心理学の研究における効果量と p 値の乖離○波田野 結花¹・岡田 謙介²(¹ 専修大学大学院文学研究科・² 専修大学人間科学部)キーワード: 効果量, p 値, 統計的仮説検定

The Gap between Effect Size and P-value in Psychological Research

Yuka HATANO¹ and Kensuke OKADA²(¹ Graduate School of the Humanities, Senshu Univ., ² School of Human Sciences, Senshu Univ.)

Key words: Effect Size, P-value, Statistical Hypothesis Testing

問題と目的

わが国の論文誌に掲載されている研究においては、統計的仮説検定の結果は通常報告されるが、効果量に関する議論は軽視されがちである。しかし、効果量に基づく結果と統計的仮説検定の結果の間には乖離が生じている可能性がある。そこで本研究では、『教育心理学研究』誌に掲載された論文の統計的仮説検定について実際に標本効果量を算出し、算出した標本効果量と p 値との関係を、相関係数および散布図によって調べることで、これらの間にどの程度乖離が生じているのかを明らかにすることを目的とする。

方 法

分析対象 論文検索用データベース CiNii を用いて、本文が公開されていた 2009 年度, 2010 年度の『教育心理学研究』(Vol. 57, No. 1 から Vol. 58, No. 4 まで) 誌に掲載された 33 編の論文を収集した。この中で報告された統計分析のうち、対応のある t 検定 (7 編の 45 の検定), 対応のない t 検定 (13 編の 94 の検定), 1 要因被験者間分散分析 (12 編の 171 の検定), 2 要因被験者間分散分析 (15 編の 112 の検定) という 4 つの仮説検定を分析の対象とした。

分析方法 上記の論文の中で報告された検定統計量とサンプルサイズから、大久保・岡田 (2012) に従って標本効果量を算出した。同様に、論文中の検定統計量とサンプルサイズから p 値を算出し、算出した標本効果量と p 値との相関係数および散布図を求めた。

結 果

例として、1 要因被験者間分散分析における標本効果量 η^2 と p 値との関係を散布図に示した (Figure 1)。その結果、2 変数間には直線的ではない関係が見られた。また、標本効果量 η^2 と p 値との相関係数を算出したところ、中程度の負の相関が見られた ($r = -.471, N = 171, SE = .060$)。同様に、対応のある t 検定, 対応のない t 検定, 2 要因被験者間分散分析においても標本効果量と p 値の間には直線的ではない関係が見られ、相関係数も全て中程度の負の相関があった (Table1)。

Figure 1 1 要因被験者間分散分析における標本効果量 η^2 と p 値との関係, $N = 171$

Table 1 4 つの仮説検定における標本効果量と p 値との相関係数及び標準誤差

仮説検定	相関係数	SE
対応のある t 検定 ($n = 45$)	-.523	.108
対応のない t 検定 ($n = 94$)	-.536	.074
1 要因被験者間分散分析 ($n = 171$)	-.471	.060
2 要因被験者間 (要因 A の主効果)($n = 112$)	-.466	.074
2 要因被験者間 (要因 B の主効果)($n = 112$)	-.600	.060
2 要因被験者間 (交互作用)($n = 112$)	-.424	.078

考 察

4 つの仮説検定において、標本効果量と p 値との散布図および相関係数から、標本効果量が大きくなるほど p 値は小さくなるという関係が見られた。ただし、2 変数間の関係は直線的ではないため、相関係数は必ずしも 2 変数間の関係を適切に表す指標ではないことには留意する必要がある。

しかしそれを踏まえても、出版された論文中の標本効果量の大きさと p 値の間にはある程度の乖離がみられるといえる。 p 値には標本効果量の大きさ以外にも、サンプルサイズなどが影響を与えていることが考えられる。また、散布図を見ると、標本効果量と p 値の双方、もしくはどちらか一方が小さい報告は数多く見られるが、その双方が大きい報告はほとんどないことがわかる。このひとつの原因としては、出版バイアスによる影響が懸念される。

以上のことから、統計分析においては p 値だけではなく効果量とともに活用していくことが望ましいと考えられる。

引用文献

大久保衛亜・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