ブートストラップ法による安定分布のパラメータ推定

安藝優子　(指導教員：吉田裕亮)

1　はじめに

安定分布

まず本研究で扱う確率分布である, 安定分布について説明する. これは特徴として, 正規分布より裾が重いこと, 株価や為替レート変動など金融資産の収益率などで表れる分布であり, フラクタル幾何学の創始者である Mandelbrot が株式市場においても見出したと言われている.

この確率密度関数は陽には定義されないが, 分布の特性関数 (Fourier 変換) が下の式のように表される.

$$\exp[i\delta z - \gamma|z|^{\alpha}(1 + i\beta\,\mathrm{sgn}(z) w(z,\alpha))]$$

$$w(z,\alpha) = \begin{cases} \tan(\frac{\pi\alpha}{2}), & \alpha \neq 1 \\ \frac{2}{\pi}\log(|z|), & \alpha = 1 \end{cases}$$

パラメータ α $(0 < \alpha \leq 2)$ は分布の裾状態を表す尺度であり, 値が小さいほど裾が重い. β $(-1 \leq \beta \leq 1)$ は非対称パラメータ, γ $(0 < \gamma)$ は規模のパラメータ, δ は平行移動のパラメータである. 本研究では平行移動・拡大縮小は考えず, 左右対称のものを扱う. このとき, 分布の Fourier 変換は

$$\widehat{f}(x) = \exp(-c|x|^{\alpha}), (c\ \text{は正規化定数})$$

のように表されることになる.

また $\alpha = 2$ のとき正規分布, $\alpha = 1$ のときコーシー分布である. Mandelbrot は金融市場の価格変動の一例として綿花の価格は $\alpha = 1.7$ の安定分布に従うことを調べた.

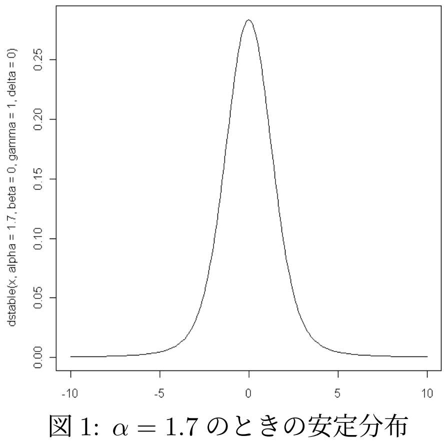


図 1: $\alpha = 1.7$ のときの安定分布

一般的な確率分布のパラメータ推定方法においては最尤法がよく用いられる. 最尤法とは確率密度関数に基づき尤度関数を求め, 尤度を最大にするようにパラメータを推定する手法である. この手法は確率密度関数が簡単な式で定義される分布に適している. しかし, 安定分布は先にも述べたように確率密度関数が単純な関数で与えられないため最尤法で推定することは, 非常に困難である. そこで最尤法に代わる効率的な推定方法が現実的な場面では必要とされている.

2　ブートストラップ法

標本集団より重複を許したリサンプリングを多数繰り返し, 得られた新たな標本集団より母集団の性質やモデルの推測の誤差などを推定する手法をブートストラップ法と呼ばれ,1979 年に Eflorn によって提唱された. 具体的には以下のような手順で行われる.

ある母集団の標本, $\{x_1, x_2, x_3, \cdots, x_n\}$ があるとする. この標本から, ランダムに重複も含めてリサンプリングを行い, リサンプリングした標本を $\{x_1^*, x_2^*, x_3^*, \cdots, x_n^*\}$ とする. このリサンプルからある推定値, 例えば標本平均値または標本中央値などの順序統計量を計算し, それを T_1 とおく. また新たなリサンプリングを行い, 同様に推定値を計算し, それを T_2 とおく. このプロセスを N 回繰り返すとすると, N 個の推定値, T_1 から T_N が得られる. この N 個の推定値から, ある推定量の標準誤差 (SD) 及び推定平均 $\overline{T}$,

$$SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(T_(i) - \overline{T})^2}, \quad \overline{T} = \frac{1}{N} = \sum_{i=1}^{N} T_(i)$$

が得られる. これらから誤差評価, 及び区間推定も行うことも可能である.

3　分位点

一般に確率分布の 100λ % $(0 < \lambda < 1)$ の位置に対応する値を 100λ パーセンタイルまたは 100λ 分位点という. 例えば四分位点の場合, 25% 分位点, 50% 分位点, 75% 分位点と表される. 本研究では与えられたデータの従う確率分布に対して様々な分位点の推定値を求めるためにブートストラップ法を用いることにする.

4　回帰分析

観測データ,

$$Z_n = (Y^{(i)}, X_1^{(i)}, X_2^{(i)}, \cdots, X_m^{(i)}), \quad i = 1, 2, \cdots, n$$

がえられたとして, 目的変数 Y が, 説明変数 X_1 から X_m にどのように依存するかを調べる手法を一般に回帰分析という.

線形回帰分析

説明変数 X_k が目的変数 Y に対して依存関係をもつとき,

$$Y^{(i)} = \alpha_1 X_1^{(i)} + \cdots + \alpha_m X_m^{(i)} + \beta + \varepsilon_i, \quad i = 1, 2, \cdots, n$$

線形回帰モデルという.
α_1 から α_m, β は回帰係数, ε_i は誤差項を表している.

