



ID	JJF00275
----	----------

論文名	積立証券投資の成功確率 －カーネル密度ブロックブートストラッピング法の適用－
	The probability to realize a life planning of individual investors through an installment security investment: A non-parametric application of the block bootstrapping and kernel density
著者名	高見茂雄
	Shigeo Takami
ページ	83-100

雑誌名	経営財務研究
	Japan Journal of Finance
発行巻号	第30巻第1.2合併号
	Vol.30 / No. 1.2.
発行年月	2011年3月
	Mar. 2011
発行者	日本経営財務研究学会
	Japan Finance Association
ISSN	2186-3792

■ 論 文

積立証券投資の成功確率 —カーネル密度ブロックブートストラッピング法の適用—

高見 茂雄
(立正大学)

要 旨

非正規・非 IID 性時系列を示す TOPIX ETF への積立証券投資において、ターゲット年月以前に目標積立額に到達する成功確率をノンパラメトリック手法(カーネル密度ブロックブートストラッピング法)で求めた。時系列特性を反映しないナイーブなモデルとは到達確率値は大きく乖離する結果が得られた。

キーワード：積立証券投資，資産形成層，非正規性，非 IID 性，カーネル密度ブロックブートストラッピング法

1 はじめに

資産形成層¹が行う持株会投資や累積投資契約は同一対象の証券に每期均等額拠出する積立証券投資の典型例である。資産形成層は、ライフプラン上の将来の住宅取得や教育資金確保などの明確な目的をもつものの、投資経験と投資知識は少なく、かつまとまった資産を保有していないので、給与を原資とする積立証券投資という形態が大方選択できる投資手法である。このような特質をもつ資産形成層は、ポートフォリオ理論が前提とするリスク回避的なプロの機関投資家とは異なり、一方では積立期間中のダウンサイドリスクに敏感であるものの、反面アップサイドリターンも期待する。Chhabra(2005)は、個人投資家のプロフィールを分析し、最低限の生活水準を犠牲にしないことを一義にし、運用中に元本割れが生じないことを求めると同時に、他方では一段階上の生活水準を目指し、たとえ少ない確率であっても期中にアップサイドのリターンも求める相反するフレームワークを提示した。これを受け、積

1 資産形成層という用語は高橋(2007)が定義しているように、ファイナンシャルプランニングの分野で用いられており、住宅取得や子供の教育などの、人生のイベントがこれから待ちかまえている 20~30 代の世代を指す。これに対し、50~60 代の資産活用層はこれまで築いてきた資産をどう取り崩し豊かな生活を実現するかが課題となる。本論文では、資産形成層を念頭において、議論を進める。

立証券投資と同義のドルコスト平均法投資の分野で、Dubil(2005)はシミュレーションを用いて投資終価の損失確率を求め、Trainor Jr.(2005)は First passage time probability(はじめて目標積立額に到達する確率)公式を適用、運用期中の損失確率を求めた。ただし、両者とも一括投資法とドルコスト平均法との伝統的二者択一問題で考えている点、ダウンサイドリスクのみしか取り上げていない点で限界がある。これに対し、高見(2009)は、アップサイドリターンイベントを目標積立額にターゲット年より早期に到達すること、ダウンサイドリスクイベントを運用中期に抛出額割れを起こすことと定義し、両者のイベントを勘案した確率を考察した。ただし、取り上げた TOPIX 時系列対数階差データに対し、定常性、正規性、IID 性の厳格な前提をおいたこと、簡単のために年 1 回サイクルの積立の前提をおいた点で課題を残す。実際の給与天引き積立証券投資では月次積立が多く、対応する月次 TOPIX 時系列対数階差データは、定常ではあるものの、非正規、非 IID 性の性質を示すことが多いからである。本論文の課題は、これら非正規性、非 IID 性の性質に対応した積立投資の運用期中における確率を考察することである。ただし、本論文では、焦点を技術的観点におくため、問題をアップサイドリターンイベントである目標積立額に早期到達する確率にしぼり、ダウンサイドリスクイベントは扱わない²。

本論文の掲げた課題のうち、非正規性への対応は、GARCH や Regime Switching をはじめ、分布関数を特定するパラメトリックモデルでも多くの研究蓄積がある。しかし、Li and Racine(2007, 序文 xvii)が指摘するように、分布関数の事前知識がある場合はパラメトリックモデルによって最良推定量が得られるが、その前提が間違っていた場合頑健性を欠く。また、分布関数の事前知識は *divine insight* といわれるように神のみぞ知る知識であり、現実には正確に分布関数を識別できない。そこで、本論文では、月次 TOPIX 時系列対数階差データの非正規性に対し、ノンパラメトリックモデルで対応する。ただし、多様なノンパラメトリック手法の中から、本論文は、カーネル密度法とブートストラッピング法にしぼり考察を進める。

他方、非 IID 性に対しては、ブロックサンプリング法で対応する。Härdle(2003)や Politis(2003)が指摘するように、時系列データにブートストラッピング法を適用する場合、暗黙に IID 性を前提にしている。しかし、時系列に自己相関がみられる場合は、その自己相関を反映したリサンプリングが求められる。その手法として、ブロックブートストラッピング法や sieve ブートストラッピング法が開発されてきたが、後者はパラメトリック時系列を前提とするので、本論文は前者を採用する。しかし、ブロックブートストラッピング法にも、欠点はある(リサンプリングブロックは自己相関を完璧に反映せず依存関係が弱くなること(Chernick(2008, p.105)), 漸近が遅いこと(Horowitz(2003, p.1051))など)。

これらのブロックブートストラッピング法の欠点に対処した手法の一つには、Paparoditis and Politis(2001), Horowitz(2003), Manzan and Zerom(2008)らが提唱した、マルコフ推移ブロックブー

2 高見(2009)は、個人投資家の性向をメインテーマとし、最終的にアップサイドリターンとダウンサイドリスクとの同時確率も求めているが、本論文は、技術的観点から、非正規非 IID 性をもつ時系列データをいかにモデル化し、成功確率を定量化するかに問題意識の焦点をおく。そのため、問題が取扱いやすく、資産形成層が運用を停止するイベントになじみやすいアップサイドリターン、すなわちターゲット年月以前に目標積立額に早期到達する成功確率に限定して分析を進める。ダウンサイドリスク発生確率も同様のシミュレーション手法で算出可能であるが、今後の課題とする。

トストップ法がある。それぞれ、手法の細部でバリエーションの違いは見られるものの、カーネル密度確率でブロックをリサンプルする点で共通している。その中から、本論文は、アルゴリズムが明示的に示されている Manzan and Zerom(2008) に準拠し、カーネル密度ブロックブートストラッピング法を適用することで、運用中に目標積立額に到達する成功確率を考察する。

なお、ノンパラメトリック手法を用いた金融時系列実証分析は数多くあるが、ブロックブートストラッピング法を適用した分析は数少ない。Mukherji(2008) は証券ポートフォリオのダウサイドリスク予測を行っている点で意義があるが、ブロックブートストラッピング法を十分吟味して適用しているとはいえない。一方、今野ら(2007) もブロックブートストラッピング法を用いて日次 TOPIX データを分析しているが、目的をカオスの検証においており、本論文の目的とは異なる。そのため、積立証券投資の成功確率にカーネル密度ブロックブートストラッピング法を適用した先行研究は皆無と思われる。

本論文の構成は以下の通りである。2 節では、積立証券投資の将来価値のキャッシュフローを明示し、先行研究の結果とその限界を述べ、3 節では対象とする TOPIX データの特徴を述べる。4 節では検討するそれぞれのノンパラメトリックモデルの特徴と計算手順を説明し、5 節ではその分析結果にもとづき考察を進める。6 節では真の分布・時系列構造を想定し分析結果の正確性を検討する。そして、7 節では結論ならびに今後の課題を述べる。

