

心理学雑誌に掲載された実験の追試を行ったところ、有意な結果が出たのは36%であった¹。これは驚くべきことなのか、いつも通りなのか。心理学者に「その心理学信じていいですか?」と尋ねたら、どう反応するだろう。物理学者に同様の質問をすると憤慨してしまいそうなこの問いは、68号(2015年)の特集タイトルである。この特集では、「疑わしい研究活動(QRPs)」が話題となったことを受けて、QRPsを紹介するとともに、「信じられる」研究にするための方略が紹介されている。ここでは特集のなかから2本の記事を紹介する。

まず、平石界氏・池田功毅氏「心理学な心理学研究 — Questionable Research Practice」では、心理学の学部生が指導教員と大学院生に研究方法について相談する様子を通じて、QRPsの実情が暴かれている。研究室でよく耳にしそうなやり取りだが、この対話に疑問を抱かず読み進めてしまうと、多重検定、確証バイアス、後知恵バイアスなどの罠にはまることになる。QRPsに陥りやすいか否かを見極める試金石のような記事である。読者には自身がQRPsに陥りやすいかどうかをぜひ試していただきたい。また、業績競争における研究者のジレンマや、QRPsに陥った研究者に必ずしも悪意があるわけではないことなど、昨今の研究環境を諷刺的に描いている。記事では言及されていないが、研究者に悪意がないのは無知ゆえであるから、改善には少なくとも適切な統計教育が必要であろう。

次に、大久保街亜氏「統計的に有意? — 帰無仮説検定でわかること・わからないこと」では、 p 値のみに基づく統計的帰無仮説検定の限界が指摘されたとうえで、効果量、信頼区間、検定力など複数の指標から包括的に検討することが勧められる。一時物議を醸したFacebookの感情伝染実験を具体例としてあげ、読みやすく工夫されている。また、ノウハウだけが流布し

て統計解析の過程がブラックボックス化されたり、有意水準0.05という基準が慣例化されたりして、研究者の思考を停止させて怠惰な精神をつくることも指摘されている。ただし、有意水準0.05には明確な根拠がないと批判的に述べられているが、効果量や検定力も同様であることに注意が必要である。統計学に頼らざるを得ないのは「真理」がわからないからであるが、統計解析法は、すべての前提が正しくても演繹のように100%正しい結論を導き出すことはできず、確率的に結論を評価するしかない点は肝に銘じておくべきである。それと合わせて、QRPsでは前提が正しいとは限らない点も忘れてはならない。また、統計的帰無仮説検定の代わりにベイズ推定を推しているが、代替案にはベイズ推定だけでなく、情報量規準によるモデル選択など多くの選択肢がある。データによる統計モデルの選択や評価として、目的に合わせた統計解析法を選ぶように促すほうがよいだろう。

領域から 科学哲学の

北海道医療大学
リハビリテーション科学部 准教授

森元良太

現代の統計的帰無仮説検定の源流である有意性検定は、約100年前にロナルド・フィッシャーが考案した。フィッシャーは当初から多くの誤用や誤解に注意を促してきたが、残念ながら研究者はいまだに誤用や誤解を繰り返している。もちろんフィッシャーだけでなく、その後も誤用や誤解への注意は幾度となく繰返し促されてきた^{2,3}。歴史を振り返ると、研究者のQRPsは常態化しているようである。心理学は行動変容に大いに役立ってきたので、心理学者には研究者がQRPsに陥らないよう行動変容させるプログラムを開発していただきたい。



もりもと・りょうた 博士(哲学)。専門は科学哲学(生物学の哲学、確率論・統計学の哲学)。共著に『生物学の哲学入門』、分担執筆に『ダーウィンと進化論の哲学』(ともに勁草書房)など。

¹ Open Science Collaboration (2015) *Science*, 349, aac4716. ² Morrison, D. E., & Henkel, R. E. (1970) *The significant test controversy*. New Brunswick. [D. E. モリソン・R. E. ヘンケル/内海廉一郎他(訳)(1980)統計的検定は有効か. 粹出版社.] ³ Gigerenzer, G. (2004) *The Journal of Socio-Economics*, 33, 587-606.