

ASA 声明と これからの統計学の使われ方

最近の心理統計分野の動向から

岡田謙介

専修大学人間科学部 准教授

1 ASA声明

2016年3月,データを扱う研究業界にとって,ひとつの「事件」があった。統計学研究の世界的な拠点団体であるアメリカ統計学会(American Statistical Association, ASA)が,学会の名のもとに「*p*値についてのASA声明」を,その機関誌 *American Statistician* に発表したのである(Wasserstein and Lazar, 2016)。この声明は,「統計的に有意」であることの本来の意味を確認し,その拡大解釈や誤用に警鐘を鳴らすものであった。

統計学的有意性

統計学において,有意である(significant)とは,統計的仮説検定を行って*p*値が事前に定めた有意水準よりも小さくなることを指す。仮説検定を行う応用研究では,慣習的に有意水準 α は $\alpha = 0.05$,つまり5%に設定されることが多い。このとき,有意であるとは,仮説検定から $p < .05$ である結果が得られたことを意味する。

20世紀前半にFisherやNeyman, Pearsonらによって仮説検定という方法が確立されて以来,結果が「統計学的に有意」であることは,実証科学のさまざまな分野において,研究の価値を左右する重要な指標と見なされてきた。しかし,後述する再現性の危機や*p*-hackingの問題など,この慣習に伴う弊害が近年とくに顕在化してきたことから,ASA自らが*p*値の本来の意味を確認し,その誤用や拡大解釈をいましめたのがこの声明である。

ASA声明の発表時,私はこの声明を主導してとりまとめた,当時のASA会長であるJessica Utts教授の所属するカリフォルニア大学アーバイン校で在外研究を行っていた。同大学ではとくに私の専門分野である,心理学での応用を念頭において統計改革やベイズ統計学の研究を主導する研究者と日常的に交流する機会に恵まれていた。そのため,ASA声明のニュースもいち早く耳に入り,同僚たちといよいよ統計学の使われ方も潮目が変わるかもしれない,と感慨深く話したことが印象に残っている。

ASA声明の概要

ASA声明は,本来的な意味で*p*値が何を表す量なのかを確認し,それを超えた拡大解釈・拡大利用に対して強く注意を喚起するものである。その本文は2ページに満たない分量であり,ぜひ読者にも一読をお勧めしたいが,本稿ではかいつまんでその内容を紹介する¹⁾。

ASA声明はまず,*p*値が何であるのかを確認する。すなわち,統計的仮説検定における*p*値は,帰無仮説のもとで,検定統計量が今回のデータから得られた観測データ以上に極端な,つまり帰無仮説と整合的でない方向のものになる確率である。また,Neyman-Pearson流の立場にたてば,*p*値はデータの,帰無仮説からの逸脱の程度を反映すると解釈できる。

同声明は次に,*p*値が何でないのかを確認する。まず,*p*値は帰無仮説が正しい確率ではない。この解釈は*p*値についてのよくある誤りである。また「統計的に有意」とは検定の手続きにした



がうと帰無仮説が棄却されることを指すのであり、ビジネスや政策のような実社会において「意味がある」ことを意味するわけではない。さらに、 p 値は効果の大きさや、モデルや仮説についての証拠の大きさを表すわけでもない。

そして声明は、単に p 値の限界を述べるだけに留まらず、 p 値に代わる、もしくは p 値を補完しうる方法について述べる。声明で挙げられているのは、(1) 検定よりも推定を重視する信頼区間や信用区間、予測区間に基づく方法、(2) ベイズ統計学の方法、(3) 尤度比やベイズファクター、そして(4) 統計的決定理論など異なる観点からのアプローチ、の4点である。

ASA声明のもつ意義

実のところ、ASA声明で述べられている p 値の性質やそれに代わる方法は、個々について言えば特段新しいものではない。むしろ、ASA声明に付随するプレスリリースでUtts会長が述べているとおり、統計学者や科学者たちはこうした問題について、何十年にもわたって警鐘を鳴らし続けてきた。ASA声明は本文わずか2ページほどの短いものであるが、その末尾には、代表的な既存の研究の系譜が参考文献リストという形でつけられている。関連して、私たちが以前、この声明に書かれたような内容の背景を掘り下げ、検定の本来の意味と誤用や、 p 値を代替・補完する方法についての和書を上梓している(大久保・岡田, 2012)。同書ではとくに私の専門である心理統計分野での動向を追ったが、同様の議論は医学や生態学など、他分野での統計学研究者からも指摘されていた。

内容的な新規性には乏しいにもかかわらず、ASA声明は大きなインパクトをもって受け止められた。*Nature*誌をはじめとする各分野の論文誌や、多くのメディアが声明をとりあげた。ASA声明のどこがそんなに衝撃的だったのだろうか？ ひとつには、世界でも最も中心的な統計学の研究者団体が声をあげた点にあると