2 積立証券投資の将来価値

2 節では本論文が扱う積立証券投資の将来価値ならびに目標積立額に早期到達する確率を定義し、高見(2009)の研究結果とその限界を示す。

2.1 積立証券投資の将来資産価値の定義

資産形成層が定期的に 1 円を同一対象に積立証券投資を行い、 $t(\leq T)$ 期までに T 円以上の資産を作る状況を想定する³。証券のリターン $\tilde{\varepsilon}_k$ を確率変数とすれば、積立証券投資の t 期における将来資産価値 \tilde{X}_t は(1)式で表わされる。

$$\tilde{X}_t = e^{\tilde{\varepsilon}_1 + \tilde{\varepsilon}_2 + \dots + \tilde{\varepsilon}_t} + e^{\tilde{\varepsilon}_2 + \tilde{\varepsilon}_3 + \dots + \tilde{\varepsilon}_t} + \dots + e^{\tilde{\varepsilon}_t} \quad (1)$$

(1) 式は t 個の項から構成される多項式である。 t 期近くのリターンは多くの項に含まれるが、反対に第 1 期の投資開始時のリターンは最初の項にしか含まれていない。そのため、 t 期近くのリターンに大きく依存し、反対に 1, 2 期などの投資開始時のリターンのウエイトは小さい。これは、たとえば、投資初期に高いリターンが続いても t 期近くで大きなマイナスリターンが続けば、拠出額割れが生じるおそれがある。すなわち、経路依存性が強い。投資開始時には、ウエイトの高い t 期近くのリターンの予測は難しく、かつ次第に明らかにされる時系列 $\tilde{\varepsilon}_k$ の情報も t 期リターンの不確実性を減少させることに

3 簡単化のため 1 円と設定している。現実的には、たとえば毎月 10 万円積立て、10 年間で 1,200 万円の財産形成を目標とすると読み替える。また $\exp(\varepsilon_1 + \varepsilon_2)$ とし、 $(1 + \varepsilon_1)(1 + \varepsilon_2)$ と複利将来価値で設定しないのは、確率変数 $\tilde{\varepsilon}$ の値によって、将来価値がマイナスになる事態を防ぐ目的からであり、デリバティブプライシングでポピュラーに用いられている設定である。

は役立たない。投資知識や経験の少ない資産形成層にとって、将来価値 \tilde{x}_t を逐次的に制御することは難しく、 t 期に目標額円に到達した場合（アップサイドリターンイベント）や、運用中期でも拠出額割れが続く場合（ダウンサイドリスクイベント）などに以降の積立拠出を停止することが、唯一とれる制御手段といえる。そのため、運用中にはじめて目標積立額に到達すれば、一段階上の生活水準が確保され、退出の意思決定を行うと想定することは資産形成層の投資プロファイルに合致している。

結局、資産形成層がとれる意思決定は、積立証券投資の運用開始と退出のみに限られる。ただし、投資知識や経験が少ないといっても、運用開始時に参考のできる情報がある。ひとつにはシミュレーションにより発生させたサンプルパスをグラフィカルに表示した情報である。それは台風の予測進路にたとえられる。ただし、台風の現在いるところより遠く離れた地域の範囲は広大に表示され、該当地域の住民はどの程度の災害を被るかわかりにくい。その点、天気予報の降雨確率のような数値情報は有益な情報である。そこで、投資開始時において、 t 期ごとに目標積立額円に到達する確率を求めることは、単に期待値や標準偏差を求める以上の意義がある⁴。 t 期に目標積立額 T 円に到達する確率は(2)式、そして目標積立額に t 期にはじめて到達する確率は、独立性の前提をおかない一般的な場合は(3)式で、独立性の前提をおく特別な場合は(4)式で定義できる。前述のように、本論文の研究対象である資産形成層は一段階上の生活水準を求め、積立証券投資の運用中に目標積立額に到達すれば、以降の運用を停止し退出すると想定するのは自然である。運用のプロではない彼らは、運用を続けることで、いったん到達した T 円を割込むリスクを避けたいがためである。そのため、(3)または(4)式のはじめて到達する確率の数値情報は特に意義がある。また、単に何パーセントという数値だけでなく、運用終期よりどのくらい早期に退出できるのかという年度情報も有益である。

$$P(\tilde{X}_t \geq T) = 1 - P(\tilde{X}_t \leq T) \tag{2}$$

$$P(\tilde{X}_1 \leq T, \tilde{X}_2 \leq T, \dots, \tilde{X}_{t-1} \leq T, \tilde{X}_t \geq T) = P(\tilde{X}_1 \leq T)P(\tilde{X}_2 \leq T | X_1 \leq T) \times \dots \times P(\tilde{X}_{t-1} \leq T | X_1 \leq T, X_2 \leq T, \dots, X_{t-2} \leq T) \{1 - P(\tilde{X}_t \leq T | X_1 \leq T, X_2 \leq T, \dots, X_{t-1} \leq T)\} \tag{3}$$

$$P(\tilde{X}_1 \leq T, \tilde{X}_2 \leq T, \dots, \tilde{X}_{t-1} \leq T, \tilde{X}_t \geq T) = \left\{ \prod_{k=1}^{t-1} P(\tilde{X}_k \leq T) \right\} \{1 - P(\tilde{X}_t \leq T)\} \tag{4}$$

2.2 先行研究結果と限界

このような問題意識から、高見(2009)は年1回1円のTOPIX ETF積立投資で、10年間で10円の資産形成を目的とし、1~10年後の各期の到達確率とはじめて到達する確率を数理的に導出している。ただし、簡単のため証券リターン $\tilde{\varepsilon}_k$ に、定常性、正規性($N(0, 0.2)$)、IID性の前提をおいている。表1は、その前提での算出結果を示している。②欄が(2)式の特定期点 t 期末に目標積立額円に到達する確率、④欄が独立性を前提とした(4)式にもとづく t 期末にはじめて到達する確率を示している。

4 個人投資家の投資目的以外にも、たとえばVaRや倒産確率のように、期待値より確率密度を求めることに意義のある分野は多い。

表 1 アップサイドリターンイベント確率

t	1	2	3	4	5	6	7	8	9	10
① t 時点確率密度	1.000	1.000	1.000	0.998	0.979	0.919	0.821	0.707	0.599	0.500
② t 時点での到達確率	0.000	0.000	0.000	0.002	0.021	0.081	0.179	0.293	0.401	0.500
③ 累積確率密度	1.000	1.000	1.000	0.998	0.977	0.898	0.737	0.521	0.312	0.156
④ はじめて到達する確率	0.000	0.000	0.000	0.002	0.021	0.079	0.161	0.216	0.209	0.156
⑤ 到達確率合計	0.844									
⑥ 平均到達年	8.156									

表 1 から、「10%を超える到達確率を期待するには、少なくとも 7 年間運用を続ける必要があり、ピークは 10 年末の 0.5 であること (②欄)、はじめて到達する確率のピークは 8 年末の 0.216 であり、以降は減少すること (④欄)」が読み取れる。これらから、仮に 8 年末までに早期に到達すれば、以降資産価値の増加を期待するよりは運用を停止する方が得策であること、8 年末までに 1 回も到達しない場合は待ち続けることにリスクがあることなどのインプリケーションが得られる。

しかし、高見 (2009) は厳格な前提をおいている点で問題を残している。たとえ証券リターンは将来の事象を表すにしても、TOPIX の過去データ情報は重視されるべきであり、正規性と IID 性の前提は十分吟味する必要がある。また、給与生活者が行う積立証券投資は通常毎月のサイクルであるにもかかわらず、年 1 回の積立サイクルに設定したことにも問題がある。そこで、3 節では TOPIX の過去データについて吟味し、これらの問題点に対応する。

