

## 「特集 極値理論」について

高橋 倫也<sup>1</sup> ・ 志村 隆彰<sup>2</sup>

### 1. 極値理論

高度に発達した社会では、希にしか起きない現象によって人は大きな被害をこうむることがある。そこで、平均ではなく希な現象を取り扱う極値理論を考える必要がある。

極値理論は古くから水文学や工学の分野で防災のために研究・応用されてきた。現在では環境や金融工学の分野でも研究・応用されている。極値理論は標本または確率過程の極値(最大値または最小値)の漸近挙動を問題にする。統計学では、限られた期間(または領域)で得られたデータを利用して、与えられた長期間(または広領域)でのデータの取りうる最大値の予測が重要である。このとき、得られたデータの最大値で長期間のそれを予測するのは賢明でなく、極値理論に基づきデータの外挿を行う必要がある。一方、応用確率論では種々の確率過程での極値統計量の漸近分布の研究が盛んである。

確率標本の極値統計量の漸近分布の研究は 1920 年代に始まり、3 つのタイプがあることが分かった。これら 3 つのタイプの極値分布を一般極値分布として一つの式で表し極値データの解析が行われている。しかし、一定期間ごとの極値データではデータ数が少なく推測の精度は良くない。そこで、データを増やす試みがなされた。ある閾値以上の超過データに適合させる一般パレート分布と各期間ごとで上位  $r(\geq 2)$  個のデータに適合させる同時漸近分布が導出され、これらの分布はデータ解析に応用されその有用性が示されている。

多変量極値の研究は 1960 年代頃に始まった。2 変量の極値分布の特徴付けから始まり、一般の多変量極値分布の研究が行われた。最近では多変量極値分布のパラメトリック・モデルが提案され環境データの解析に用いられている。また、時系列での極値理論の研究も行われ金融工学の分野で応用されている。

現在、極値理論の研究はヨーロッパやアメリカの少数の有力な研究者を中心にしたグループにより活発に行われている。日本では古くから水文学や工学の分野で研究・応用されてきた。また、統計や確率の分野での研究もあるがそれほど多くはない。ところで、極値理論の研究者は多くはないが、次節で紹介する統計数理研究所の共同研究集会「極値理論の工学への応用」は 10 年間続いており最近同研究所よりレポートを発行している。

### 2. 共同研究集会「極値理論の工学への応用」

ここでの紹介は、高橋・渋谷(2002)に従いその一部を書き直したものである。

#### 2.1 研究経過

1993 年 5 月、ワシントン DC 郊外にあるアメリカ商務省の National Institute of Standards and Technology (NIST)において、シンポジウム Extreme Value Theory and Applications が開

<sup>1</sup> 神戸大学 海事科学部：〒658-0022 兵庫県神戸市東灘区深江南町 5-1-1

<sup>2</sup> 統計数理研究所：〒106-8569 東京都港区南麻布 4-6-7

かれた。極値理論の古典時代が終わって新しい時代に入っただけでなく、新しい理論が多くの応用分野に浸透して利用分野を広げると同時に、現実問題解決要求が研究者を刺激している状況を示す会議であった。その後国際会議がより盛んとなり、専門雑誌 *Extremes* (Kluwer) も 1998 年に創刊された。

この会議に、予想より多くの日本人が出席し、しかも互いに接触のない分野の人々であったので、翌年から統計数理研究所で共同研究集会を開催させていただき交流を続けている。

以下、この研究集会等での講演をもとに、参加者等が工学分野でどのような極値理論の研究課題に関心があるかを述べる。研究の成果よりも未解決問題を中心とした紹介を行う。

## 2.2 自然と災害

気象学：降水量、気温、風速などの極値は、土木、建築の安全性規準の指標として重要である。当然季節変動があり、週、月の範囲で強い従属性があり、主に年間最大、最小を議論するが、国内の利用可能データは高々 100 年(国土庁調査)に限られ、常に制約となる。風速など年度を越えた極値を議論し始めているが、過去の年間最大値との対応が問題となる。 $k$  日間雨量、 $k = 1, 2, 7$ , の年間極値も重要視される統計量であるが、これらの間の関係は明らかでない。

