研究紹介

確率分布の裾の研究

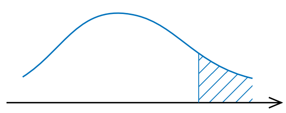
数理・推論研究系　志村　隆彰

近年理論応用とにかかわらず、裾が重い確率分布（heavy tail）の研究が盛んになってきているが、筆者の最近の研究対象も確率分布の裾の挙動である。分布の裾とは100歳以上の人の割合とか、年収いくら以上の人の割合といった、ある値以上（或いは以下）をとる確率のことである。上限のない確率分布の裾は、値を大きくしていくと正を保ちながら、0に近づいていく。このときの速さは様々であり、速いとき分布の裾が軽い、遅いとき裾が重いという。裾が軽ければ、極端に大きい事象が起こる確率は十分に小さく無視できるので平均的に起こることが重要であるが、裾が重い場合は極端に大きい値をとる確率が意味を持ってくる。確率法則として有名な大数の法則、中心極限定理はランダムな数値の平均はサンプルの数を増やしていけば、それらが従う確率分布の平均に近づく、或いはそれからのずれが正規分布に従うことを主張していて、それぞれのクラスに様々な体格の人がいるのに、クラス平均値はどこもほとんど変わらないといった例からも理解しやすい。ところが、これは裾が軽いから成立つもので、裾が重い世界では、ランダムな値の和の挙動が正規分布以外の分布へ近づいたり、値の和がそのうちの最大のものただひとつとほとんど変わらないことさえも起こりうる。

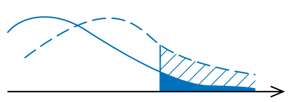
裾の挙動が主役の分野に極値統計学がある。世界新記録、長寿日本一などは人々の関心を引く極端な値であるが、大地震や大雨などのように、数多く起こる小さなものの蓄積よりも稀に起こる極端に大きいことが現実的意味をもつ現象は少なくない。一般に、建築工学、信頼工学、保険数学、ファイナンスなどの安全性やリスクを扱う工学の分野では極値理論が大きな役割を果たす。

さて、筆者は確率論の中の無限分解可能分布と関連する分野を研究してきた。これは名前の通り、合成積（convolution）の意味で任意の数に分解可能な分布で、ある確率変数 X の分布 μ が無限分解可能であるとは、任意の自然数 n に対して、n 個の独立同分布確率変数列 $X_1,\ldots,X_n$ が X と $X_1+\cdots+X_n$ の分布が同じに取れるときをいう。正規分布、複合ポアソン分布、安定分布などが代表的な例である。その中で、ここ数年は渡部俊朗氏（会津大学）と共同で subexponential 分布の一般化の研究をしている。Subexponential分布とはそれに従う独立な非負値確率変数を X と Y としたとき $P(X+Y>x)\sim 2P(X>x)$ となるときをいい、その全体を subexponential class S をいう。片側安定分布、パレート分布、ワイブル分布（助変数1未満）などがそうであり、重い裾の典型である。一般に分布の合成積をとれば裾は元の裾と比べて漸近的に少なくとも2倍になるから、S の分布はその下限であり、これを2より大きい定数まで広げたものが convolution equivalent class である。我々はさらに拡張を試み、比 $P(X+Y>x)/P(X>x)$ が有界である分布を O-subexponential OS と名づけた。半安定分布はこれに含まれる。S に関しては、無限分解可能分布の裾とその Lèvy 測度の裾の漸近挙動同士の関係が知られていたが、OS で対応することを考察し、S の場合とは本質的に異なるものを含む興味深い結果を得た。最近は、もっぱら分布族の合成積に関する性質を研究している。確率変数でいえば、独立確率変数列とその和の分布の裾の相互関係であり、部分と全体、原因と結果との関係と見ることも出来る。ある分布族に含まれる分布同士の合成積が再び元の分布族のものになるのか、或いは合成積がある分布族に入っていれば元の分布自身もそうなのかといった基本的な問題である。理論的整備という意義もある、簡単そうに見えて、難解な問題である。S が合成積について閉じていないことはよく知られた意外な結果であるが、我々も独立同分布に従う確率変数和の分布の裾が指数的であっても各々の裾はそうとは限らないことを示した。これも古くからの予想に反する意外な結果であった。

尚、統計数理研究所では、極値理論及び無限分解可能分布をテーマとした共同研究集会を開催している。ともに10年以上続いていて、報告集共同研究リポートを発行している。極値理論については新設されたリスク解析戦略研究センターとの連携をとった活動も始めている。詳しいことは統計数理研究所のホームページ等で告知しているのでご覧頂きたい。



確率分布の裾



重い裾と軽い裾