3 分析対象データ

本論文も先行研究と同様、10 年間の積立証券投資を分析対象とするが、考察の出発点として、積立サイクルを考える。現実に合わせて、月 1 回 120 回の積立 TOPIX 投資と設定すべきか、単純化して年 1 回 10 回の積立投資の設定で許容されるのか、月次と年次の TOPIX 対数階差の過去データの態様を調べ、両者が類似していれば許容されると判断する。まず、表 2 は月次 241 ヶ月の TOPIX 過去データから、240 個の対数階差を対象にその統計値を掲げている。241 ヶ月の標本数を調べた理由は、10 年間 120 年間の予測を行う場合には、その 2 倍程度の過去データを重視すべきと考えたからである。

表 2 月次 TOPIX 対数階差データ統計値

対象期間	平均 リターン (年率)	平均 リターン (月率)	標準偏差 (年率)	標準偏差 (月率)	歪度	尖度	Jarque- Bera p 値
① 1982/1-2001/12	2.97%	0.2472%	19.15%	5.5281%	-0.28	4.182	0.000
② 1983/1-2002/12	1.75%	0.1462%	20.00%	5.7727%	-0.25	4.003	0.002
③ 1984/1-2003/12	1.77%	0.1479%	20.16%	5.8183%	0.06	3.896	0.006
④ 1985/1-2004/12	1.15%	0.0959%	19.90%	5.7453%	-0.22	3.996	0.003
⑤ 1986/1-2005/12	2.27%	0.1894%	20.12%	5.8072%	-0.23	3.899	0.006
⑥ 1987/1-2006/12	0.37%	0.0305%	19.71%	5.6895%	-0.25	3.997	0.002
⑦ 1988/1-2007/12	-0.78%	-0.0652%	19.12%	5.5209%	-0.24	4.198	0.000
⑧ 1989/1-2008/12	-5.05%	-0.4205%	20.13%	5.8124%	-0.38	4.458	0.000

表 2 を観察すれば、標準偏差は年率約 20% で安定的ではあるものの、対象期間すべてにわたり、Jarque-Bera p 値がすべてゼロ近く、正規分布を想定する帰無仮説は棄却される。

これに対し、表 3 は TOPIX 開始以来の最大限の 23 個の標本数をもとに、各月末ごとの年次対数階差データの統計値を示している。

表 3 年次TOPIX対数階差データ統計値

1949or 1950- 2008	平均 リターン (年率)	標準偏差 (年率)	歪度	尖度	Jarque- Bera p 値
1月末	8.30%	23.68%	0.33	3.202	0.56
2月末	8.13%	22.08%	0.34	2.880	0.55
3月末	8.14%	21.63%	-0.08	2.197	0.45
4月末	8.31%	21.16%	-0.09	2.292	0.52
5月末	7.09%	22.91%	-0.70	3.994	0.03
6月末	7.34%	22.57%	-0.54	3.024	0.24
7月末	7.41%	20.78%	-0.25	2.561	0.58
8月末	7.14%	20.02%	-0.01	2.790	0.95
9月末	7.02%	23.10%	-0.39	3.111	0.47
10月末	6.83%	23.61%	-0.32	3.538	0.42
11月末	6.95%	24.11%	-0.01	4.239	0.15
12月末	7.12%	24.67%	0.04	3.323	0.87

表 3 でも標準偏差はほぼ 20% で安定的ではあるが、表 2 とは対照的に Jarque-Bera p 値は 5 月末データを除きすべて棄却されず、正規分布とみなすことができる。表 2 と 3 の観察結果から、月次と年次では統計値と分布の態様は異なり、高見 (2009) のおいた年 1 回積立サイクルの前提は適切ではなく、月次データに限定してさらに調べる必要があり、以降月次積立をサイクルに考察を進める。

表 2 の月次データにもどり、まず定常性を確認すると、表 2 のすべての期間にわたって、ADF, PP, ERS 単位根検定は棄却され、定常性は前提としてよいといえる。次に、IID 性の確認であるが、コロログラムを調べたところ、Ljung-Box 統計量では自己相関は認められず、AR または MA のパラメトリック構造は同定できない。ただし、相関というよりも独立性を意識し、BDS 検定を試みたところ⁵、表 2 ⑥、⑦、⑧の対象期間では独立であるという帰無仮説は次元 3 まで棄却され、少なくとも、3 次の独立性は否定して考察すべきという結論になる。そこで、本論文では、いちばん最近のデータセット⑧にもとづき、非正規、非 IID 性 (次元 3 の自己相関の非独立性、ただし分布の同一性は仮定する) を前提におき、積立証券投資の成功確率を求める。ただし、4 節に移る前に、以下では表 2 ⑧欄の対象期間 TOPIX 対数階差データの特性を述べる。

5 Brock, Dechert, Scheinkman and LeBaron 検定の略称である。IID なら任意の 2 点の距離が想定した誤差以下になる確率は一定であることを原理としており、特定のパラメトリックモデルの前提は必要としない。Diks and Manazan(2002, p.10 ダウンロード版)によれば、はじめて棄却されない次元の直前次元の自己相関があるとみなすことができる。

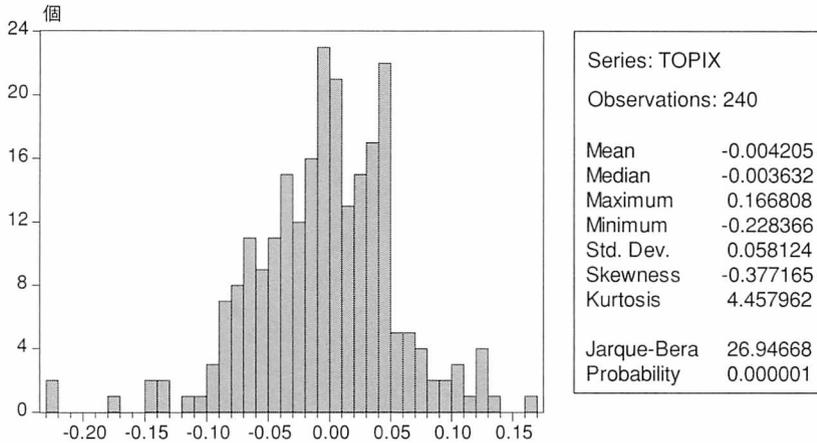


図 1 対象データの特徴

図 1 のヒストグラムは左に裾を引き、2 峰型で、間隔をおいた両端に異常値のある特異な形状を示している。両端の頻度は -0.05 以下が 47 個、 -0.1 以下が 9 個、 0.1 以上が 10 個、 0.05 以上が 28 個含まれており、両端の値が均等確率で出現する場合は、マイナス値の方が絶対値が大きく、ダウンサイドに振れやすい性質にある。表 4 は絶対値が 0.1 以上の異常値をその出現時点とともに示している。

表 4 両端値出現実績

最小値から	月次リターン	最大値から	月次リターン
1990年 9月	-0.2284	1990年10月	0.1668
2008年10月	-0.2264	1991年 2月	0.1361
1993年11月	-0.1712	1992年 8月	0.1278
2008年 9月	-0.1431	1994年 1月	0.1239
1990年 3月	-0.1413	1993年 4月	0.1239
1990年 8月	-0.1320	1999年 3月	0.1235
1998年 8月	-0.1315	2008年 4月	0.1134
1990年11月	-0.1165	1995年 7月	0.1099
1992年 6月	-0.1074	1993年 3月	0.1088
		2005年 9月	0.1052

4 分析対象モデルの特徴と計算手順

3 節の TOPIX 対象データの考察を踏まえ、本論文は非正規性と非 IID 性の双方の性質を考慮し、カーネル密度ブロックブートストラッピング法を適用し、積立証券投資の早期達成成功確率を求めることを目的とする。加えて、非正規性と非 IID 性の一部しか考慮しない他のモデルも検討し、モデル間での成功確率を比較する。比較することで、誤識別モデルにはどのようなバイアスがかかるかを確認するためである。そのため、本論文では 6 種類のモデルを検討するが、それらの位置関係を図 2 で示す。

