

フィナンシャルエンジニアリング特論第2

中間レポート

内海佑麻^{*}
(学籍番号: 82018398)

澤屋敷友一[†]
(学籍番号: 82019220)

November 27, 2020

概要

ポートフォリオ選択モデル (1 期間モデル) を構築し, それを用いて資産運用を行う金融商品を組成し, そのパフォーマンスを検証した.

目次

1	モデルの構築方法	2
1.1	ポートフォリオ選択モデル	2
1.2	ソフトウェアとデータ	3
1.3	バックテストの設定	4
2	バックテストとパフォーマンス評価	5
2.1	運用期間における東京証券取引所の動向	5
2.2	ポートフォリオ選択モデルのパフォーマンス評価	6
3	結論と考察	6

^{*} 所属: 環境開放科学専攻 情報工学専修 連絡先: uchiumi@ailab.ics.keio.ac.jp

[†] 所属: 環境開放科学専攻 オープンシステムマネジメント専修 連絡先: yashiki@keio.jp

1 モデルの構築方法

1.1 ポートフォリオ選択モデル

1.1.1 Markowitz の平均分散モデル

Markowitz の平均分散モデルでは、「ポートフォリオの期待収益率 (Expected return) が一定値以上となる」という制約条件の下で、「ポートフォリオの分散を最小化する」最適化問題を考える。一般に、 n コの資産で構成されるポートフォリオの場合、ポートフォリオの分散は n コの資産間の共分散行列の二次形式となるので、この最適化問題は二次計画問題 (Quadratic Programming, QP) のクラスとなり、次のように定式化される。

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} \quad \sigma_p (= \mathbf{x}^T \Sigma \mathbf{x}) \quad (1)$$

$$\text{subject to} \quad \bar{r}_p = \bar{\mathbf{r}}^T \mathbf{x} = \sum_{i=1}^n \bar{r}_i x_i \geq r_e \quad (2)$$

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n x_i = 1 \quad (3)$$

$$x_i \geq 0 \quad (i = 1, \dots, n) \quad (4)$$

- $\Sigma \in \mathbb{R}^{n \times n}$ - n コの資産の共分散行列
- $\mathbf{x} \in \mathbb{R}^n$ - n コの資産の投資比率ベクトル
- $\bar{\mathbf{r}} \in \mathbb{R}^n$ - n コの資産の期待収益率ベクトル
- $x_i \in \mathbb{R}$ - 資産 i の投資比率
- $\bar{r}_i \in \mathbb{R}$ - 資産 i の期待収益率
- $r_e \in \mathbb{R}$ - 投資家の要求期待収益率
- $\bar{r}_p \in \mathbb{R}$ - ポートフォリオの期待収益率
- $\sigma_p \in \mathbb{R}$ - ポートフォリオの標準偏差

1 つ目の制約式は、ポートフォリオの期待収益率が一定値 ($= r_e$) 以上となることを要請している。2 つ目、3 つ目の制約式はポートフォリオの定義からくる自明なものである。資産の空売りを許す場合、3 つ目の制約式を除くこともある。

1.1.2 Sharpe-ratio 最大化モデル

シャープレシオ (Sharpe ratio, SR) は、最もよく使われるポートフォリオに対するリスク調整済みパフォーマンス尺度である。任意のポートフォリオ p のシャープレシオ θ_p は、無リスク資産の収益率 r_f とポートフォリオの期待収益率 \bar{r}_p 、ポートフォリオの標準偏差 σ_p を用いて

$$\theta_p := \frac{\bar{r}_p - r_f}{\sigma_p} \quad (5)$$

と定義される。シャープレシオ最大化問題は、目的関数に n コの資産間の共分散行列が含まれるため、二次計画問題 (Quadratic Programming, QP) のクラスとなり、次のように定式化される。

