

意外と身近な“極値”の世界

回答者：志村 隆彰（統計数理研究所 数理・推論研究系 リスク解析戦略研究センター 准教授）

いただいたご質問には極値統計に限ったことではないように思えますので、広く統計一般の視点からお答えします。また、今日の世界は一年前には全くの「想定外」です。安城哲三先生のお話（研究トーク3）のように、新型コロナが気候変動にも影響を及ぼすというのですから、未来予測がいかに困難であるかわかります。統計予測は、数値だけで未来を予測するという極めて困難な問題に挑戦しているので、おのずと限界もありますが、にもかかわらず極めて有効な手法であると私は考えています。

1. 極値理論で予測ができるということですが、それも外れたりするということでした。ということは、極値理論に極値理論を適用すると、予測の当たりはずれも予測できることになりませんか？

申し訳ありませんが、「極値理論に極値理論を適用する」の意味がわからないので、回答が難しいです。

2. 自然災害に関する統計学上の数値に誤差が問題になることはありますか。

統計理論を現実の問題へ適用する場合、統計理論は、それを適用する現実の問題の数値データに、測定精度をはじめとする様々な条件を求めます。人工的に管理された工業製品のデータなどでは比較的望ましいデータが得られるので理論通りの解析がしやすいのですが、一般的には、必ずしも理論の仮定を満たす理想的なデータが得られるとは限りません。むしろそのような場合は少ないので、理論の仮定を満たす、質の良いデータを得ることは統計解析において、普遍的な課題です。中でも自然現象は、多くの要因が複雑に絡み合っ起こる複雑ものですし、理論が要求するデータを得ることが難しいと言えます。

さて、誤差にもいろいろな種類があり、質問者のいう誤差がどのような誤差か正確にはわかりませんが、測定誤差のようなものとして話を進めます。工業製品のような場合とは違い、自然災害の場合、数値データを収集すること自体が困難な場合も少なくなく、測定精度に限らず、理論の仮定を満たす、質の良いデータを得ることが非常に困難です。ですから、誤差も大きいといえますが、それ以外の要因がそれ以上に大きいので、相対的には測定誤差の問題は小さいように思います。

このような理由で、自然災害に対する統計解析はとても難しいのです。

3. 統計的な極値を用いた地震の規模の予測は、例えば、地球物理学などによる力学的に予測される上限と一致したりするものでしょうか？

私は力学的なことは知りませんし、一致するかどうかは、程度問題でもありますが、おそらくある程度のずれが出ると思われま。講演でスポーツの記録の限界についてお話ししましたが、スポーツの記録の上限予想も生理学的な視点からの予測がありますし、一般に予測は様々な手法でなされます。予測に用いられる情報は、記録（数値のみ）以外にも力学的なもの、生理学的なものと同様に予測対象によって多様で豊富ですが、統計的予測では数値しか使いません。その統計的予測の中でさえ、同じデータに基づいても違う統計手法を用いれば厳密には一致せず、異なる予測値が得られるのが普通です。私は極値理論を用いた地震規模の予測について知らないのですが、統計数理研究所では統計学を用いた地震の研究をしています。

4. 地震予測、温暖化の到達温度予測にも活用できそうでしょうか？

地震予測が、いつ、どこで、どの程度の規模の地震が起こるかを予測する意味なら、残念ながら極値統計でそのようなことを実現するのは難しいと思います。しかし、極値統計で、長い目を見て（非常に長期）、どれくらいの規模の地震がどれくらいの数が起こるだろうという予測をすることは可能ではないかと思。ですから、洪水対策同様、巨大地震の発生頻度を予測により、発生時のインフラ整備等には役立つと思。私はこの方面に詳しくはないのですが、この方面の研究は既にいくつかは行われています。

地球上の最高気温の記録は摂氏56.7度（アメリカ デスバレー 1913.7.10）とされていますが、このような昔の局所的な記録と近年のグローバルな地球温暖化の到達温度とは別の話だと思います。温暖化の到達温度予測は時間の流れの中で統計的には必ずしも急とはいえない変化を見るものではないかと思。極値解析よりも時系列解析など他の統計手法の方が適して、実際に用いられていると思。

5. 大水害の危険を4000年に一回レベルにまで下げる災害対応とリスク評価ができることは素晴らしいと思いますが、記録上まだ一回も起きていないかもしれない災害（数万年に一回、数億年に一回？）というものがあるとすれば、化石など以外にほとんどデータがなく、解析が難しい感もあります。この考え方の路線が正しいとすれば、極値理論にもとづく災害対応について一定の適用可能範囲があるようにも思いましたが、これについて何か定量的な基準などがありますでしょうか。

質問された方も難しいと感じられている通り、極値統計における予測は、これまで起こったことがないものを含む、珍しい出来事（rare event）を少ないデータから予測することになるので、統計予測の中でも難しく、その有効性や正確性には限界があります。さらに、極値理論が用いられるような頻度の少ない現象のリスク管理は、長期的視点から行われます。しかし、長い時間の中に、当初とは状況が変わってしまうことも珍しくはありません。実際、講演でお話したオランダのデルタ計画もリスクを大幅に下げたとされていますが、近年の地球温暖化により、改めて対策を立て直さなければならない状況になっているようです。

“定量的な基準”はわかりませんが、時間経過による状況の変化に応じてリスク管理の方法も柔軟に変えることが肝要と思います。

6. 【その他】「想定外」とは極値の世界何ですか！ 目から鱗です。

「想定外」という言葉は発言者によっては、責任逃れにとられるせいか、最近では聞かれなくなっていると感じます。甚大災害についていえば、それ自体が想定外といってしまうものですから、その対策は従来の想定外を想定内にしなければなりません。極値理論はそのようなものです。

以下は言葉の問題ですが、人間の想定内は、自分や周りを経験したものになりがちですから、「これまでに経験したことがない」は「想定外」の言い換えといってもいいかもしれません。