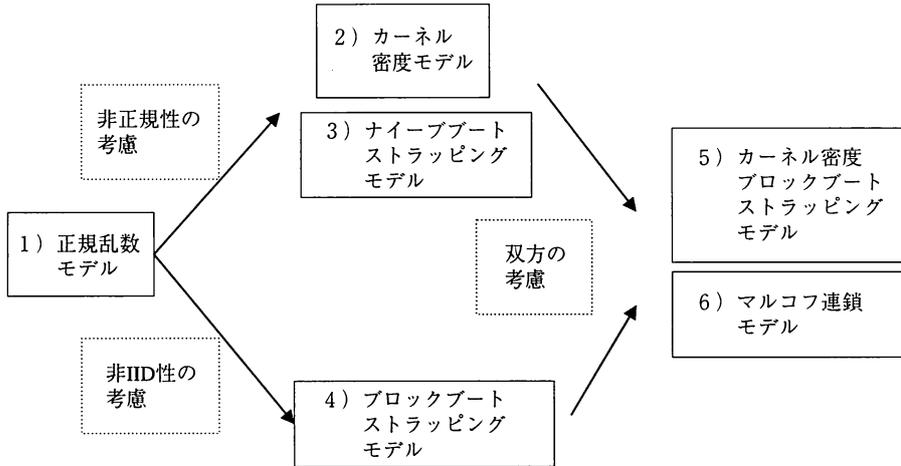


図2 検討モデルの相互関係

図2にそってモデル間の関係を説明する。1) 正規乱数モデルは表2⑧欄の情報から平均値と分散の情報だけしか使用していない点で、非正規性と非IID性とも考慮しない全くの誤識別モデルと位置づけられる。1)モデルに非正規性を考慮したモデルは図1の中央上の2)カーネル密度モデルと3)ナイーブブートストラッピングモデルである。これに対し、非IID性を考慮したモデルが中央下の4)ブロックブートストラッピングモデルである。そして、双方の性質を考慮したモデルが、本論文が最終的に適用する5)カーネル密度ブロックブートストラッピングモデルと6)マルコフ連鎖モデルである。以上のモデルのうち、1)のみがパラメトリックモデルで、他のすべてがノンパラメトリックモデルに分類される。

次に、それぞれのモデルの特徴を説明する。ただし、便宜のため分析対象とする表2⑧欄のデータセットを(5)式で定義する。(5)式では、過去のTOPIX対数階差を記号 y で表し、現時点を2009年1月と想定し、現時点に近い方から逆向きに対数階差の番号 s をつけている。これに対し、将来の予測リターンは記号 ε で、将来の時点は t で表している。なお、本論文は、すべてのモデルでシミュレーション手法を用いて分析する。すなわち、乱数にもとづきサンプルパスを発生させ、目標積立額120円に到達したパスをカウントすることで到達成功確率を求める。その際、発生させるサンプルパスの本数が問題になる。Mukherji (2008)は1,000本に限定しているのが、1,000本ではランダムシード間のばらつきが顕著であったため、本論文では頑健性を担保すべく10,000本発生させている。

$$\begin{aligned}
 y_1 &= \log(TOPIX_{2008.12}) - \log(TOPIX_{2008.11}) \\
 y_2 &= \log(TOPIX_{2008.11}) - \log(TOPIX_{2008.10}) \\
 &\dots\dots\dots \\
 y_s &= \log(TOPIX_{2008.12から逆に数えてs-1番目}) - \log(TOPIX_{2008.12から逆に数えてs-2番目}) \\
 &\dots\dots\dots \\
 y_{240} &= \log(TOPIX_{1989.1}) - \log(TOPIX_{1988.12}) \quad (s=1,2,\dots,240)
 \end{aligned} \tag{5}$$

1) 正規乱数モデル:表2⑧欄の平均リターンと標準偏差のデータをパラメータとして設定(月次ベースで $N(-0.004205, 0.058124)$), 乱数を発生させる。それを証券リターン ε_k に使用し(1)式に代入, 積立証券投資将来価値 $X_t(t=1, 2, \dots, 120)$ のサンプルパスを10,000本作る。

2) カーネル密度モデル: データセット $(y_1, y_2, \dots, y_{240})$ は, 表 4 の通り, 最小値 -0.2284 から最大値 0.1668 の間に分布するが, 最小値 -0.230 から最大値 0.168 の間に均等間隔 0.008122 合計 50 個の点 $(z_1 = -0.23, z_2, \dots, z_i, \dots, z_{50} = 0.168)$ を作成, 点 z_i に対応する確率密度 $\text{pdf}(z_i)$ をカーネル密度法により求める。そして, その密度関数にしたがい乱数を発生させ, k 番目に発生した証券リターン $\varepsilon_k = z_i^k$ とおき, (1)式に代入する。1本のサンプルパスは 120 の証券リターン ε_k が必要なので, トータル 1,200,000 回の乱数を発生させる。ただし, ソフトウェア R のパッケージ **ks** のうち **kde** 関数を使用して求めた。Duong(2007)は, 自ら作成した **kde** 関数を説明しているが, カーネル関数にはガウス型が内蔵されており, 他の関数は選べない。ただし他の関数を使っても密度推定にはあまり違いはないといわれる⁶。これに対し, バンド幅には複数のオプションが用意されているが, 3 種類のクロスバリデーション法のうちスムーズクロスバリデーション法がデフォルト設定なので, この方法を選択した⁷。また, 240 点の標本点 y_s ではなく, 50 個の均等間隔点 z_i を新たに設定しその場所の密度を求めたのは, 標本点を発生させるブートストラッピング法と峻別するためであり, 点の数を減らす方が密度の変化がより顕著になると考えたためである。図 3 のグラフはカーネル密度推定結果を示している。図 1 のヒストグラムと比較すると, 左に裾を引いている点は保存されているが, 2 峰型が単峰型に修正され, 両端の出現密度が低くおさえられている。

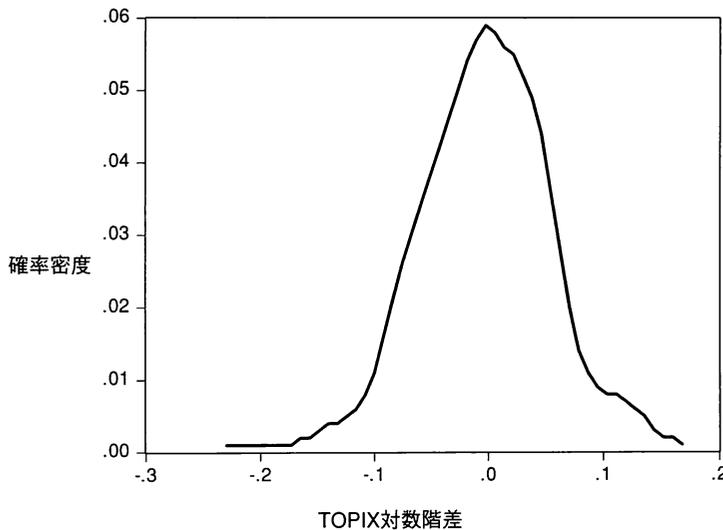


図3 カーネル密度推定結果

6 Li and Racine (2007, p.15) 参照。

7 バンド幅の選択は盛んに議論されている。Li and Racine (2007, p. 14) は, 最小二乗誤差が最小値をとる $1.06(\text{標準偏差})/(\text{標本数})^{1/5}$ の固定バンド幅を提唱している。次に, バンド幅が小さくとるとカーネル密度は多くのこぶのある変動の大きな分布を示し, 逆に大きくとるとなめらかになる半面, 特徴の少ない分布を示す。これに対し変動バンド幅の手法は, ローカルなデータ個数に依存させてバンド幅を決定させる方法で, プラグイン法とクロスバリデーション法がある。