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad \frac{\bar{r}_p - r_f}{\sigma_p} \left(= \frac{\bar{\mathbf{r}}^T \mathbf{x} - r_f}{\sqrt{\mathbf{x}^T \Sigma \mathbf{x}}} \right) \quad (6)$$

$$\text{subject to} \quad \|\mathbf{x}\|_1 = \sum_{i=1}^n x_i = 1 \quad (7)$$

$$x_i \geq 0 \quad (i = 1, \dots, n) \quad (8)$$

- $r_f \in \mathbb{R}$ - 無リスク資産の収益率

上式の目的関数を二次形式で表すため、ポートフォリオの期待リスクプレミアム λ と各資産の比重ベクトル $\mathbf{w} = \mathbf{x}/\lambda$ を導入して変形すると、次のように Sharpe-ratio 最大化問題のコンパクト分解表現を得る。

$$\underset{\mathbf{w} \in \mathcal{W}}{\text{minimize}} \quad \mathbf{w}^T \Sigma \mathbf{w} \quad (9)$$

$$\text{subject to} \quad \|\hat{\mathbf{r}}^T \mathbf{w}\|_1 = \sum_{i=1}^n (\bar{r}_i - r_f) w_i = 1 \quad (10)$$

$$w_i \geq 0 \quad (i = 1, \dots, n) \quad (11)$$

- $x_i \in \mathbb{R}$ - 資産 i の投資比率
- $w_i \in \mathbb{R}$ - 資産 i の投資比率と期待リスクプレミアムの比率 ($w_i = x_i/\lambda$)
- $\lambda \in \mathbb{R}$ - ポートフォリオの期待リスクプレミアム ($\lambda = \sum_{i=1}^n \hat{r}_i x_i = \sum_{i=1}^n (\bar{r}_i - r_f) x_i$)
- $\hat{r}_i \in \mathbb{R}$ - 資産 i の期待リスクプレミアム ($\hat{r}_i = \bar{r}_i - r_f$)

ここで、定義式 $x_i = \lambda w_i$ と予算制約式 $\sum_{i=1}^n x_i = 1$ に注意すると、 $\lambda = 1/(\sum_{i=1}^n w_i)$ となるから、元問題の実行可能解 \mathbf{x} とコンパクト分解表現の実行可能解 \mathbf{w} の間に以下のような 1 対 1 対応が成り立つ。

$$\mathbf{x} = \lambda \mathbf{w} = \frac{\mathbf{w}}{\sum_{i=1}^n w_i} = \frac{\mathbf{w}}{\|\mathbf{w}\|_1} \quad (12)$$

1.2 ソフトウェアとデータ

Python を用いてポートフォリオ選択モデルの実装とヒストリカルデータに対するバックテストを行った。今回作成した Python スクリプトではすでに公開されているパッケージを使っていない。主な理由は、資産運用やシステムトレード向けの Python パッケージはいくつかあるが、米国の株式市場を対象としたものが多く、日本の株式市場に対応しているものは少ないためである。^{*1} また、作成したスクリプトや実行ファイルは Github 上にリポジトリ (https://github.com/yumaloop/afe2_backtest) として公開した。

1.2.1 pandas-datareader パッケージによるデータ取得

株式銘柄のヒストリカルデータは、pandas-datareader を用いて取得した。pandas-datareader は、IEX, World Bank, OECD, Yahoo! Finance, FRED, Stooq などが公開している Web API を内部で呼び出すことで、Python スクリプト上に取得したいデータを読み込みことができる。

^{*1} "The Top 22 Python Trading Tools for 2020", <https://analyzingalpha.com/python-trading-tools>

1.2.2 cvxopt パッケージによる二次計画問題の球解

銘柄選択時に必要となる凸最適化問題の球解には, cvxopt (<https://cvxopt.org>) を使った. cvxopt は二次計画問題 (QP) を含む一般の凸最適化問題に対する高速ソルバである. cvxopt で二次計画問題を扱う場合は, 解きたい最適化問題を以下の一般形式に整理して,

$$\underset{\mathbf{x}}{\text{minimize}} \quad \frac{1}{2} \mathbf{x}^T P \mathbf{x} + \mathbf{q}^T \mathbf{x} \quad (13)$$

$$\text{subject to} \quad G \mathbf{x} \leq \mathbf{h} \quad (14)$$

$$A \mathbf{x} = \mathbf{b} \quad (15)$$

パラメータ P, q, G, h, A を計算し, cvxopt.solvers.qp() 関数を実行することで最適解と最適値を求める. たとえば, Markowitz の平均・分散モデルの場合は,

$$P = 2 \cdot \Sigma, \quad q = \mathbf{0}_n, \quad G = -1 \cdot \begin{pmatrix} \bar{r}_1 & \cdots & \bar{r}_n \\ 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}, \quad h = -1 \cdot \begin{pmatrix} r_e \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad A = \mathbf{1}_n^T, \quad b = 1 \quad (16)$$