思われる。統計学の論文誌では、現在でも新しい検定が日々開発され、論文として出版されている。新たな問題に対して検定を構成することは、統計学者の重要な仕事のひとつであり続けてきた。それにもかかわらず、統計学の学会が、検定の濫用・誤用への危機感から、まかり間違えば自らの存在意義を低下させかねないような声明を出したのである。しかも、このように重要な統計学の問題について、学会としての推奨事項を明示したのは、ASAの177年の歴史の中で初めてのことだったという(Baker, 2016)。

今後はどうなるか？

ASA声明が出てから1年後の2017年3月、ASAとイギリス王立統計学会(Royal Statistical Society, RSS)が共同出版している*Significance*誌にASA声明のその後についての記事が掲載された(Matthews, 2017)。そこでは、*Lancet*や*PNAS*(米国科学アカデミー紀要)といった一流誌の最新号を見る限り、何事もなかったかのように p 値を報告している論文が依然として多いことが述べられている。もっとも、研究を計画し、実行して統計的データ分析を行い、論文が査読を経て論文誌に掲載されるまでには年単位の時間がかかるのが一般的であろう。そう考えると、ASA声明の実際的な影響を評価するにはやや時期尚早かもしれない。

また、ASA声明の直接の効果とは言えないまでも、実証研究における統計学の使われ方に近年さまざまな変化が見られているのは事実である。本稿では以下、ASA声明でもとりあげられていた、ベイズ統計学的方法と、研究の透明性を高めるためのオープンサイエンスというふたつのテーマについて、とくに心理統計分野の動向を見ていきたい。

2 ベイズ統計学的アプローチ

従来の p 値の役割を補完または代替する方法としてASA声明が挙げたものの多くは、ベイズ統計学の枠組みと関連している。ベイズ統計学は、未知パラメータを確率変数として、既知の定数であるデータから推定する統計学の立場である。従来の伝統的な統計学（頻度論）に基づく検定では、パラメータを未知の定数として、データを母集団から確率的に得られる確率変数として扱っていた。つまり、ベイズ統計学と従来の統計学とでは、確率変数として扱うものが逆の関係になっている。

ベイズ統計学の設定は、データ分析に際して通常研究者が持つ、直感的な考えに合致しているといわれる。たとえば、頻度論の枠組みでは、パラメータについての仮説である「帰無仮説が正しい確率」を考えることはできないが、ベイズ統計学の枠組みではそれが可能である。また、頻度論の p 値がサンプリング・デザインに依存する問題に対して、ベイズ統計学の枠組みではこれを回避しうる (Rouder, 2014)。

こうしたことは、 p 値の拡大解釈に対する批判と同様に、ベイズ統計学の利点として長い間統計学者たちが論じてきたことである。一方で、こうした理由もさることながら、近年のベイズ統計学の隆盛を支えているのは、統計モデリングという考え方の浸透と、汎用的な推定アルゴリズムを搭載したソフトウェアの普及という2点が大きな要因となっていると考えられる。

統計モデリングの考え方

従来の統計学の応用では、2条件間での差を調べたいのであれば t 検定を行う、といった具合に、データの形式的設定に応じた形で「よく使われる分析」があった。そうした分析において、統計モデルとは、データの生成過程を表現するというよりは、差の有無のような調べたいこ

とを調べるための便宜的な仮定、という位置づけであることが多かった。

これに対して、統計モデリングの考え方では、分析者の手元にあるデータがどのようなメカニズムによって得られたものなのかを、明示的に統計モデルによって表現しようとする。これによって、データに適用される統計モデルは、各データの文脈や特徴を反映した、個別性や複雑性の高いものであることが必然的に多くなる。よい統計モデリングは、データが生成される過程についての理論を検証し、そこに量的な知見を提供したり、将来の量的な予測を行ったりすることを可能にする。こうした方法論の具体的な考え方や例は、和書では久保 (2012) や松浦 (2016) などに詳しい。

推定法とソフトウェア

ベイズ統計モデルの汎用的な推定法として、マルコフ連鎖モンテカルロ (Markov chain Monte Carlo, MCMC) 法と呼ばれる一群の方法がある。MCMC法を用いれば、ベイズ統計で目的とする分布（事後分布）からの乱数を多数とりだすことができ、これを用いて目的の推論や予測を行うことができる。MCMC法は計算資源を要求する数値的な方法であるが、コンピュータの普及と相まって、1990年代以降、ベイズ統計学を一般の研究者にとって実用的なものにすることに大きく貢献した。

