

ホットな研究テーマほど「偶然の有意差」が論文になりやすい？

著者：関勝寿 公開日：2026年9月19日 ソース：<https://sekika.github.io/2026/09/19/publication-bias/>

ある研究者が、「Aという現象が起きるのではないかと予想して実験をしたとする。帰無仮説を「Aという現象はない」として、有意水準5%で検定したところ、 $p \geq 0.05$ で帰無仮説を棄却できなかった。そこで「予想を支持する結果は得られなかった」と考え、論文にはしなかったとする。しばらくして、別の研究者が同じことを思い付き、同じような実験をした。この人も有意な結果を得られず、論文にはしなかった。このようなことが何度も繰り返されたら、どうなるだろうか。

仮に、実際にはAという現象が存在しないとする。それでも、有意水準5%で検定していれば、帰無仮説が正しいにもかかわらず、偶然 $p < 0.05$ となることがある。これが第一種過誤である。1回の実験でその確率が5%なら、独立した実験を20回行ったとき、少なくとも1回は偶然 $p < 0.05$ となる確率は $1 - 0.95^{20} \simeq 0.642$ となり、約64%である。50回なら約92%、100回なら約99.4%になる。

もちろん、一人の研究者が100回実験して有意だった結果だけを論文にすれば問題である。しかし、もっと厄介なのは、100人の研究者がそれぞれ独立に1回ずつ、適切な方法で実験した場合である。誰も不正をしていないし、誰も「有意になるまで実験を繰り返した」わけでもない。それぞれが有意水準5%を守って正しく検定している。それでも、本当はAという現象が存在しなければ、平均すれば100件のうち5件程度は偶然 $p < 0.05$ になる。ここで、有意な結果が出なかった95人が「予想した結果にならなかったの、論文にするほどの成果ではない」と考えて投稿しなかったとする。一方、偶然 $p < 0.05$ になった5人は「Aという現象を支持する結果が得られた」として論文を投稿する。すると、後から文献を調べた人に見えるのは「Aという現象を支持する論文が5本ある」という事実だけである。しかし、本当に知りたいのは「Aを調べた研究が100件あり、そのうち5件で有意、95件で非有意だった」という情報である。前者と後者では意味がまったく違う。論文にならず研究者の「引き出し」にしまわれた研究が見えないために生じるこの問題は、**file-drawer problem**と呼ばれ、出版バイアス (publication bias) の代表的な形の一つである。

ここから、研究テーマの注目度との関係も見えてくる。同じアイデアを思いつく研究者がほとんどいないテーマなら、この問題はそれほど大きくならない。しかし、非常に注目されているホットなテーマなら、世界中の研究者が「もしかするとAなのではないか」と考え、それぞれ独立に研究を始めるかもしれない。各研究者は1回しか検定していなくても、研究コミュニティ全体では同じような仮説が何十回、何百回と検定されていることになる。これは、ある意味では研究コミュニティ全体で行われる「見えない多重検定」と考えることができる。しかも、普通に論文を読んでいるだけでは、その「検定回数」が分からない。100件の研究が行われて5件だけ出版されたとしても、読者に見えるのは5件だけだからである。そのため、他の条件が同じなら、多くの研究者が同じアイデアを試すホットなテーマほど、偶然生じた有意な結果がどこかで出現する可能性は高くなる。そして非有意な結果が出版されにくければ、それが文献に残り、「Aを支持する研究がいくつもある」という外観を作る可能性がある。

この問題の重要なところは、個々の研究だけを点検しても発見できない場合があることである。実験計画にも問題がなく、統計解析にも問題がなく、研究者は有意水準を事前に5%と決め、その通りに解析し、**p-hacking**もしていない。それでも、「どの研究が論文として表に出るか」という選択が加わるだけで、公開された科学文献全体には偏りが生じる。個々の論文に書かれた $p < 0.05$ だけを見るのではなく、「この結果が表に出てくるまでに、世界中で同じ問いが何回試されたのだろうか」と考えることも、研究結果を読むうえで重要な視点である。