となり*2, シャープレシオ最大化モデルの場合は次のようになる.

$$P = 2 \cdot \Sigma, \quad q = \mathbf{0}_n, \quad G = -1 \cdot \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}, \quad h = \mathbf{0}_n, \quad A = \begin{pmatrix} \bar{r}_1 - r_f \\ \vdots \\ \bar{r}_n - r_f \end{pmatrix}^T, \quad b = 1 \quad (17)$$

1.3 バックテストの設定

以下のような設定下で, ポートフォリオ選択モデルを過去の価格データへ適用し, バックテストを行った.

- 運用期間 - 2011 年 10 月 31 日～2020 年 10 月 31 日 (過去 10 年間)*3
- 対象資産 - 運用開始時点 (2011 年 10 月 31 日) で, TOPIX 500 (Core30) に採用されている内国株式
- 無リスク資産の利回り - 0.01% 日本国債 10 年物利回り*4
- リバランス単位 - 1 カ月ごとにポートフォリオを再調整する.
- パラメータの推定期間 - リバランス時点から過去 36 カ月
- パフォーマンス評価 - 運用期間におけるポートフォリオのシャープレシオ (年率)

*2 cvxopt の公式ドキュメント参照. <https://cvxopt.org/userguide/coneprog.html#quadratic-programming>

*3 . TOPIX シリーズの構成銘柄は毎年 10/31 に更新されるため 10/31 を基準日とした.

*4 出典: Bloomberg LP, <https://www.bloomberg.co.jp/markets/rates-bonds/government-bonds/japan>

2 バックテストとパフォーマンス評価

2.1 運用期間における東京証券取引所の動向

東証 TOPIX シリーズの構成銘柄に対して収益率および累積収益率の変動をみると、2007-2009 の大きな下降、2012-2013 の大きな上昇を除くと緩やかに増加していることがわかる。