地球温暖化によって発生している超異常気象についても、なにを尺度として異常の程度を示すかよい方法がない。その研究のために地表の自然にある痕跡から過去の気象を推定するプロジェクトなどもある。

水文学、海岸水理学、海洋学：これらの分野は気象と工学の間にあって、極値データ解析が日常的に行われており、問題の提起、手法の開発が盛んであり、理論への影響が強い。水文学で始まった Peaks Over Threshold (POT) 手法の開発が、現在の極値統計学に与えた影響は大きい。

土木工学：日本の土木産業・技術の後進性のために、長い間データが隠匿された歴史がある。最近になりようやく諸設計の情報公開化に向けて、データ解析が進むようになっている。しかし自然環境が破壊されており、気象条件が定常でなく、局所地形の影響があるなど、難問が多い。

建築工学：耐震建築の設計では、建築地点における最大震度の予測が必要であるが、地震データベースが整備されているために、可能となっている。強風にたいする強度は当然風向を考慮する必要がある。風向風速を同時に測定したデータがあれば、たとえば 8 方向に分類して多変量年間最大風速を解析する。多変量、特に 2 変量極値の理論は 1960 年代から始まったが、実用化は最近のことである。

## 2.3 実験データ

信頼性、寿命：信頼性の中心は、最小寿命の解析であり、そのためもっぱらワイブル分布モデルが利用されている。実験データの解析という点で、2.2 節の分野と異なる。しかし非正規理論を扱い、実験は長時間を要することが普通であり、ユーザーが使用する環境では粗い測定しかできない。主流は加速寿命試験の開発であり、そのためのモデル、実験、測定法が議論されている。

以下の分野も広い意味での信頼性である。

絶縁：電気関係は比較的实验が容易であり、古典的極値理論も適用しやすく、数値解析法を含めて、極値解析が研究されている。

腐食：この分野では、古典的な方法にもとづくソフトウェアが開発、市販されたために、普及したものの応用範囲の拡張が遅れている。実験の計画に役立つ理論、ソフトウェア開発が必要である。

宇宙工学：人工衛星上のコンピュータ素子は、強い宇宙線で常に破壊される。宇宙線の強度

は太陽活動に依存し、11年周期で変化する。人工衛星の寿命も考慮して、最大宇宙線量にたいする安全設計をする。

金属疲労：鉄鋼の疲労強度がそのなかの非金属介在物の最大寸法に依存するという理論に基づきその予測法が研究されている。

#### 2.4 その他の分野

超高齢：社会が豊かになり、非衛生活が改善され、注意をはらえば食生活を健全にでき、致命的疾患の医療が進歩したため、人間の寿命にたいして極値理論を適用する条件が整っている。

金融工学、損害保険：金融工学の目的はリスクの制御であると言われているが、もちろんその中で極値の役割は大きい。損害保険は伝統的に統計学、応用確率論に貢献してきたが、たとえば再保険は極値理論が重要な役割を果たしている。

文献：研究集会で紹介された極値理論に関する最近の本としては、Kotz and Nadarajah (2000)、Embrechts et al. (2001) や Finkenstädt and Rootzén (2003) 等がある。これらの本から極値理論の最近の研究動向がうかがえる。

ソフトウェア：次の様な極値データ解析用ソフトが紹介された。ソフト Xtreme は Reiss and Thomas (2001) の書物付属の CD-R で配布されている。これに基づく解析が書物で説明されている。Coles (2001) には S-PLUS によるデータ解析ソフトをダウンロードできるサイトの情報がある。また、Stephenson (2003) のホームページでは R によるソフトをダウンロードできる。