3) ナイーブブートストラッピングモデル：標本 $(y_1, y_2, \dots, y_{240})$ から復元を含めて無作為抽出（リサンプリング）する。抽出の順番にしたがい証券リターン ε_k の列ができる $(\varepsilon_1 = y_1^*, \varepsilon_2 = y_2^*, \dots)$ 。ここで、 y_k^* は k 番目の無作為抽出で出現した値を表し、 $1/240$ の確率で y_k が当たる。以下の手順は同様である。なお、3)~6) のモデルでの計算は R の base パッケージの sample 関数を用いている。

4) ブロックブートストラッピングモデル：3) モデルでは、無作為抽出後の列 y_1^*, y_2^*, y_3^* …はばらばらに並んでおり、自己相関関係が反映されていない。それに対し、3 節で検討した 3 次元の自己相関関係を 3 次元のベクトル列 $Y_s = (y_{s+2}, y_{s+1}, y_s)$ で導入する。すなわち、(6) 式の 238 個のベクトル列が得られる⁸。

$$\begin{aligned} Y_1 &= (y_3, y_2, y_1)' \\ \dots\dots\dots \\ Y_s &= (y_{s+2}, y_{s+1}, y_s)' \\ \dots\dots\dots \\ Y_{238} &= (y_{240}, y_{239}, y_{238})' \end{aligned} \tag{6}$$

このベクトル列に関して復元無作為抽出を行い、抽出の順番にしたがい、(7) 式の証券リターンの列ができる。以下の手順は同様である。ただし、1 本のサンプルパスを得るには、40 個のベクトル列で足りるため、10,000 本のサンプルパスは 400,000 回の無作為抽出で得られる。ブロックブートストラッピングには 1 節で述べた欠点があるが、それに加え MacKinnon(2006, S15) が指摘するように、ベクトルの要素数が少ない場合不連続点が発生することがあげられる。(7) 式で、 $\varepsilon_3 = y_1^*, \varepsilon_4 = y_6^*$ という並びが現れるが、両者とも同じベクトルに含まれない場合、もともとの標本になかった並びが発生する。

$$\begin{aligned} (\varepsilon_1, \varepsilon_2, \varepsilon_3)' &= Y_1^* = (y_3^*, y_2^*, y_1^*)' \\ (\varepsilon_4, \varepsilon_5, \varepsilon_6)' &= Y_2^* = (y_6^*, y_5^*, y_4^*)' \\ \dots\dots\dots \end{aligned} \tag{7}$$

この点を修正し、なめらかさを加えるひとつの手段として、ベクトル列 Y_s の発生に確率を割当る手法が考えられる。それが、カーネル密度ブロックブートストラップ法であり、5) と 6) のモデルにあたる。

5) カーネル密度ブロックブートストラッピングモデル：5) と 6) のモデルでは 4) モデルに該当するベクトル列 Y_s の発生に確率を割当てる点で共通しているが、5) は 6) モデルの簡略版という位置づけにある。5) モデルでは、238 個のベクトル列 $Y_s = (y_{s+2}, y_{s+1}, y_s)$ の 3 次元の各点のカーネル密度を求める。そして、その確率密度でベクトル列のブートストラッピングを行う。以下の手順は同様である。ただし、2) モデルでは新たに均等間隔点で確率密度を設定したが、5), 6) モデルではブートストラッピングを行うので、標本点でなければならない。ここでも、kde 関数を用い、スムーズクロスバリデーション法バンド幅を選択した。4) モデルの難点は完全には解決しないが、発生頻度の高いベクトル同士が並ぶ

8 (5) 式のベクトル列では要素が 3 つのベクトルにオーバーラップされている。これに対し、オーバーラップを認めない手法もあるが、その手法では標本ベクトルの数が 40 個に減少する難点があり、採用しなかった。

可能性が大きくなり、不連続点が生じて、近いベクトル同士がつながる可能性が高まる。そのため、なめらかな時系列が得られる効果が期待できる。

6) マルコフ連鎖モデル：Manzan and Zerom(2008, pp.537-9)で提案しているモデルのうち、Remark 3に該当する。5)モデルでは、カーネル密度が高いところでは、同じベクトルが複数回連続して発生する可能性がある。たとえ同じベクトルが連続していても、株価に該当するレベル値でなく、変化率を示す対数階差につき同じベクトルが続いてもさほど不都合はないと考えられるが、6)モデルでは、この点を厳格に考慮しその可能性を排除する。出現値ベクトルを $E_1=(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ 、 $E_2=(\varepsilon_4, \varepsilon_5, \varepsilon_6), \dots$ と定義する。つぎに、ベクトル列 $Y_s=(y_{s+2}, y_{s+1}, y_s)'$ のうち、第1の要素にを含まない237個のベクトル列 $Y^{-s}_u=(y_{u+2}, y_{u+1}, y_u)'$ ($u=1, 2, \dots, 237$)を用意する。たとえば、 Y_1 を含まなければ、 $Y^{-1}_u=(y_{u+2}, y_{u+1}, y_u)'$ $= Y_2, Y_3, \dots, Y_{238}$ である。このような組合せは238通りある。そして、それぞれの Y^{-s}_u に対応するカーネル密度 pdf (y_u^{-s})を求める。はじめに $E_1=(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ を発生させるときは、 Y_1 が発生しないように、 Y^{-1}_u に対応する pdf (y_u^{-1})にそって、1回だけブートストラッピングを行い、その出現値を $E_1=Y_1'$ とする。第2回目、 Y_1 を含まない、 Y^{-1}_u に対応する pdf (y_u^{-1})にそって、再度1回だけブートストラッピングを行う。以降これを繰り返す⁹。この手法によれば、同じベクトルが連続しないとともに、高い確率で近いベクトルが次回出現する効果がある。以下の手順は5)モデルと同様である。ただし、これまでのモデルでは、10,000本分のサンプルパスをまとめて作りそれを切り離す手順で計算したが、6)モデルではサンプルパスの初期設定が必要なため、1本ずつサンプルパスを作成する。カーネル密度の計算は5)モデルと同様の設定で行った。

5 分析結果

4節の6種類のモデルに対し、1)モデルではRのrnorm関数、2)~6)モデルではsample関数を用い、5通りのランダムシードで10,000本のサンプルパスを発生させ、到達確率を求めた¹⁰。表5は、(2)式のt時点での到達確率を表すが、1)~6)のすべてのモデルで任意に選んだランダムシード201番の乱数列にもとづく分析結果を示している。ただし、表1の先行研究では、各期の運用期間は年単位であったが、本論文が採用する運用期間はより現実的な月単位である。このため、120カ月分の到達確率が得られたが、表示の見やすさから、年ごとの表1と対応させ、6年末(72ヵ月末)以降の応答年末出力値のみ掲げる。

9 Manzan and Zerom(2008)が提唱する手法は証券リターン1期 $\varepsilon_1, \varepsilon_2, \dots$ ずつ、確率密度を変更しブートストラッピングを行う手法であるが、3次元の自己相関の意味が希薄になると考え、ベクトル列ごとにマルコフ連鎖を用いた。

10 10,000本と数多くのサンプルパスを発生させているため、任意に選んだ5通りのランダムシード間のばらつきは、表5~6で掲げた平均変動係数で、高々8%の範囲内に収まっている。これは標本数を多くしたことにより、分布の漸近が進んでいることを示唆している。

