更新日：2022年1月29日

#### An Application of Extreme Value Theory for Measuring Financial Risk

Manfred Gilli, Evis KÄellezi,

# 概要

本論文「[An Application of Extreme Value Theory for Measuring Financial Risk](chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=http%3A%2F%2Fwww.unige.ch%2Fses%2Fdsec%2Fstatic%2Fgilli%2Fevtrm%2FGilliKelleziCE.pdf&clen=856477&chunk=true)」は，主要な株式市場指数に極値理論（Extreme Value Theory）ををtailリスク指標にどのように利用できるのか，その手法の提案をしたものである．本論文で使用している株式市場指数とデータ取得期間は以下のとおりである．

|  |  |  |
| --- | --- | --- |
| 株式指数 | Start | End |
| Euro Stoxx50 | 1987/1/2 | 2004/8/17 |
| FTSE100 | 1984/1/5 | 2004/8/17 |
| Hang Seng | 1981/1/9 | 2004/8/17 |
| Nikkei225 | 1970/1/7 | 2004/8/17 |
| Swiss Market Index | 1988/7/5 | 2004/8/17 |
| S$P500 | 1960/1/5 | 2004/8/17 |

また，論文著者はMatLabにてデータの分析を実施しており，論文中に使用したコード等を載せたサイト[[1]](#footnote-2)を記載しているが，現在は使用できない．

本論文では，tailリスクの指標として「Value at Risk」，「期待ショートフォール」，「リターン・レベル」を使用している．次章にて3つの指標について簡単に説明する．3章にて，極値理論を用いた場合におけるtailリスク指標の評価式を導入する．

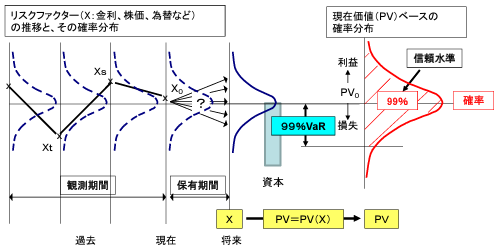
# 3つのtailリスク指標

## Value at Risk（VaR）

バリュー・アット・リスク法は，金融商品の種類に関わらず同じ基準で市場リスクを計測し，一つの数値（VaR）で表すことができるため，国際的にも広く認められている手法である．VaRとは，統計的考え方に基づき，一定の確率で今後ある保有期間内に生じる可能性のある最大の損失額のことを指す．

連続分布関数をもつ確率変数は，ある時点におけるある金融商品の損失を表すとする．このとき，を分布の番目の分位数として定義する．

ここでは，分布関数の逆数として定義，つまり分位関数である．社内のリスク管理のため，多くの金融機関では1日の保有期間に対して5％，1%のVaRを算出している．VaRのイメージ図[[2]](#footnote-3)は以下の通り．



## 期待ショートフォール

期待ショートフォールとは，VaRを超えるような巨大損失が実現した場合，「平均的」にどの程度の損失額が想定されるかを表すリスク尺度である．期待ショートフォールは，を超える損失の期待値として定義される．

Artznerら(1999)は，VaRとは対照的に，期待ショートフォールが首尾一貫したリスク指標であると主張している．また，バーゼルⅢ最終合意における市場リスクの見直し[[3]](#footnote-4)では，トレーディング勘定と銀行勘定の境界の見直し，内部モデル方式についてVaRモデルから期待ショートフォールへの移行，標準的方式の全般的見直しなどが行われる予定である．

## リターン・レベル（テキストでいう「再現レベル」）

を同一の長さかつ重複のない期間内に観察される最大値の分布であるとする．つまり，ある期間における最大値の分布関数とする．このとき，以下で定義されるをリターン・レベルと定義する．ただし，はブロックの個数，を1ブロックのおけるデータの個数，を分布関数の逆関数とする．

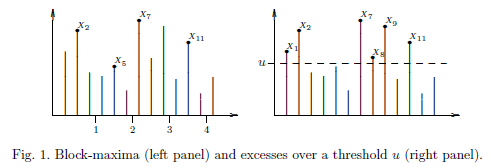
リターン・レベルは，1期間でVaRなどを超過すると予想される水準であり，ポートフォリオの最大損失の尺度として利用することができる．