### 3. 論文紹介

「統計数理」の今回の企画は「極値理論」というタイトルで、この分野の現状と日本における研究の一端を紹介し、より多くの統計学の研究者にこの分野に関心を持ってもらいたいという思いから生まれた。特集号への執筆依頼を、前節で紹介した研究集会の参加者を中心に行った。そのため、応用が伝統的な工学の分野に偏ったものになっている。最近極値理論の研究・応用が盛んである金融工学等の社会科学の分野の論文は残念ながら時間の関係で載せることが出来なかった。しかし、古くからの極値理論の研究者の論文を多く掲載することが出来たと思う。

以下、掲載されている論文の紹介を行う。論文のテーマは、極値理論におけるデータから、確率、統計そして工学への応用にまたがっている。

木下論文 は、自然災害への極値理論を応用するさいに利用する降水量、河川流量といった観測データの紹介と測定における非ランダム誤差について概説している。また、全国的な観測状況と歴史的な災害の記述データについても述べている。

吉原論文 は、標本の独立性の仮定を弱めた弱従属性のもとでの極値統計量の漸近理論とその応用について、独立な場合と弱従属の場合を比較しながら概説している。

前島論文 は、独立同分布確率変数の和から大きなものあるいは小さなもの(すなわち極値データ)をいくつか取り除いた trimmed sum に関する概収束と弱収束の広範な結果を概説している。

松縄・中村論文 は、順序統計量の中で標本数の増加に応じて変化する小さい(あるいは大きい)複合極値統計量の確率分布について、情報ネグトロピの意味での方向近似を中心に論じている。

河村・岩瀬論文 は、最大エントロピー法による一般化パレート分布の特徴付けを行い、それに基づくパラメータの推定法を構成し実データに適用している。

高橋・渋谷論文 は、上位  $r$  個のデータを用いる上側微小確率点の推定法を紹介し、 $r$  個を用いることにより極値データのみを用いる場合と比べどの程度推定精度が改善されるかを論じ、腐食深さデータの解析を行っている。

渋谷・華山論文は、連続変数に対して議論されている極値理論を超高齢者の年齢時代区分データへ適用する方法を提案し、日本人男性の寿命分布の限界を求めている。

北野論文は、海岸工学の長期波浪統計解析で用いられる極値分布と通常ワイブル分布に対し提案されている裾長さの意義を説明し、新たな指標として裾厚度を提案している。これらの指標により分布特性の把握や比較が可能になることを示し実データに適用している。

神田・西嶋論文は、多変量極値データ解析法について議論し、建築物の構造設計で想定すべき外力として代表的な強風と地震について、それぞれ実データとシミュレーションで多地点危険度解析を行っている。

廣瀬論文は、絶縁分野におけるインパルス破壊電圧の推定について、従来の段階昇圧法と最近の破壊値の情報が得られる高速な電圧測定装置により可能となった方法との比較を行い、後者の優位性について述べている。

極値理論は今後ますますその応用範囲を広めていくと考えられる。このような時点で、今回の特集がきっかけとなり、極値理論が統計学者の興味を引きこの分野の研究に参加される方が出てこれたら、筆者たちの望外の喜びである。

#### 参 考 文 献

- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, London.
- Embrechts, P., Klüppelberg, C. and Mikosch, T. (2001). *Modelling Extremal Events for Insurance and Finance*, 3rd ed., Springer (1st ed., Springer, 1997), Berlin.
- Finkenstädt, B. and Rootzén, H. (2003). *Extreme Values in Finance, Telecommunications, and the Environment*, Chapman & Hall, New York.
- Kotz, S. and Nadarajah, S. (2000). *Extreme Value Distributions: Theory and Applications*, Imperial College Press, London.
- Reiss R.-D. and Thomas, M. (2001). *Statistical Analysis of Extreme Values*, 2nd ed., Birkhäuser, Berlin.
- Stephenson, A. (2003). Software for Extreme Value Theory, <http://www.maths.lancs.ac.uk/~stephena/software.html>
- 高橋倫也 編(2004). 極値理論の工学への応用, 統計数理研究所共同研究レポート, No.169.
- 高橋倫也, 渋谷政昭(2002). 極値理論の工学への応用, 2002年度統計関連学会連合大会講演報告集, 17-18.