

伝えるための心理統計：効果量・信頼区間・検定力

大久保 街亜・岡田 謙介 (2012)

勁草書房.

p 値が0.05 よりも小さくなってくれば、論文が書ける。0.05 より小さくならなければ、論文が書けず、せっかくとったデータが「お蔵入り」になってしまう——。

これってなにかがおかしい気がする。そう感じたことのある方は、おそらく多くいらっしゃるのではないのでしょうか。

これは、本稿で紹介する標題の書籍「あとがき」の冒頭部である。研究者で集まって話をする時、このような問題意識を持っている人は実に多いことがわかる。実際のところ、統計的仮説検定におけるp値は標本サイズの関数になっており、理論上は標本サイズさえ十分に大きくとることができれば、どんなに小さな差でも（それが0でない限りは）有意という結果を出すことができる。しかし、そのように標本サイズにまかせて、力技で出した「有意」に実際に心理学的な意味があるとは考えにくいだろう。また、1つの実験にたくさんの時間がかかったり、なんらかの特徴を持つ集団に関心がある場合など標本サイズを増やすことが難しい分野では、標本サイズが小さいために有意差を出すことができず、貴重な結果が公表されなくなってしまっていることも考えられる。

したがって、統計的仮説検定でp値が0.05よりも小さいかどうかだけによって、結果の意義が、そして結果が論文として出版されるかどうかが決まってしまうような風潮は非常に危険である。しかしながら、こうした問題意識は多くの研究者に共有されているにもかかわらず、それではどうすればよいのかについて、正面から論じられることはわが国では少なかった。

本書は、この問題に対して正面から向きあった書籍である。上記の問題意識は世界的にみても多くの研究者に共有されており、そうした研究者達によって、統計的データ分析をより実質的で実践的なものにしようという取り組みがなされてきた。そのまとめ役のひとつとして、米国心理学会（American Psychological Association, APA）が機能してきており、こうした流れはいわゆる心理学における統計改革（statistical reform）と呼ばれる。本書は、こうした統計改革の考え方や流れを概説し、とくに効果量・信頼区間・検定力を用いて仮説検定を補完した代替する方法論について解説している。また、より発展的な内容として、メタ分析・ベイズ統計学・ p_{rep} という3種類の話題にもそれぞれ1節を割り当てて扱っている。書籍中で出された数値例については、付録にRを用いた計算プログラムが掲載されており、単に頭で理解するのみならず実際に手を動かして計算をすることが可能である。本書が心理学研究者の日々の研究に役立ち、よりよい心理学研究のために少しでも貢献できることを願ってやまない。