# 極値理論

極値統計学は，異常に大きな値をとるデータの出現（とる値のその確率，すなわち確率分布）に興味がある．極値理論は，1953年のオランダの大洪水の経験から，長期間にわたって海抜０メートル以下の国土の安全性を保つために，防潮堤の高さをいくらにすればいいかを数学的に求めたことが出発点である．

確率変数の最大値をモデル化するとき，極値理論は中心極限定理と同様に基本的な役割を果たす．つまり，極値理論と中心極限定理のどちらも，極限分布の具体的な形を示している．

一般的に，取得データで外れ値を識別するためには，2つの方法が考えられる．1つ目は，例えば月や年単位での連続した期間における最大値を考えてみる．取得期間をブロックに分けて，各ブロックにおける最大値を集めると，その集合は極端な事象で構成される．下図の左パネルにおける観測結果は， 4つの期間のブロック最大値を表している．2つ目は，閾値を超えるデータに焦点を当てる方法である．上図の右パネルにあるデータ値はすべて閾値を超えており，極端な事象で構成されている．



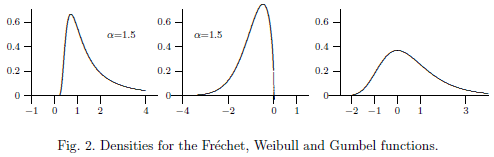
どのようにデータを取得するかを「データ取得デザイン」という．極値統計学でよく用いられるデザインは3つあるが，ここでは，ブロック最大データと閾値超過データを利用する．ブロック最大データは，ブロック数が少ない場合，効率的にデータを利用できないため，最近では閾値超過データが用いられることが多い（らしい）．

## Distribution of Maxima

ブロックの大きさをとし，をブロック最大値とすると，以下の定理が成り立つ．

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ≪Theorem 1(Fisher and Tippett (1928), Gnedenko (1943))≫  を独立同分布な確率変数とする．このとき，以下を満たすある定数と非減少な分布関数が存在するとする．  このとき，は以下の3つのうちどれか一つの極値分布に従う．   |  |  |  | | --- | --- | --- | |  | ： |  | | Weibull | ： |  | | Gumbel | ： |  | |

フレシェ分布，ワイブル分布，グンベル分布の確率密度関数の形状は以下の通り．



分布関数，確率密度関数の形より，ワイブル分布では上限があり，それ以上の値はとらない．グンベル分布では上限はないが，減速が早く大きい値を取る確率は小さい．フレシェ分布については，上限はなく，減速はあまり早くないため大きい値を取る確率が小さくない．このことより，ワイブル分布は有限分布の漸近的な分布，グンベル分布は指数関数的に減衰する分布の裾，フレシェ分布は多項式的に減衰する裾，つまり裾が暑い分布の表現に適している．

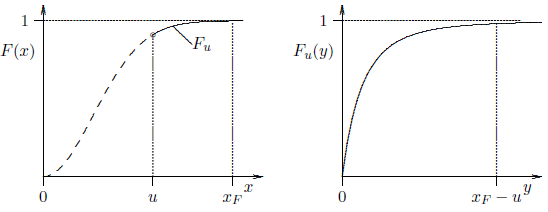
Jenkinson (1955)とvon Mises (1954)は，以下の関数で3つの分布の一般形（GEV）を導入した．

一般に，サンプルの最大値の極限分布（または極値分布）がどのようなものか事前に分からないので，一般化された表現は，最尤推定値を計算しなければならない場合に特に有用である．更に，上式で定義された標準的なGEVは，正規化された最大値の極限分布である．実際は真の分布は分からず，加えてとについてもわからないことから、新たに2つのパラメータとを追加してGEVを表現する．は位置パラメータであり，はスケールパラメータである．

パラメータの推計結果をそれぞれとしたときに，それによって得られるリターン・レベルは以下のとおりである．の値が7であれば、1年間に観測される最大損失が平均して10年に1度7％を超えることを意味する．

