ゾーン

最後に、バックテスト結果の評価において監督当局の判断の余地が存在するイエロー・ゾーンとは対照的に、レッド・ゾーンに突入したバックテスト結果（超過回数 10 回以上）については、一般に、銀行のモデルに問題があると看做される。これは、250 のトレーディング結果のうち超過回数が 10 回以上独立に発生することは、正確なモデルでは滅多に起こり得ないことと考えられるからである。

したがって、一般に、銀行のモデルがレッド・ゾーンに突入した場合、監督当局は自動的にそのモデルに対して適用するマルチプリケーション・ファクターを 1 だけ上積みし、結果として 3 から 4 に引き上げるべきである。言うまでもなく、監督当局は、当該行のモデルが何故多くの超過回数を発生させたかについて調査するとともに、銀行に対して即座にモデルを改善するよう要求する必要がある。

250 の観測データ数において 10 回という超過回数は極めて大きい。正確なモデルがそのように多くの超過回数を発生させることを正当に説明できる場合が極めて稀ながら存在する。特に、金融市場においてレジーム・シフト (Regime Shift) が生じた場合には、ボラティリティや相関係数が大きく変動すると考えられる。銀行が自行モデルにおけるボラティリティや相関係数の推計値を即座に更新しなければ、そうしたレジーム・シフトが原因となって短期間に超過回数がいくつも発生するような事態が起こり得る。しかしながら、これらの超過回数は本質的に全て同一の事由によって発生しているため、監督当局がとるべき適切な対応は、たとえ

10回の超過回数であったとしても、各々異なる原因によって10回の超過が発生している場合と同じ扱いにすべきではないであろう。例えば、こうした場合に監督当局が採りうる1つの対応策としては、モデルを更新する内部手続きの堅確性を維持しつつ、できる限り早急に、かかるレジーム・シフトを銀行のモデルに織り込むことを要求することとなる。

しかしながら、当委員会では、こうした特例措置は極めて異常な状況に限り適用され、(原則として)レッド・ゾーンに入ったバックテスト結果に対しては、自動的かつ裁量の介在する余地なく当該行の所要自己資本額が加算されるべきであると考えていることを、強調する。

#### ・結論

上記フレームワークは、マーケット・リスクに対応する所要自己資本額の算出に用いる内部モデル・アプローチの中にバックテストを組み込む上で、首尾一貫したアプローチを提供することを企図したものである。同フレームワークは、銀行自らが直面するリスクを計測する努力に拠るところが少なくないが、現在利用可能な手段に内在する限界を認識し、銀行に求められる負担やコストを最小限に抑えつつ、必要かつ適切なインセンティブを組み込むことを目的としている。

当委員会は、上記フレームワークが、こうした点に関し十分バランスのとれたものとなっていると確信する。それにも増して当委員会は、このアプローチが、銀行の(リスク管理における)パフォーマンスを評価する手法を、銀行監督上のガイドラインの中に一段と明確な形で取り込むための非常に重要な第一歩となったと確信している。

以 上

表 1

## モデルが正確な場合

## モデルが不正確な場合：信頼水準の代替的な水準

超過回数 (250の内)	信頼水準 = 99% exact type 1		超過回数 (250の内)	信頼水準 = 98% exact type 2		信頼水準 = 97% exact type 2		信頼水準 = 96% exact type 2		信頼水準 = 95% exact type 2	
0	8.1%	100.0%	0	0.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
1	20.5%	91.9%	1	3.3%	0.6%	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%
2	25.7%	71.4%	2	8.3%	3.9%	1.5%	0.4%	0.2%	0.0%	0.0%	0.0%
3	21.5%	45.7%	3	14.0%	12.2%	3.8%	1.9%	0.7%	0.2%	0.1%	0.0%
4	13.4%	24.2%	4	17.7%	26.2%	7.2%	5.7%	1.8%	0.9%	0.3%	0.1%
5	6.7%	10.8%	5	17.7%	43.9%	10.9%	12.8%	3.6%	2.7%	0.9%	0.5%
6	2.7%	4.1%	6	14.8%	61.6%	13.8%	23.7%	6.2%	6.3%	1.8%	1.3%
7	1.0%	1.4%	7	10.5%	76.4%	14.9%	37.5%	9.0%	12.5%	3.4%	3.1%
8	0.3%	0.4%	8	6.5%	86.9%	14.0%	52.4%	11.3%	21.5%	5.4%	6.5%
9	0.1%	0.1%	9	3.6%	93.4%	11.6%	66.3%	12.7%	32.8%	7.6%	11.9%
10	0.0%	0.0%	10	1.8%	97.0%	8.6%	77.9%	12.8%	45.5%	9.6%	19.5%
11	0.0%	0.0%	11	0.8%	98.7%	5.8%	86.6%	11.6%	58.3%	11.1%	29.1%
12	0.0%	0.0%	12	0.3%	99.5%	3.6%	92.4%	9.6%	69.9%	11.6%	40.2%
13	0.0%	0.0%	13	0.1%	99.8%	2.0%	96.0%	7.3%	79.5%	11.2%	51.8%
14	0.0%	0.0%	14	0.0%	99.9%	1.1%	98.0%	5.2%	86.9%	10.0%	62.9%
15	0.0%	0.0%	15	0.0%	100.0%	0.5%	99.1%	3.4%	92.1%	8.2%	72.9%