MCMC法を実装した最初の汎用的なソフトウェアとして、WinBUGSはよく知られている。また最近では、JAGSやStanといった、より新しいMCMC法を用いたベイズ推定のソフトウェアが多くの研究者に利用されている。とくにStanは、著名なベイズ統計学者である、コロンビア大学のAndrew Gelman教授のチームが、ハミルトニアン・モンテカルロ法という新しく高性能なMCMCアルゴリズム (和書では豊田, 2015) を実装して活発に開発を行っている。こうしたことから、Stanは近年わが国でも、大学・企業を



問わず、研究者やデータサイエンティストの間で普及してきている。一方、アルゴリズムの特性上、離散的なパラメータを明示的にモデリングできないことや、心理学でよく利用されるような潜在変数に対する制約の導入法の一部が使えないことといった、発展的なソフトウェアであるがゆえの問題もある。その点では、ひと世代前の推定アルゴリズムを用いているが高い安定性を誇るJAGS (やWinBUGS) も、現在でもその利用価値は高い。

このように、ソフトウェア環境が整い、統計モデリングの利点が広く意識されてきたことで、ベイズ統計学は実用的なものになってきた。たとえばここ1,2年のうちに、*Journal of Mathematical Psychology* 誌、*Psychological Methods* 誌、*Psychonomic Bulletin & Review* 誌、*European Journal of Psychotraumatology* 誌といった心理学の論文誌がベイズ統計の特集号を組んでいる。

3 オープンサイエンス

ASA声明では、研究の透明性を高めることの重要性も指摘されていた。近年、心理学は再現性の危機 (reproducibility crisis) を経験した。教科書に載るような心理学、とくに社会心理学分野の知見の中に、再現性が担保できないものが少なくないことが多くの研究者から指摘された。その原因が調査され、 p 値だけが過度に重視されたために生じた p -hackingのような不適切な研究慣習や、ひどい場合には研究捏造があったことも明らかになった。このあたりの事情について、和文では『心理学評論』誌における心理学の再現可能性特集号が多数の論者を掲載している (友永・三浦・針生, 2016)。

こうした経緯を受けて、心理学でもオープンサイエンスの機運が高まり、データや研究プロトコルの公開が急速に普及しつつある。インターネットのなかった時代には、生データや調査票などを論文誌の誌面にすべて掲載することは、

物理的にも経済的にも不可能であった。しかし現在、インターネット時代の我々は、こうしたデータや研究素材などを簡単な操作でWebにアップロードし、公表することができる。

たとえば、心理学のトップジャーナルである *Psychological Science* 誌は、「オープンデータ」「オープンマテリアル」「事前登録」の3点について、各々を達成した論文に対してオープンプラクティス・バッジと呼ばれるカラーの目立つアイコンをつけて、これを奨励する取り組みを2014年から始めている。

オープンデータとオープンマテリアルは、読んで字のごとく、研究で収集したローデータと、調査票や実験刺激といった研究素材 (マテリアル) を公開するものである。また研究の事前登録 (pre-registration) とは、まず研究デザインやサンプルサイズを外部の機関やWebサイトに登録し、その登録内容に沿った形で以降の研究、つまりデータ収集や統計分析を行うものである。サンプルサイズを決めるためには、統計学的な例数設計が重要となる。このような手順を踏むことで、有意になるまでデータをとり続けたり、有意になった結果だけを発表するような問題のある研究報告を減らすことができると考えられている。

データや研究素材の公開、研究の事前登録のために研究者が無償で使えるサービスには、すでに多くのものである。代表的なものひとつに Open Science Framework (Open Science Framework, 2017) があり、私もこれを利用している (図)。同Webサイトの事前登録のページ²⁾を見ると、毎日いくつもの研究が事前登録されるようになっていくことが実感できるだろう。