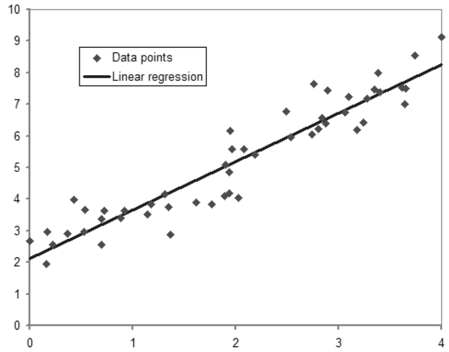


図 2: 回帰直線

線形回帰分析の利点として, 変数間の関係を数式モデル化することで, 既知の説明変数から未知の目的変数の値を予測できることや目的変数に対する説明変数の影響の大きさを評価できることが挙げられる.
そのため, 観測値からの事象の予測, シミュレーション, 検証, 要因分析など広く用いられる手法である.

5　提案手法

まず, $\alpha = 1.1 \sim 1.9$ の安定分布に従う乱数を発生させ, それぞれの場合についてブートストラップ法を用いた後, 各 α に対する安定分布の様々な分位点の推定値を求め, これを各安定分布の分位点の理論値として代用する. (裾の重い分布について調べたいので, 分布の上位・下位とも細かく取る.)

	α=1.3	α=1.7	α=1.8	α=1.9
1%	−12.55345	−5.14542	−4.513100	−3.66875
2%	−7.337125	−3.84315	−3.382028	−3.09097
5%	−3.759650	−2.64376	−2.503592	−2.40847
10%	−2.288432	−1.93054	−1.886334	−1.85544
15%	−1.656507	−1.50367	−1.494462	−1.49122
20%	−1.263432	−1.99485	−1.203143	−1.20230
25%	−0.975438	−0.95145	−0.958410	−0.96736
50%	0.0000000	0.000000	0.0000000	0.000000

図 3: 分位点の推定値の一部

観測されたデータの従う分布についても同様に分位点の推定値を求める. $\alpha = 1.1 \sim 1.9$ の各安定分布の分位点の理論値と観測データに基づく分位点の推定値に対して回帰分析を行う. その結果もっともあてはまりが良い α を, 観測されたデータの安定分布指数と推定する手法を用いる.

6　実データへの応用

2007 年から 2009 年の円ドル為替レートの日終値の前日比での対数分布で実験を行った.

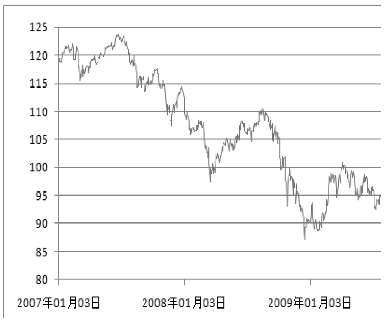


図 4: 円ドル為替レートの日終値の変動

分位点	観測値
1%	−0.00980
2%	−0.00787
5%	−0.00566
10%	−0.00407
15%	−0.00312
20%	−0.00259
25%	−0.00200
50%	0.00000
85%	0.00291
90%	0.00356
95%	0.00497
98%	0.00700
99%	0.00887

図 5: 分位点の推定値の一部

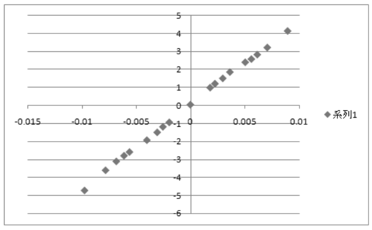
パラメータ	α=1.1	α=1.2	α=1.3	α=1.4	α=1.5	α=1.6	α=1.7	α=1.8	α=1.9
決定係数	0.837718	0.866082	0.898525	0.93298	0.958416	0.985441	0.996778	0.997843	0.989933

図 6: 回帰分析の結果 1

ただし決定係数とはモデルとデータとの適合度の良さを測る尺度であり, 1 に近いほどあてはまりが良い.
表より, まず, $\alpha = 1.7 \sim 1.8$ の間であると推定される. さらに細かく見てみると $\alpha = 1.77, 1.78$ が適当であると考えられる.

α=1.79	α=1.78	α=1.77
0.998264	0.99899	0.998838

図 7: 回帰分析の結果 2



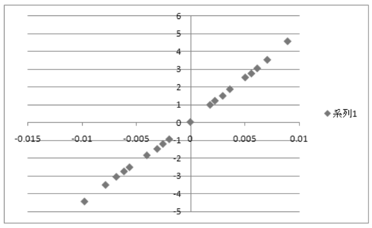


図 8: $\alpha = 1.77$(上), 1.78(下) のときの回帰直線 (ほぼ直線上にデータが並んでいる)

7　まとめ

円ドル為替レートの日終値の他にも日経平均株価終値や金の価格変動などの金融データで試したところ, 同様に $\alpha = 1.7 \sim 1.8$ という結果が得られた. 一般に金融市場の変動の多くは $\alpha = 1.7$ に近い安定分布に基づいていると言われている.
よって, 本研究は簡便でありながら, かなり良い推定を与えていると考えられる.

参考文献

1. 研究報告,"ブートストラップ法を用いた分布の裾指数の推定手法の改良"
2. http://www.motorwarp.com/koizumi/quantile.html
3. http://d.hatena.ne.jp/ryamada/20101126/1290237997