注 上表は、実際の信頼水準に関し幾つかの仮定を置いた上で、250個の独立した観測データについて、トレーディング結果がモデルによるリスク量を超過する回数（超過回数）が生じる正確な確率と、これらの正確な確率におけるタイプ1及びタイプ2のエラーが生じる確率を示している。

表の左部分には、モデルが正確であり、実際の信頼水準が99%である場合が示されている。すなわち、いかなる観測期間についても超過回数が発生する確率が1%（100% - 99% = 1%）となる場合である。“exact”と記された欄は、この仮定の下で250個の独立した観測データからある水準の超過回数が発生する正確な確率を示している。“type 1”の列は、モデルを排除する基準として250個の独立した観測データから発生する超過回数の水準を用いた場合に、正確なモデルを誤って排除してしまう確率を示している。例えば、モデルを排除する基準を、250個の独立した観測データにつき超過回数5回以上と規定した場合、正確なモデルを誤って排除してしまう確率は10.8%となる。

右部分は、不正確なモデルについて示したものである。特に、上表では、実際の信頼水準が98%、97%、96%、95%である不正確なモデルに限定されている。各々の不正確なモデルに対して、“exact”の欄は、この仮定の下で250個の独立した観測データからある水準の超過回数が発生する正確な確率を示している。“type 2”の列は、モデルを排除する基準として250個の独立した観測データの中から発生する超過回数の水準を用いた場合に、不正確なモデルを誤って容認してしまう確率を示している。例えば、モデルを排除する基準を、250個の独立した観測データにつき超過回数5回以上と規定すると、信頼水準97%におけるtype 2の欄によれば、97%の信頼水準しか有していないモデルを誤って容認してしまう確率は12.8%となる。

表 2

	超過回数	スケーリング・ファクターの増分 (プラス・ファクター)	累積確率
グリーン・ゾーン	0	0.00	8.11%
	1	0.00	28.58%
	2	0.00	54.32%
	3	0.00	75.81%
	4	0.00	89.22%
イエロー・ゾーン	5	0.40	95.88%
	6	0.50	98.63%
	7	0.65	99.60%
	8	0.75	99.89%
	9	0.85	99.97%
レッド・ゾーン	10 以上	1.00	99.99%

## 注：

上表は、グリーン、イエロー、レッドの各ゾーンを定義したものであり、監督当局はこれに基づいて、マーケット・リスクに対する所要自己資本額を算出するために用いられる内部モデル・アプローチに付随する、バックテスト結果について評価することとなる。この表に示された境界線は、250 個の観測データを基準にしたものである。観測データ数がそれ以外の場合については、イエロー・ゾーンは累積確率が 95%以上の点から、またレッド・ゾーンは同 99.99%以上の点を各々基点とする。

累積確率とは、実際の信頼水準が 99%である場合に 250 個の観測データにおいて、超過回数が一定水準以下となる確率をいう。例えば、超過回数 4 回に対応する累積確率は、超過回数が 0 回から 4 回までとなる確率である。

注意が必要なのは、こうした累積確率や表 1 に示されているタイプ 1 エラーの発生確率を合計しても 1 にはならないことである。何故なら、特定の超過回数に対応する累積確率には、丁度その超過回数となる確率が含まれているからである。これはまた、タイプ 1 のエラーについても同様である。したがって、これら 2 つの確率値を合計した場合、丁度その超過回数となる確率の分だけ 1 を超えることになる。