

## 地震前兆宏観現象の統計的検定に関するいくつかの問題 On Statistical Hypothesis Testing of Earthquake Precursory Phenomena

○加藤 護

○Mamoru KATO

We discuss that improper statistical hypothesis testing methods are often used in studies of earthquake precursory phenomena. Specifically, we present cases of data snooping, which is researchers decide to perform statistical inference after looking at the data, and multiple inference, which is statistical testing without adjusting the Type I error rate accordingly.

### 1. はじめに

平成 25 年 5 月 28 日に公表された南海トラフ沿いの大規模地震の予測可能性に関する調査部会の報告書は「現在の科学的知見からは地震の規模や発生時期を高い確度で予測することは困難である」と結論した。これは地震学コミュニティのコンセンサスに近いと考えられ、この報告書は海溝型地震の対策を進める際に直前予測がないことを想定する必要性をあらためて認識する契機となった。

本報告書では地震の予測に資する可能性がある現象の例がいくつか紹介されている。過去の海溝型地震の前兆現象と考えられている観測事例に加え地震と関連する物理機構が定量的に解明されていない現象（以下、宏観現象と呼ぶ）も含まれている。例えば地球電磁気学的な観測結果と地震との関連を統計的に検証した研究例が紹介されている。

本発表では宏観現象の統計的検証手法について考察する。過去に報告された宏観現象の研究事例を取り上げ、妥当とはいえない統計的検定手法が使われた例があることを紹介する。これは地震との関係が統計的に有意であるとの結論された報告の中に再検討が必要であるものが含まれていることを意味する。なお本発表で取り上げる統計的検証手法の問題は地球電磁気学的以外の宏観現象を扱うときにも用いられており、本発表の考察はこれらの事例についても適用されるであろう。

### 2. 宏観現象の閾値の設定

宏観現象には一般にその発生時期や信号の大きさについての理論的な根拠がなく、地震との関係

では閾値が探索的に考察される。閾値の設定には過去のデータの検証が必要となるが、結果を知ってから仮説を設定する統計的検証では Love and Thomas (2013) が指摘するように data snooping と呼ばれる統計的な誤謬を避ける必要がある。具体的には仮説の設定とその検証において同じデータを繰り返し用いてはいけない。前述の報告書で紹介された事例では神津島における電磁気ノイズを検討した Orihara et al. (2010) がこのケースに該当すると考えられ、彼らの結論は再検証が必要であろう。

### 3. 多重検定に伴う補正

統計的検定で用いる有意水準は帰無仮説が正しいにも関わらず仮説検定で棄却してしまう確率（第一種の誤りを犯す確率）だが、同じ値を設定しても同時に多くの統計的検定を行うと第一種の誤りを犯す可能性は大きくなる。統計的に有意であることは観測事実が偶然とは考えにくいことを意味するが試行回数を増やすと偶然有意水準より小さな  $p$  値を得る期待値がそれに伴って大きくなる。Love and Thomas (2013) による太陽活動と地震との関係の検討では小さな  $p$  値が得られる閾値の組み合わせがあるが、複数の検定を同時に行う際の補正を行うと統計的に有意ではないと判定できる。この補正を行っていないと考えられる事例には Nagao et al. (2011) が含まれる。

宏観現象の研究事例には多重検定を行ったことが陽に示されていないものもある。閾値の設定の理論的根拠が明示されていない場合統計的に有意になる閾値を探索的に探した可能性を念頭に置き、その結論の採用を保留することが賢明であろう。