## Distribution of Exceedaneces

POT（Peak Over Threshold）法と呼ばれる別のアプローチでは，ある閾値を超える値の分布を考える．下図に示されているように，確率変数の（未知の）分布関数を考える．ここでの目的は，ある閾値を超えるの値の分布関数を推定することである．



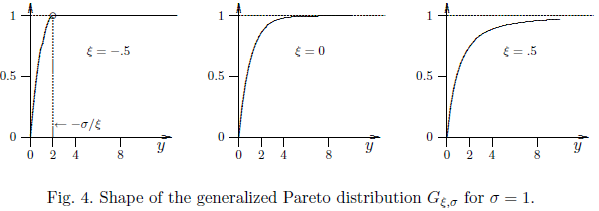
分布関数は条件付き超過分布関数（conditional excess distribution function）と呼ばれ，は確率変数，は閾値、は過剰分，は分布関数の右端点である．また，はを用いて表現できる．

確率変数の実現値は，だいたいとの間にあるので，この区間での分布関数の推定には一般的に問題はない．しかし、分布関数の一部分であるの推定は，一般にこの区間での観測がほとんどないため，難しくなる．

この難点において極値理論は，次のような条件付き超過分布関数に関する強力な定理を示してくれている．

|  |
| --- |
| ≪Theorem 2 (Pickands (1975), Balkema and de Haan (1974))≫  条件付き確率変数の分布関数は，が十分大きいとき  ただし，  は一般化パレート分布（GPD）と呼ばれる．ならば，GPDは以下の通りに表現できる． |

下図は，一般化パレート分布の形状を示したもので，形状パラメータまたは極値指数と呼ばれるが，負，正，ゼロの値をとるときのものである．ただし，スケールパラメータは1とする．



極値指数は，分布の裾の厚さを示すもので，が大きいほど裾が暑くなる．一般に，金融機関の損失に上限を設けることはできないため，金融機関の収益分布のモデルには，形状パラメータを持つ分布のみが適している．

裾の分布にGPD関数を仮定すると，およびの分析式はGPDパラメータの関数として定義できる．条件付き超過分布関数をについて解くと，

をGPDに，を推定値に置き換えると，（ただし，は観測データの総数，は閾値以上の観測データ数）次式が得られる．

この式は以下の様により簡単にできる．

分布の上側確率点は，次の通り．

上式をについて解くことで，を求めることができる．

期待ショートフォールを以下の通りに書き換える．右辺の第2項は閾値の超過分の期待値である．

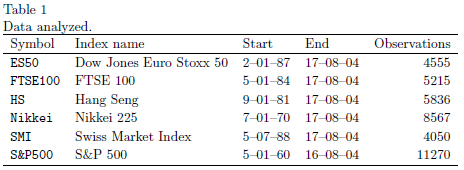
極値指数のGPDの平均超過関数は以下であることが知られている[[4]](#footnote-5)．

この関数は，閾値の値を変化させたときの確率変数の過剰分の平均を与えている．モーメント（ここでは1次モーメント）の存在に関するもう一つの重要な結果は，がGPDに従う場合，であるようなすべての整数に対して，個の1次モーメントが存在する．

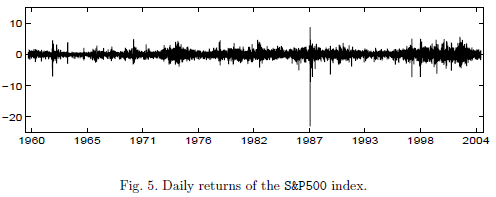
同様に，期待ショートフォールの定義が上式で与えられ，平均超過関数を用いて およびに対するの超過を表すについて，以下が成り立つ．

# Application

論文の目的は，日次リターンより裾の分布を推定し，その結果を市場リスクの定量化に利用することである．以下の表は，論文中で分析したインデックスのリストである．ここでは，主にS&P500指数に焦点を当て，信頼区間と推定値のグラフによる視覚化を行っているが，他の系列については点推定値のみ記載している．