表 5 目標積立額到達確率

	6 年末	7 年末	8 年末	9 年末	10 年末	年度平均 変動係数
1) 正規乱数モデル	0.013	0.051	0.107	0.185	0.268	0.044
2) カーネル密度モデル	0.015	0.054	0.120	0.199	0.278	0.031
3) ナイブブートストラップモデル	0.013	0.046	0.103	0.182	0.254	0.036
4) ブロックブートストラップモデル	0.016	0.057	0.121	0.192	0.274	0.029
5) カーネル密度BBモデル	0.009	0.058	0.160	0.301	0.443	0.032
6) マルコフ連鎖モデル	0.009	0.054	0.158	0.301	0.447	0.035

表 5 の数値は t 時点での目標積立額到達確率を示している。到達確率 (2) 式 $P(\bar{x}_t \geq \tau)$ に $T=120$ を代入し、 $P(\bar{x}_t \geq 120) = \#$ (到達したサンプルパス / 10,000) にもとづき計算している。 $N(0,0.2)$ の前提の表 1 ②欄と比較すれば、同じ正規性を前提とした 1) 正規乱数モデルでも到達確率は、マイナスの平均リターンパラメータを適用しているため、6 年末から $0.081 \rightarrow 0.013$, $0.179 \rightarrow 0.051$, $0.293 \rightarrow 0.107$, $0.401 \rightarrow 0.185$, $0.500 \rightarrow 0.268$ と一貫して減少している。つぎに、1) モデルと 2)~4) のモデルの間では年度ごとの分布にあまり差異はみられない。ところが、5), 6) モデルでは、1) モデルと比較し、9 年末で $0.185 \rightarrow 0.301$, 10 年末で $0.268 \rightarrow 0.443$, 0.447 と大幅に増加している。このため、誤識別の正規乱数モデルでは、成功確率を過小評価していることになるが、データの非正規性と非 IID 性を双方を考慮した 5), 6) モデルではじめて年度ごとの分布差異が顕著になり、ここにカーネル密度ブロックブートストラッピング法の意義が確認できる。加えて、5) と 6) モデルではあまり違いはなく、計算上負荷の少ない 5) モデルで十分であるといえる¹¹。相違が小さい点は、238 個と数多くの標本点からなるカーネル密度は、5), 6) モデル間であまり変わらないと考えられるからである。

つぎに、表 6 の数値ははじめて到達する確率を表す。ただし、シミュレーションごと 120 カ月分の出力値が得られているが、特定月にはじめて到達する確率値は微小なため、比較しにくい。そこで、表 1 ④欄との対応をとるため、特定年 12 カ月分の合計を掲げている。たとえば、1) 正規乱数モデル、6 年目、0.101 とあるのは、6 年 1 カ月 (61 カ月) 月末から 6 年 (72 カ月) 月末までの確率値を合計した数値である。そのため、表 1 と表 5 では「年末」と記載してあるが、表 6 では「年目」という表現を用いた。各月末の確率計算は、先行研究とは異なり、独立性の前提はおかないので、一般的な場合の条件付き確率の積からなる (3) 式が該当する。具体的な計算は、パスの個数をカウントすることで求めた。すなわち、(8) 式にしたがって、10,000 本のパスからはじめ、逐次的に 120 円に達したパスを以降のカウント対象から除いていき、 $t-1$ 期まで月次ごと一度も到達しない事象の条件付確率の累乗積を求め、 t 期の到達確率をかけ計算している。

11 4 節で説明した計算は、すべて標準的のパソコン (2GM, Windows Vista) で、ソフトウェア R を用いた。1 回のシミュレーションにかかった時間は、5) の方法では約 2 分であるのに対し、6) の方法では約 9 時間を要した。6) の方法では、R プログラミングで、行列処理に多くの if 文や for 文を用いざるを得なかったからと思われる。このため、計算コストの面からも、5) のカーネル密度ブロックブートストラッピング法は有益と考える。

$$P(\tilde{X}_1 \leq 120, \tilde{X}_2 \leq 120, \dots, \tilde{X}_{t-1} \leq 120, \tilde{X}_t \geq 120)$$

$$= \frac{\#運用開始から1期末まで未達のパス}{10,000} \times \frac{\#運用開始から2期末まで一度も到達していないパス}{\#運用開始から1期末まで未達のパス} \times \dots \times \frac{\#運用開始からt-1期末まで一度も到達していないパス}{\#運用開始からt-2期末まで一度も到達していないパス} \times \left(1 - \frac{\#運用開始からt期末まで一度も到達していないパス}{\#運用開始からt-1期末まで一度も到達していないパス} \right) \quad (8)$$

表 6 はじめて到達する確率

	6年目	7年目	8年目	9年目	10年目	平均到達年数	年度平均変動係数
1) 正規乱数モデル	0.101	0.381	0.398	0.108	0.005	7.060	0.065
2) カーネル密度モデル	0.132	0.400	0.365	0.078	0.003	6.979	0.061
3) ナイーブブートストラップモデル	0.086	0.350	0.416	0.135	0.009	7.160	0.081
4) ブロックブートストラップモデル	0.112	0.394	0.390	0.092	0.004	7.001	0.046
5) カーネル密度BBモデル	0.042	0.348	0.502	0.105	0.002	7.214	0.072
6) マルコフ連鎖モデル	0.038	0.328	0.520	0.111	0.002	7.249	0.050

1)モデルを比較すると表1④欄から変化が見られる。平均到達年数は8.156年から7.060年に短くなっている。確率密度に移ると、6年末から8年末までは、0.079 → 0.101, 0.161 → 0.381, 0.216 → 0.398と増加するが、9年末と10年末では、0.209 → 0.108, 0.156 → 0.005と減少している。つまり、はじめて到達する確率は7,8年目のピークが顕著になっており、この傾向は他のモデルでも共通してみられる。つぎに、1)モデルと他のモデルを比較すると、7,8年目の密度は概ね増加しているが、増加度合いは、8年目で5),6)モデルで0.398 → 0.502, 0.520とより顕著であり、分布の年度別コントラストがより顕著になっている。これらの観察事項から、積立投資開始時点において、8年末までに目標積立額に到達しなければ、以降到達することは難しいというインプリケーションが得られる。そして、8年末の高い水準の確率値はカーネル密度ブロックブートストラッピング法によってはじめて得られ、ここでもこの手法の意義が確認できる。

ところで、表6のモデルはすべて、はじめて到達する確率をすべての時点にわたって足すと、どのモデルでも0.99台の1に近い値になっている。ただし、かならず10年以内に目標積立額に到達することは意味しない。あくまで確率の計算であって、特定のパスを追跡して求めているわけではなく、パスによっては、運用期間中にいったん120円に到達しても、それ以降120円を割り込むこともありうるからである。そこで、分析の角度を変え、各モデルでサンプルパスが全運用期間120カ月にわたり、目標積立額120円を何回超えているかを調べる。表7は10,000本のサンプルパスの目標積立額到達回数の代表的統計値を示している¹²⁾。

12 最終欄の平均変動係数は、各モデルでの5つのランダムシードあるいは5回のシミュレーションで、平均、標準偏差、尖度、歪度、最大到達回数、到達回数ゼロのパス数の変動係数項目の平均値を表す。

表7 モデルごとのサンプルパス統計値

単位回数	1)モデル	2)モデル	3)モデル	4)モデル	5)モデル	6)モデル
平均	6.0	6.4	5.7	6.3	9.1	9.0
中央値	0	0	0	0	2	2
標準偏差	10.9	11.4	10.7	11.3	12.3	12.1
歪度	2.14	2.05	2.20	2.01	1.34	1.32
尖度	4.23	3.87	4.50	3.48	3.90	3.88
最大到達回数	66	74	65	66	68	71
到達回数ゼロのパス数	6,084	5,970	6,163	6,027	4,585	4,553
サンプルパス数	10,000	10,000	10,000	10,000	10,000	10,000
平均変動係数	0.023	0.034	0.026	0.029	0.029	0.035