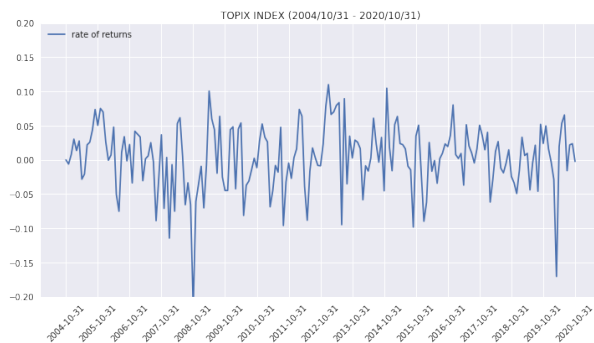


図1 TOPIX Index の収益率

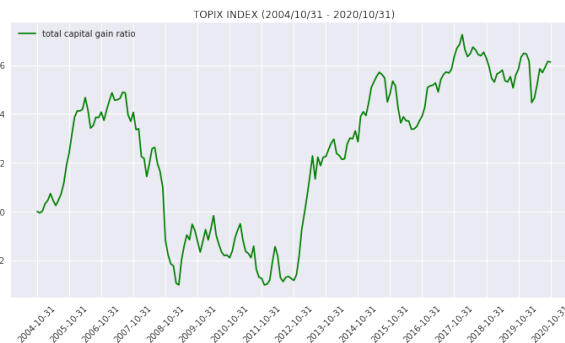


図2 TOPIX Index の累積収益率

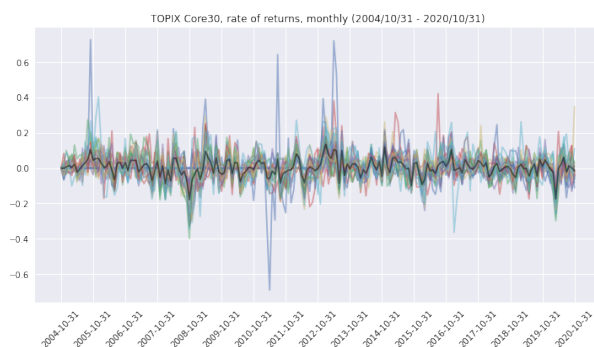


図3 TOPIX Core30 構成銘柄の収益率

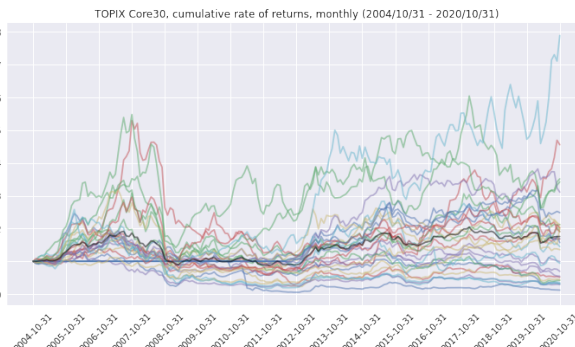


図4 TOPIX Core30 構成銘柄の累積収益率

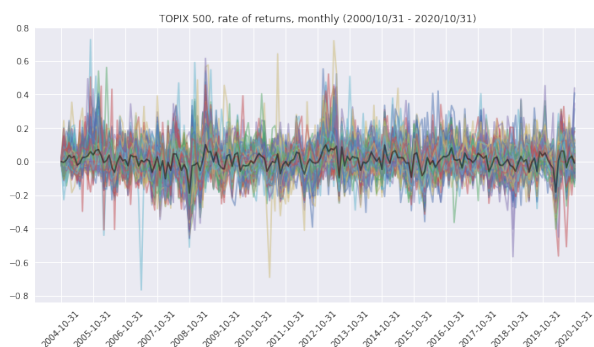


図5 TOPIX 500 構成銘柄の収益率

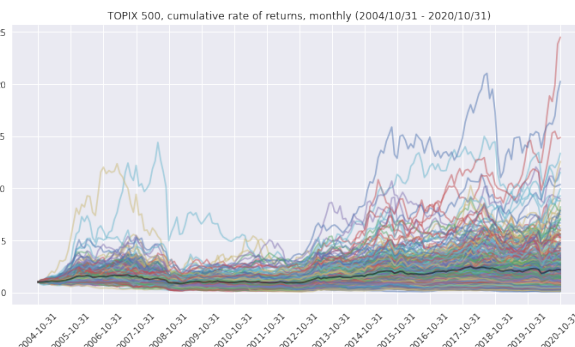


図6 TOPIX 500 構成銘柄の累積収益率

実際、収益率のヒストグラムをみると、平均が正となる単峰性の形状を示しており、「収益率は正規分布に従う確率変数である」と仮定することが妥当であると判断できる。すなわち、平均・分散をリターン・リスクの計算に用いることが正当化される。

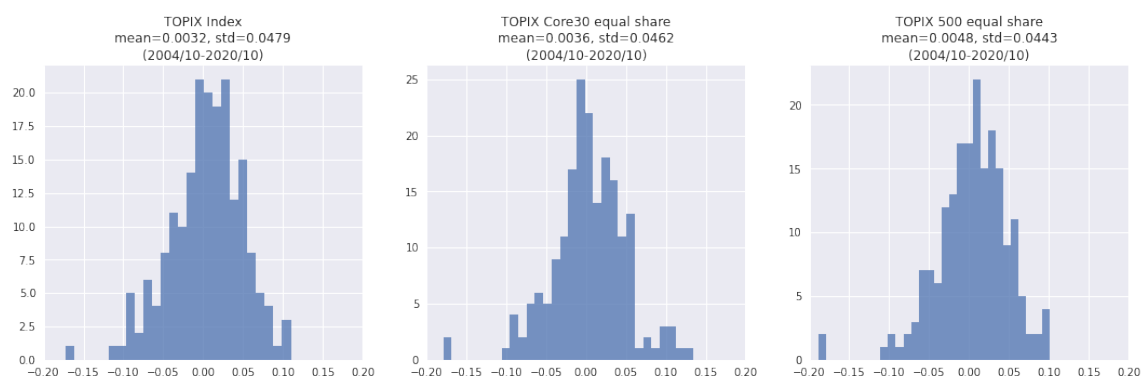


図7 TOPIX シリーズの月次収益率分布 (期間:2004/10 - 2020/10)