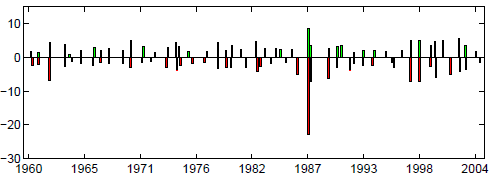


S&P500インデックスの 個の日次リターンをプロットしたものは以下の通り．



## Method of Block Maxima

ブロック最大値の適用により，の点推定値と区間推定値を計算する．季節効果を避けるため，また定理1が成立するのに十分な大きさの年単位の期間を選択する．したがって，すべてのブロックが同じ長さではない．S&P500のデータサンプルは45の重複しないサブサンプルに分割されており，それぞれのサブサンプルには1年間の日次リターンが含まれている．各ブロックの最大値は，一般化極値分布（GEV）の推定に使用される．下図は，S&P500の左テールと右テールの年間最大値をプロットしたものである．



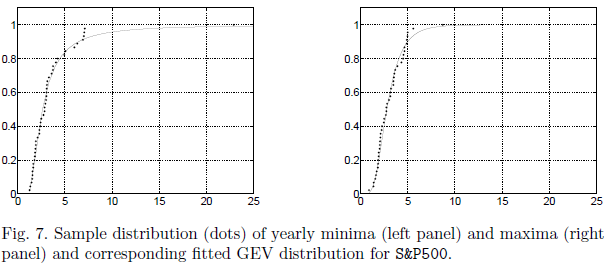
3つの未知のパラメータに関して最大化する対数尤度関数は以下の通り．

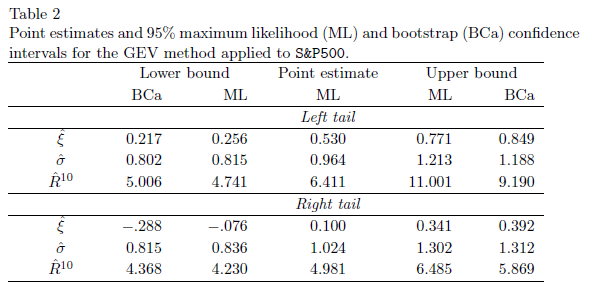


ただし，，のとき以下は確率密度関数となる．

の時，確率密度関数は以下の通り．

下図は，標本分布関数とそれに対応するGEV分布をプロットしたものである．また，パラメータの点推定値と区間推定値を以下の表に示す．

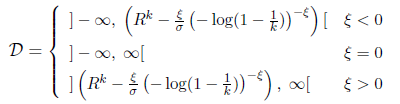




### Interval Estimates

区間推定を計算するためには，GEV 分布を未知のリターン・レベル の関数として直接再パラメータ化することで，分位点推定問題にアプローチすることが有用である．3章で示した一般化極値分布の式とリターン・レベルを用いてGEVの分布関数を表すと以下の通り．

は以下で定義され，これよりの最尤推定値を直接求めることができる．



その後，プロファイル対数尤度関数を用いて各パラメータの信頼区間を計算することができる．例えば，関心のあるパラメータがの場合，プロファイル対数尤度関数は以下で定義される．

導き出された信頼区間には，以下の条件を満たすのすべての値が含まれている．

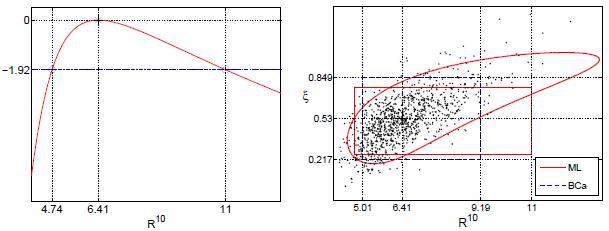
ただし，は自由度１のカイ二乗分布のレベル分位点を表す．関数

は，相対プロファイル対数尤度関数と呼ばれ，下図の左パネルにプロットされている．の6.41%という点推定値は、かなり大きな区間に含まれている．この区間は，高次のクオンタイルほど観測データが少ないため，非対称になっており最大損失の上限値がより不確実であることを示している．