1) モデルでは到達ゼロの件数が 6,084 件、平均 6.0 回、標準偏差 10.9 回という特徴をもつ。この傾向は 2), 4) のモデルではあまり変わらないが、3) ナイーブブートストラッピングモデルを用いる方法では、ゼロの件数が 6,163 件に増え、そのため平均と標準偏差の減少がみられるが、これは証券リターン ε_k のうち、表 4 に掲げた絶対値の大きいマイナス値を均等確率で出現するからと思われる。対照的に、よりなめらかさの加わったカーネル密度ブロックサンプリング法による 5) と 6) のモデルでは、ゼロの件数が 4,585, 4,553 件と大幅に減少、そのため平均と標準偏差の増加と中央値の 0 → 2 への増加がみられる。これは、両者とも絶対値の大きなマイナス値を含むベクトルが出現する確率密度を抑えているためと推察される。ただし、今回の対象データを重視する立場にたつ分析では 5), 6) モデルでは到達回数ゼロの件数が大幅に減少しているが、それは、基本的にはもともになるデータの分布形状と自己相関に依存すると考えられ、どのような場合でも減少するわけではない。

6 真の分布・時系列構造を想定した比較

5 節では、目標積立額到達確率、はじめて到達する確率ともに、5) カーネル密度ブロックブートストラッピングモデルと、6) マルコフ連鎖モデルの結果が、1) 正規乱数モデルの結果とは顕著に異なるパターンを示すことを確認した。この点は本論文がいちばんに主張する点である。しかしながら、5), 6) モデルの示す確率値が正しい値か、またどのような条件下で有効といえるのか、さらに検討すべき課題が残されている。そこで、6 節では、非正規の真の分布と非 IID の真の時系列構造を想定し、出力する確率値が 5), 6) モデルの出力値と近いかを調べる。ただし、真の分布や時系列構造は永遠に未知なので、3 節で検討した TOPIX 過去データに近い理論分布と時系列構造を推定し、それを真の分布・時系列構造と想定する。そして、5 節同様、10,000 本のパスを発生させ確率値を求めるとともに、out of sample を使って求めた 1) ~ 6) のモデルの出力結果と比較するアプローチをとる。

まず、3 節で述べた TOPIX 対数階差データセット表 2 ⑧を対象に、ソフトウェア Cristal Ball® を用いて、いちばん適合的な理論分布を検索した結果、3 種類の検定方法のうち 2 種類で (1 種類は次点として)、10 種類の理論分布候補からロジスティック分布が導出された。そこで、真の分布には、パラメータ (平均 = -0.004205, 尺度 = 0.0320454 (= $\sqrt{3 \times \text{分散} / \pi^2}$)) のロジスティック分布を選択した。3 節で議論したように、データセット ⑧ の非正規性は歪度と尖度に顕著に見られる。このうち、ロジスティック分布では、正規分布と同様に歪度 = 0 なので、分布の非対称性の特性は反映できない。しかし、尖度 = 4.2 (正規分布 3) のファットテール特性で非正規性を反映させることができる。

次に、データセット⑧の時系列特性に移る。3 節で述べたように、コレログラムでは、自己相関は認められない。仮に認められても、自己相関係数、偏自己相関係数ともに絶対値が高々 0.09 台で、長期のラグにわたり弱い依存関係しか検知されず、顕著な AR または MA のパラメトリック構造は特定できない。そこで、データ特性からは乖離するが、簡単のため AR(1), $\rho = 0.1$ をもって真の時系列構造と想定した。

以上の真の分布・時系列構造にもとづき、4, 5 節同様、120 ヶ月分を 1 本のとするパスを 10,000 本発生させたところ、5 種類のランダムシード間で、目標積立額達成確率値、はじめて到達する確率値の変動係数は表 5, 6 に対応するレベル未満であった。このため以降の分析は、任意に選んだランダムシード 51 番の乱数列を用いて行った。比較対象の 1)~6) モデルでは、データセット⑧を用いずに、51 番の乱数列が出力した 10,001 本目からの out of sample データのうち 240 ヶ月分の 2 本をサンプルとして用いている¹³。

表 8 目標積立額到達確率（真の分布との比較）

	6 年末	7 年末	8 年末	9 年末	10 年末
ロジスティック AR(1)モデル $\rho = 0.1$	0.022	0.062	0.118	0.186	0.260
1) 正規乱数モデル	0.089	0.213	0.357	0.503	0.612
2) カーネル密度モデル	0.006	0.049	0.169	0.345	0.517
3) ナイーブブートストラップモデル	0.091	0.215	0.365	0.502	0.618
4) ブロックブートストラップモデル	0.075	0.199	0.357	0.506	0.632
5) カーネル密度 BB モデル	0.006	0.054	0.180	0.369	0.552
6) マルコフ連鎖モデル	0.006	0.053	0.186	0.374	0.558

表 8 の数値は表 5 と対応し、t 時点での目標積立額早期到達確率を示している。真の分布は、6 年末から徐々に確率値が増加し、10 年末は 0.260 程度のレベルである。これに対し、1)~6) のモデルはどれも 9, 10 年末のレベルが高く、10 年末では 0.5 を超えている。絶対的な意味では、どのモデルも真の分布に近いとはいえない。しかしながら、5), 6) モデルでは、8 年末の数値が真の分布に近いこと、9, 10 年末の数値では、他のモデルで極端に離れた数値を真の分布・時系列構造の数値の方向に戻している。このため、表 8 の範囲内では、5), 6) モデルは相対的に真の分布・時系列構造に近い確率値を出力していると評価する。

13 データセット⑧を用いれば、表 5, 6 と同様な結果しか得られず意味がない。また、10,000 本の in sample は想定した真のデータでは、1)~6) モデルで確率値を推定する意味合いが希薄になるので、out of sample をもってサンプルとした。24 ヶ月分を用いたのはデータセット⑧の個数と対応させるためである。なお、表 8, 9 のサンプルは 10,001 本目と 10,002 本目のサンプルをとっている。4)~6) で用いるブロックサイズは、真の時系列構造でラグ 1 を想定したので、4 節の 3 次元ベクトルではなく、2 次元ベクトルを用いている。

表 9 はじめて到達する確率（真の分布との比較）

	6 年目	7 年目	8 年目	9 年目	10 年目	平均到達年数
ロジスティックAR(1)モデル $\rho=0.1$	0.186	0.404	0.316	0.066	0.003	6.763
1) 正規乱数モデル	0.476	0.397	0.031	0.000	0.000	5.899
2) カーネル密度モデル	0.030	0.295	0.554	0.120	0.001	7.309
3) ナイブブートストラップモデル	0.499	0.363	0.027	0.000	0.000	5.848
4) ブロックブートストラップモデル	0.432	0.446	0.045	0.000	0.000	5.998
5) カーネル密度BBモデル	0.030	0.320	0.550	0.099	0.001	7.265
6) マルコフ連鎖モデル	0.035	0.326	0.544	0.094	0.001	7.237

つぎに、表 9 の数値は表 6 と対応し、はじめて到達する確率を示している。真の分布・時系列構造は平均到達年数 6.763 年で、7 年目にピーク値 0.404 を示し、9 年目と 10 年目はほとんどゼロというパターンを示している。これに対し、1)~6) のモデルは高々 0.1 台と、9、10 年目は適格的であるが、6 年目の 2) カーネル密度モデル、8 年目の 1) 正規乱数モデルで 0.03 台の過小誤識別が散見される。また、5)~6) モデルでも、6 年目が 0.05 未満と過小、8 年目が 0.5 台と過大識別を起こしている。この意味では、表 8 と同様、絶対的に真の分布に近いとはいえない。ただし、ここでも表 8 と同様、たとえば 7 年目に、2) モデルが過小、4) ブロックブートストラッピングモデルが過大識別を起こしているが、5)~6) モデルはそれらを修正するように相対的に真の分布・時系列構造に近い確率値を出力していると考ええる。