- (1) TOPIX Index(東証全銘柄に対する加重平均)
- (2) TOPIX Core30 構成銘柄に対する等配分ポートフォリオ
- (3) TOPIX 500 構成銘柄に対する等配分ポートフォリオ

2.2 ポートフォリオ選択モデルのパフォーマンス評価

表1 バックテストの結果: 運用期間:2011/10 - 2020/10 (年率)

ポートフォリオ選択モデル (対象銘柄, パラメータ推定期間)	シャープレシオ θ_p	リターン \bar{r}_p	リスク σ_p
平均分散モデル (TPX Core30, 過去 12 カ月)	0.5137	0.0829	0.1612
平均分散モデル (TPX Core30, 過去 36 カ月)	0.6201	0.0847	0.1364
平均分散モデル (TPX Core30, 過去 60 カ月)	0.6172	0.0854	0.1382
シャープレシオ最大化モデル (TPX Core30, 過去 12 カ月)	0.6181	0.1121	0.1812
シャープレシオ最大化モデル (TPX Core30, 過去 36 カ月)	0.7335	0.1245	0.1697
シャープレシオ最大化モデル (TPX Core30, 過去 60 カ月)	0.6604	0.1201	0.1817

3 結論と考察

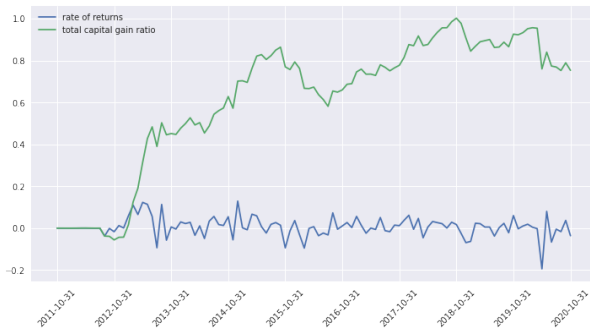


図 8 平均分散モデル [$\theta_p = 0.5137$]
(TOPIX Core30 構成銘柄, 推定期間:過去 12 カ月)

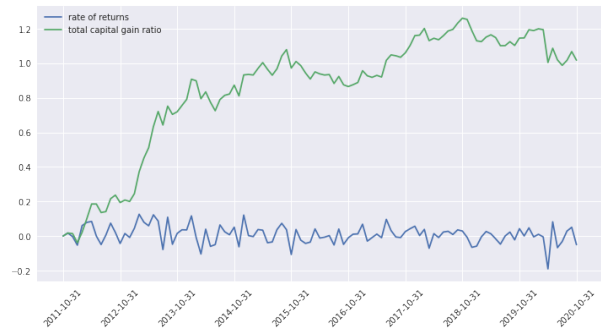


図 9 シャープレシオ最大化モデル [$\theta_p = 0.6181$]
(TOPIX Core30 構成銘柄, 推定期間:過去 12 カ月)

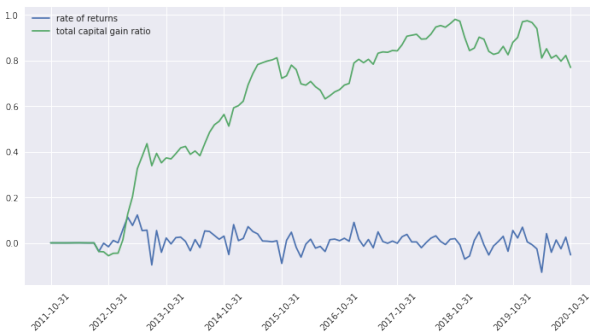


図 10 平均分散モデル [$\theta_p = 0.6201$]
(TOPIX Core30 構成銘柄, 推定期間:過去 36 カ月)