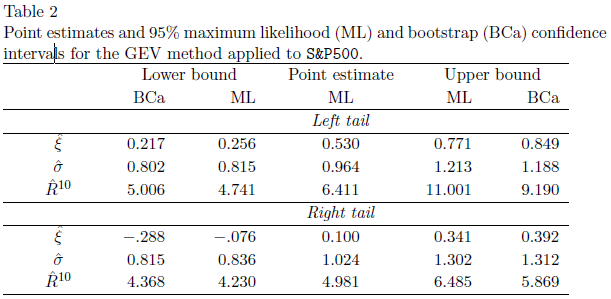
場合によっては，極値分分布の裾の厚さを特徴づけるの値にも関心がある．この場合，との両方に対する信頼区間が必要となる．この場合，プロファイル対数尤度関数は次のように定義される．

下図の右パネルでは，S&P500のとについて，95%の信頼区間を表している．同じグラフに，1000個のブートストラップ・サンプルで推定されたペアもプロットしている．信頼区間は，ブートストラップペアの約95%をカバーしており，信頼区間を計算することで取り得るパラメータ値を理解できる．

また，サンプルサイズが小さいことを考慮して，BCa（bias-corrected and accelerated）ブートストラップ法を用いて，信頼区間を算出した．その結果，RについてBCa区間はに縮小した．極値指数については，その差はそれほど顕著ではない．しかし，いずれの場合も極値指数は生の値となっており，Frechet分布であることを示している．

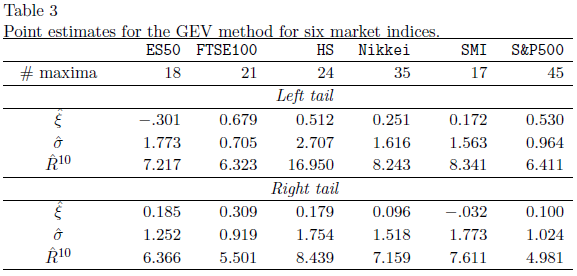


S&P500のGEV分布の点推定値と信頼区間を以下表にまとめた．



の場合，となり1年間に観測された最大損失がを超えるのは，平均して10年に1回ということになる．同様に，の損失を超えるのは平均して100年に一度だけであることが導き出される．

以下の表は，6つのインデックスすべてについてGEVの点推定値をまとめたものである．ES50とSMIは観測数が少ないため，点推定値の信頼性は低い．またハンセン(HS)の左テールのRの値が高く，次にリスクの高い指数の2倍になっている．



データサンプル中の極端な値に関する情報をよりよく利用する方法の一つとして，POT法がある．Coles (2001, p. 81) は，GPD を用いたリターン・レベルの推定を提案しています。しかし、データセットが十分に大きい場合、ブロックが十分に大きい場合には、データのクラスタリングの問題を避けることができるため、GEVが有用であることがわかります。さらに、閾値uの選択が不要なため、推定が簡素化される。

1. サイト（ただし，現在は利用できない）：[www.unige.ch/ses/metri/gilli/evtrm/](https://d.docs.live.net/5a6fcaf58e44607b/デスクトップ/自己研鑽用/www.unige.ch/ses/metri/gilli/evtrm/) [↑](#footnote-ref-2)
2. 参考サイト：[chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A%2F%2Fwww.boj.or.jp](chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A%2F%2Fwww.boj.or.jp%2Fannouncements%2Frelease_2015%2Fdata%2Frel150929a2.pdf&clen=534467&chunk=true) [↑](#footnote-ref-3)
3. 参考文献：[chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A%2F%2Fwww.dir.co.jp](chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A%2F%2Fwww.dir.co.jp%2Freport%2Fresearch%2Flaw-research%2Fregulation%2F20190125_020597.pdf&clen=314218&chunk=true) [↑](#footnote-ref-4)
4. 証明は以下のサイトを参照されたし．

   [chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=http%3A%2F%2Fwww.msi.co.jp](chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=http%3A%2F%2Fwww.msi.co.jp%2Fsplus%2FusersCase%2Ffinance%2Fpdf%2Furatani.pdf&clen=98984&chunk=true) [↑](#footnote-ref-5)