6 節の結論として、調べた範囲内では、5)、6) モデルは、絶対的な意味で真の分布・時系列構造の確率値を表してはいない。しかし、5 節で検討した 5)、6) モデルが間違っているともいえない¹⁴。両者は非正規性、非 IID 性を織り込んでいるので、他のモデルとの比較で、相対的な意味で真の分布・時系列構造に近い確率値を出力していると考ええる。

7 結 論

資産形成層にとって、積立証券投資開始時における目標積立額早期到達確率は有益な数値情報といえる。高見 (2009) は、年度ごとの到達確率とはじめて到達する確率を数理的に導出した。ただし、証券リターンを示す確率変数の時系列に、正規性かつ IID 性という厳格な前提をおき、年 1 回の積立サイ

14 カーネル密度ブロックブートストラッピング法に限定し、同じサイズの他の 4 通りの out of sample で調べたところ、特に はじめて達成する確率値でパターンは大きく変化することが認められた。また、サンプルサイズの R プログラム実行上の上限 26 本分まで増やしても、真の分布の確率値に収束する傾向はみられない。この点は out of sample に依存するところが大きく、サンプルに異常な値が含まれると、ノンパラメトリック法はその特徴を増幅して表現するものと考ええる。加えて、相関係数を 0.4 に上げて調べたところ、パターンの変動は大きくなる。この点はブロックごとのつなぎ目で独立に近い挙動が見られるからではないかと考える。これら諸点から、6 節の真の分布・時系列構造の前提は強すぎ、5 節での 5)、6) モデルの正確性を否定あるいは肯定する材料にはならない。どのような設定で 5 節での 5)、6) モデルの正確性を検証すべきかについてはいまだ研究は未着手であり、今後の課題として取り扱う。

クルに単純化した点に問題を残していた。一方、TOPIX 対数階差の過去データは、非正規、非 IID 性を示す場合が多く、年次と月次間でスケール則は働かない。そのため、過去データを重視する立場で将来の予測をとらえれば、数理解析には限界がある。そこで、本論文はシミュレーションにより、目標積立額到達確率を求めることに取り組んだ。用いた手法は、非正規性にはカーネル密度法、ブートストラッピング法を、非 IID 性にはブロックブートストラッピング法を、双方を統合してカーネル密度ブロックブートストラッピング法を適用した。そして、6 種類のモデルで、目標積立額到達確率、はじめて到達する確率ならびにサンプルパスの態様を比較した。

得られた結果は、誤識別の正規乱数モデルでは、到達確率を過小評価していること、はじめて到達する確率で、過去データを重視した立場では 8 年末により密度のピークがあると推論できること、真の姿に絶対的にどれだけ近いかにについては保証できないものの、相対的な意味において近づくには、非正規性と非 IID 性の双方を考慮する必要があること、一方、双方を考慮したマルコフ連鎖モデルとカーネル密度ブロックブートストラッピングモデル間にはあまり差異がなく簡便法の 5) モデルで十分であり、計算コストの観点から有益であること、などである。これらは、総じてカーネル密度ブロックブートストラッピング法の意義を示している。

本論文が扱ったデータは TOPIX の 1 区間にすぎないこと、ダウサイドリスクイベントも考察すべきこと、真の分布・時系列構造とは絶対的な意味でどれほど乖離があるか検証できていないこと、など本論文は課題を残す。今後は、これら課題に対処するとともに、ボーナス積立併用の場合、安全資産をポートフォリオに組み入れた場合などのより具体的な積立証券投資に拡張して行きたい。しかし、本論文の範囲における貢献は、非正規性と非 IID 性の双方の時系列特性をカーネル密度ブロックブートストラッピング法により確認したこと、ならびにモデル間で比較検討を行い、非正規性と非 IID 性の双方の手法を適用してはじめて、到達確率パターンに変化が現れることを示したことにある。

【謝辞】

本論文は日本経営財務研究学会第 33 回全国大会(山口大学)での報告を元にしてしています。まず、全国大会にてコメンテーターをつとめられた佐々木一郎先生(同志社大学)に感謝申し上げます。また、論文審査過程にて、2 名のレフェリー先生方より有意義なコメントを頂き、誤りの修正と研究の拡張につながり、論文の質は確実に向上しました。とくに、第 6 節はコメントを受け新たに付け加えた箇所があります。ここに感謝の念を表明いたします。なお、残された誤謬については筆者に責任の所在があります。

【参考文献】

- [1] Chernick, Michael R. (2008) "Bootstrap Methods: a Guide for Practitioners and Researchers, 2nd Edition," John Wiley and Sons.
- [2] Chhabra, Ashvin B. (2005) "Beyond Markowitz: A Comprehensive Wealth Allocation Framework for Individual Investors," *Journal of Wealth Management*, Vol.7, No.4, 8-34.
- [3] Diks, Cees and S. Manzan (2002) "Tests for Serial Independence and Linearity Based on Correlation Integral," *Studies in Nonlinear Dynamics and Econometrics*, Vol.6, No.2 (available at <http://ssrn.com/abstract=1269053>).

- [4] Dubil, Robert (2005) "Lifetime Dollar-Cost Averaging: Forget Cost Savings, Think Risk Reduction," *Journal of Financial Planning*, Vol. 18, No.10, 86-90.
- [5] Duong, Tarn (2007) "ks: Kernel Density Estimation and Kernel Discriminant Analysis for Multivariate Data in R," *Journal of Statistical Software*, Vol.21, No.7,1-16.
- [6] Härdle, Wolfgang, J. Horowitz and J-P. Kreiss (2003) "Bootstrap Methods for Time Series," *International Statistical Review*, Vol.71, No.2, 435-459.
- [7] Horowitz, Joel L. (2003) "Bootstrap Methods for Markov Processes," *Econometrica*, Vol.71, No.4, 1049-1082.
- [8] Li, Qi and J. S Racine (2007) "Nonparametric Econometrics," Princeton University Press.
- [9] MacKinnon, James G. (2006) "Bootstrap Methods in Econometrics," *Economic Record*, Vol.82, S2-S18.
- [10] Manzan, Sebastiano and D. Zerom (2008) "A Bootstrap-Based Non-Parametric Forecast Density," *International Journal of Forecasting*, Vol.24, 535-550.
- [11] Mukherji, Sandip (2008) "A Study of Time Diversification with Block Bootstraps and Downside Risk," *Business Review*, Vol. 10, No.1, 55-60.
- [12] Paparoditis, Efstathios and D. N. Politis (2001) "A Markovian Local Resampling Scheme for Nonparametric Estimators in Time Series Analysis," *Economic Theory*, Vol. 17, 540-566.
- [13] Politis, Dimitris N. (2003) "The Impact of Bootstrap Methods on Time Series Analysis," *Statistical Science*, Vol.18, No.2, 219-230.
- [14] Trainor Jr, William (2005) "Within-horizon Exposure to Loss for Dollar Cost Averaging and Lump Sum Investing," *Financial Services Review*, Vol.14, No.4, 319-330.
- [15] 今野義彦・森康久仁・松葉育雄 (2007) 「ブートストラッピング法による少数データのリアプノフ指数精度向上法と経済時系列への応用」『信学技法 NLP2006-146 電子情報通信学会』, 21-26.
- [16] 高橋希代子 (2007) 「資産形成層への運用アドバイスはこんな考えで進めよう」『近代セールス』2007年12月1日号, 12-15.
- [17] 高見茂雄 (2009) 「ドルコスト平均法投資におけるリスク・リターン把握—資産形成層の立場から」『日本ファイナンス学会第 17 回大会予稿集』, 455-464.