Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 1634–1646
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Effect magnitude: A different focus☆

Roger E. Kirk

Baylor University, Waco, TX 76798-7334, USA

Available online 9 October 2006

Abstract

The role of measures of effect magnitude in the research enterprise is examined. Measures of effect magnitude are used for four purposes: (a) to estimate the sample size required to achieve an acceptable power, (b) to integrate the results of empirical research studies in meta-analyses, (c) to supplement the information provided by null hypothesis significance tests, and (d) to determine whether research results are practically significant. The advantage of focusing on effect magnitude and practical significance instead of statistical significance and p values is illustrated with an example.

Keywords: Effect magnitude; Effect size; Meta-analysis; Null hypothesis significance testing; Practical significance

1. Introduction

This paper examines the role of effect magnitude in the research enterprise. A variety of statistics are used to measure effect magnitude. Many of the statistics fall into one of two categories: measures of effect size (typically, standardized mean differences) and measures of strength of association. In addition, there is a large group of statistics that do not fit into either category. A partial listing of effect magnitude statistics is given in Table 1.

The first measure of effect size that was explicitly labeled as such appeared in the social and behavioral sciences literature in 1969. Among researchers who work in substantive areas, interest in measures of effect magnitude has increased over the last 35 years. I have published an article in 1996 in which I identified over 40 measures of effect magnitude that are used in psychology and education journals (Kirk, 1996). Today the term "effect size" appears in the index of virtually every introductory statistics textbook written for the social and behavioral sciences.

Many statisticians do not share this interest in measures of effect magnitude. Those of a Bayesian persuasion represent the exception to this generalization. The lack of interest is apparent when we examine our premiere journals and textbooks. Consider, for example, the seventh edition of Freund's (2004) *Modern elementary statistics* and the fourth edition of Moore and McCabe's (2003) *Introduction to the practice of statistics*. These popular introductory textbooks, which are widely used in statistics departments, do not even include the term "effect size" in the index. Applied researchers in the social and behavioral sciences have a different view regarding the importance of measures effect magnitude. According to Bruce Thompson (2006), as of 2005, 24 substantive journals have adopted editorial policies that *require* the reporting of effect sizes or other measures of effect magnitude. Obviously the editors of these journals think that reporting effect sizes is important.

☆ This article is based on my keynote address at the 2003 International Conference on Statistics, Combinatorics, and Related Areas that was held at the University of Southern Maine.

E-mail address: Roger_Kirk@baylor.edu.

0378-3758/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2006.09.011

Table 1
Measures of effect magnitude

Measures of effect size	Measures of strength of association	Other measures
Cohen's (1988) d, f, g, h, q, w	$r, r_{pb}, r_s, r^2, R, R^2, \eta, \eta^2, \hat{\eta}^2_{\text{mult}}, \phi, \phi^2$	Absolute risk reduction (ARR)
Glass's (1976) g'	Chambers' (1982) r_e	Cliff's (1993) p
Hedges's (1981) g	Cohen's (1988) f^2, κ	Cohen's (1988) U_1,U_2, U_3
Mahalanobis's D	Contingency coefficient (C)	Doksum's (1977) shift function
$\text{Mean}_1 - \text{mean}_2$	Cramér's (1946) V	Dunlap's (1994) common language effect size for bivariate correlation (CL_R)
$\text{Median}_1 - \text{median}_2$		
$\text{Mode}_1 - \text{mode}_2$		
Rosenthal and Rubin's (1989) Π	Fisher's (1921) Z	Grissom's (1994) probability of superiority (PS)
Tang's (1938) ϕ	Friedman's (1968) r_m	Logit d'
Thompson's (2002) d^*	Goodman and Kruskal's (1954) λ, γ	McGraw and Wong's (1992) common language effect size (CL)
Wilcox's (1996) $\hat{\Lambda}_{Mdn,\sigma_b}$	Hays' (1963) $\hat{\omega}^2, \hat{\omega}^2_{Y\|A\cdot B, AB}, \hat{\rho}_\text{I}, \hat{\rho}_{\text{I}Y\|A\cdot B, AB}$	Odds ratio ($\hat{\omega}$)
Wilcox and Muska's (1999) $\hat{Q}_{0.632}$	Herzberg's (1969) R^2	Preece's (1983) ratio of success rates
	Kelley's (1935) ε^2	Probit d'
	Kendall's (1963) W	Relative risk (RR)
	Lord's (1950) R^2	Risk difference
	Olejnik and Algina's (2003) $\hat{\omega}^2_\text{G}, \hat{\eta}^2_\text{G}$	Sánchez-Meca et al.'s (2003) d_Cox
	Pillai–Bartlett's V	Rosenthal and Rubin's (1982) binomial effect size display (BESD)
	Rosenthal and Rubin's (2003) $r_\text{equivalent}$	Rosenthal and Rubin's (1994) counternull value of an effect size ($\text{ES}_\text{counternull}$)
	Rosnow et al.'s (2000) $r_\text{alerting}, r_\text{contrast}, r_\text{effect size}$	Wilcox's (1996) probability of superiority ($\hat{\lambda}$)
	Roy's Θ	
	Tatsuoka's (1973) $\hat{\omega}^2_{\text{mult. } c}$	
	Wherry's (1931) R^2	