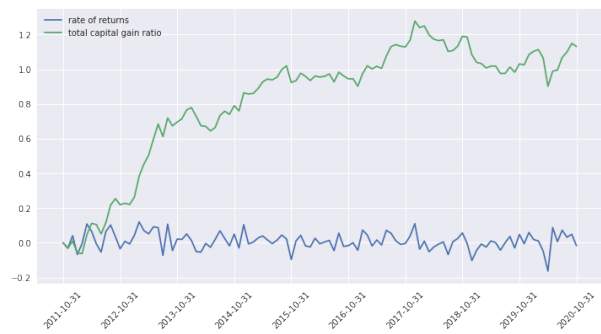


図 11 シャープレシオ最大化モデル [$\theta_p = 0.7335$]
(TOPIX Core30 構成銘柄, 推定期間:過去 36 カ月)

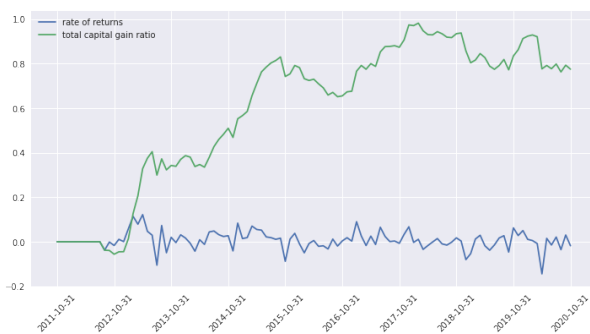


図 12 平均分散モデル [$\theta_p = 0.6172$]
(TOPIX Core30 構成銘柄, 推定期間:過去 60 カ月)

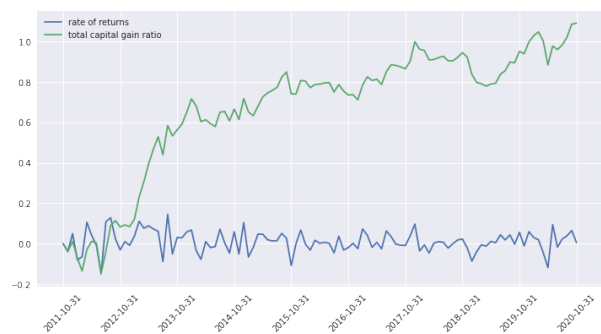


図 13 シャープレシオ最大化モデル [$\theta_p = 0.6604$]
(TOPIX Core30 構成銘柄, 推定期間:過去 60 カ月)

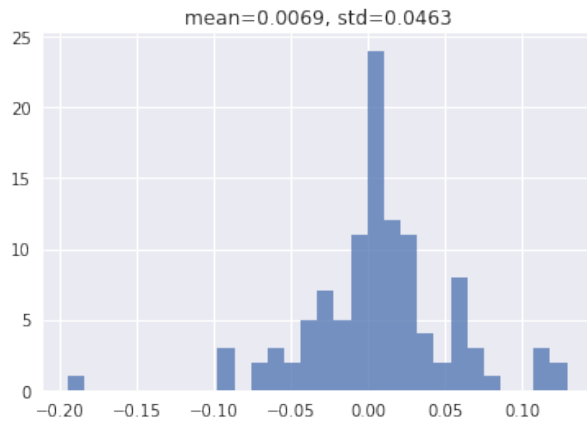


図 14 平均分散モデル [$\theta_p = 0.5137$]
(TOPIX Core30 構成銘柄, 推定期間:過去 12 カ月)

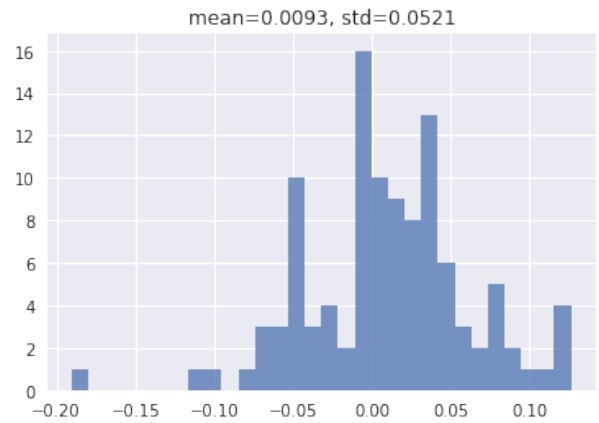


図 15 シャープレシオ最大化モデル [$\theta_p = 0.6181$]
(TOPIX Core30 構成銘柄, 推定期間:過去 12 カ月)