4 むすびに

本稿では昨年出されたASA声明を題材として、仮説検定や p 値と、ベイズ統計学・心理統計学にまつわる最近の話題を概観した。

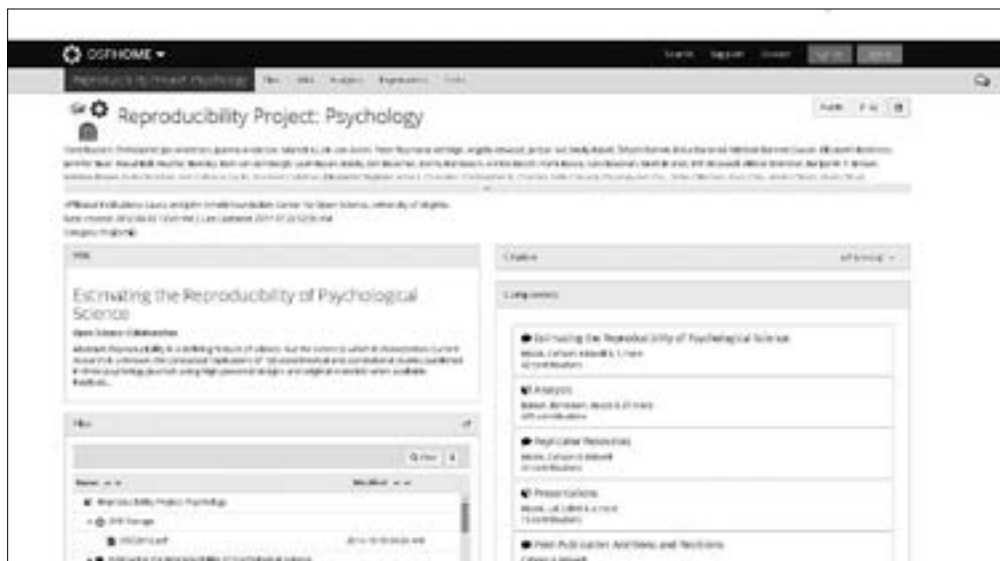


図 Open Science Frameworkを著名にした、心理学の論文誌3誌に2008年以降刊行された100の研究を世界各国の270の研究者が追試したプロジェクトのページ(<https://osf.io/ezcuj/>)

私たちが最近発表した研究 (Okada et al., 2017) では、心理学のオープンデータを用いて、オンライン調査の回答行動のベイズ統計モデリングを行った。1次研究にはなかった統計モデリングの視点を導入して既存のデータを再活用することにより、調査回答行動のメカニズムをよりよく理解し、また予測できることを同論文では示した。このような統計モデリングやオープンサイエンスの動向を背景とした研究は、今後とも増加が予想される。

統計学の応用は、従来の p 値一辺倒であった時代から、モデリングの方向へ、そしてオープン化の方向へと変化しつつある。従来型の研究では、たとえば2変数間の差や相関といった、仮説検定を使って評価できる形で研究計画や研究仮説が準備されることが少なかった。しかし現代では、関心のある現象それ自体に注目したモデリングや予測を行うことが可能である。オープンサイエンスの輪が広がれば、単一の研究では検証しえないことを研究コミュニティの力で解明することが可能になるだろう (たと

えばOpen Science Collaboration, 2015)。

統計学の応用は大きく変わりつつある。こうした新しい動向を、関心下の現象の解明・理解という研究コミュニティが共有する目的へ向かって役立てていきたい。

【謝辞】 本稿の内容は2016年度統計関連学会連合大会における企画セッションでの研究発表「心理学におけるオープンサイエンスの進展と、統計学にできること」に基づき、これを再構成したものである。本研究の実施にあたってはJSPS 科研費17H04787の助成を受けた。

注

- 1) 本稿の締切日当日に、ASA 声明の和訳が日本計量生物学会のホームページにて公開された (<http://biometrics.gr.jp/news/all/ASA.pdf>)。原文と併せて参照いただきたい。
- 2) <https://osf.io/activity/#newPublicRegistrations>



文献

- Baker, M., 2016, "Statistician Issue Warning over Misuse of P values", *Nature*, 531: 151.
- 久保拓弥, 2012, 『データ解析のための統計モデリング入門——一般化線形モデル・階層ベイズモデル・MCMC』岩波書店。
- 松浦健太郎, 2016, 『StanとRでベイズ統計モデリング』共立出版。
- Matthews, R., 2017, "The ASA's p -value Statement, One Year on", *Significance*, 14 (2): 38-41.
- Okada, K., Vandekerckhove, J. and Lee, M.D., 2017, "Modeling When People Quit: Bayesian Censored Geometric Models with Hierarchical and Latent-Mixture Extensions", *Behavior Research Methods*, Online Ahead of Print, doi:10.3758/s13428-017-0879-5.
- 大久保街亜・岡田謙介, 2012, 『伝えるための心理統計——効果量・信頼区間・検定力』勁草書房。
- Open Science Collaboration, 2015, "Estimating the Reproducibility of Psychological Science", *Science*, 349, aac4716.
- Open Science Framework, 2017, "Resource Review: Open Science Framework", *Journal of Medical Library Association*, 105 (2): 205-206.
- Rouder, J. N., 2014, "Optional Stopping: No Problem for Bayesians", *Psychonomic Bulletin & Review*, 21: 301-308.
- 友永雅己・三浦麻子・針生悦子, 2016, 「心理学の再現可能性——我々はどこから来たのか 我々は何者か 我々はどこへ行くのか」『心理学評論』59 (1) : 1-2。
- 豊田秀樹, 2015, 『基礎からのベイズ統計学——ハミルトニアンモンテカルロ法による実践的入門』朝倉書店。
- Wasserstein, R. L. and Lazar, N. A., 2016, "The ASA's Statement on p -values: Context, Process, and Purpose", *American Statistician*, 70 (2): 129-133.