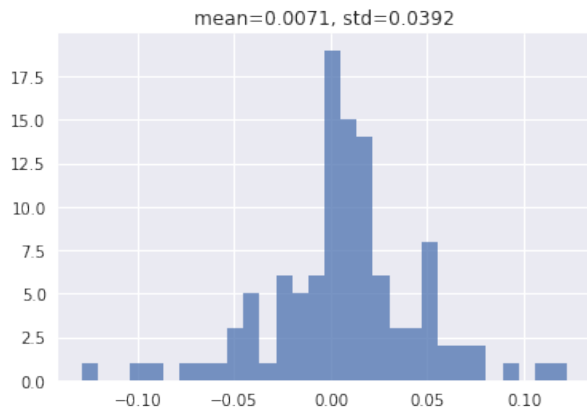


図 16 平均分散モデル [$\theta_p = 0.6201$]
(TOPIX Core30 構成銘柄, 推定期間:過去 36 カ月)

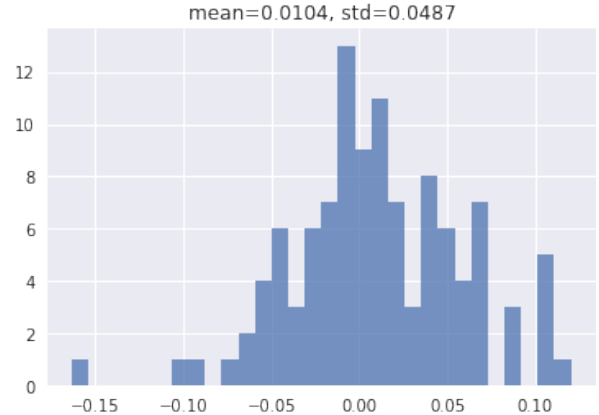


図 17 シャープレシオ最大化モデル [$\theta_p = 0.7335$]
(TOPIX Core30 構成銘柄, 推定期間:過去 36 カ月)

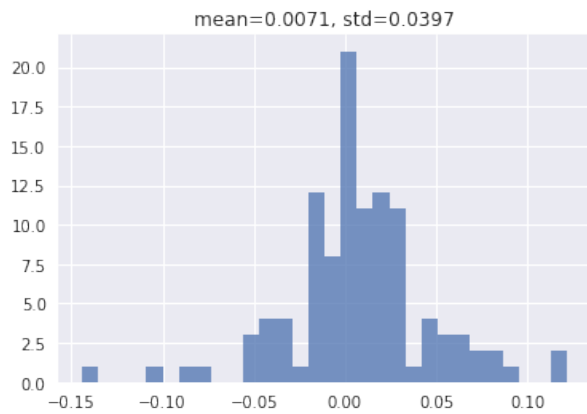


図 18 平均分散モデル [$\theta_p = 0.6172$]
(TOPIX Core30 構成銘柄, 推定期間:過去 60 カ月)

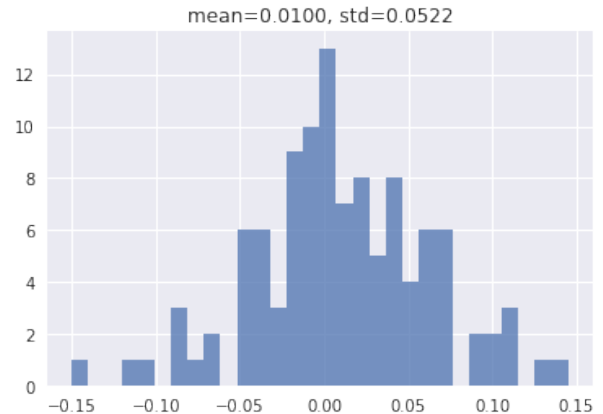


図 19 シャープレシオ最大化モデル [$\theta_p = 0.6604$]
(TOPIX Core30 構成銘柄, 推定期間:過去 60 カ月)