Research Article

Approximate Confidence Interval for Effect Size Base on Bootstrap Resampling Method

Suvimol Phanyaem*
Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

* Corresponding author. E-mail: suvimol.p@sci.kmutnb.ac.th
DOI: 10.14416/j.ijast.2018.08.001
Received: 27 October 2016; Accepted: 17 August 2017; Published online: 7 August 2018

Abstract
This paper presents the confidence intervals for the effect size base on bootstrap resampling method. The meta-analytic confidence interval for effect size is proposed that are easy to compute. A Monte Carlo simulation study was conducted to compare the performance of the proposed confidence intervals with the existing confidence intervals. The best confidence interval method will have a coverage probability close to 0.95. Simulation results have shown that our proposed confidence intervals perform well in terms of coverage probability and expected length.

Keywords: Effect size, Bootstrap method, Meta-analytic confidence interval

1 Introduction

Meta-analysis is a systematic technique for reviewing, analyzing, and summarizing quantitative research studies on specific topics or questions. It is usually to estimate the overall treatment effect and make inferences about the difference between the effects of the two population. There are three types of data that are commonly to meet with binary data, ordinal data and normally distributed data. In this paper the responses can be considered to be approximately normally distributed. Therefore, the commonly effect size estimator is standardized mean differences. The standardized mean difference is the mean difference between groups in standard score form such as the ratio of the difference between the treatment and control means to the standard deviation [1].

Hedges [2] suggested fixed-effects meta-analytic confidence interval for an average standardized effect size. Bond [3] described a fixed-effects meta-analytic confidence interval for an average unstandardized effect size. The Fixed-effect meta-analysis confidence intervals for standardized mean differences are based on unrealistic assumption of effect size homogeneity. Confidence intervals for standardized mean differences that do not assume homoscedasticity are given in Bonett [4]. Bonett [5] was proposed meta-analytic confidence intervals for standardized mean differences.

Efron [6] proposed a bootstrap computational technique that can be used to effective estimate the sampling distribution of a statistic.

The objective of this paper is to construct the confidence intervals for effect size by using Glass's method, Bonett's method and Bonett-Bootstrap method. Bonett-Bootstrap has been proposed to construct the confidence interval for the effect size parameter. The performance of the proposed method is investigated through a Monte Carlo simulation based on various evaluation criteria such as coverage probability and expected length. A confidence interval is better than another if its coverage probability is closer to the desired value (95%) and its expected length is short.

The organization of the paper is as follows: in Section 2, we presented *Glass's* and *Bonett's* estimator

Please cite this article as: S. Phanyaem, "Approximate confidence interval for effect size base on bootstrap resampling method," *KMUTNB Int J Appl Sci Technol*, vol. 11, no. 4, pp. 257–261, Oct.–Dec. 2018.

for estimating the confidence intervals of effect size. The proposed confidence intervals for effect size are presented in Section 3. A Monte Carlo simulation study is conducted in Section 4. Finally, some concluding remarks base on simulation is given in Section 5.

2 Confidence Intervals for Effect Size

In situation the outcome is reported on a meaningful scale and all studies in the analysis use the same scale, the meta-analysis can be performed directly on the raw difference in means. The advantage of the raw mean difference is that it is intuitively meaningful, either inherently such as blood pressure, which is measured on a known scale.

Consider a study that reports means for the treatment and control groups and suppose we wish to compare means of these two groups. Let and be the population means of treatment and control groups. The standardized difference between two population means is defined as [Equation (1)]

$$\delta = \frac{(\mu_1 - \mu_2)}{\left[(\sigma_1^2 + \sigma_2^2)/2\right]^{1/2}} \tag{1}$$

where σ_1^2 and σ_2^2 be the population variance of the treatment and control groups, respectively [7].

In study i, assume that the random sample is obtained from specific study population [Equation (2)]

$$\delta_i = \frac{(\mu_{i1} - \mu_{i2})}{\left[(\sigma_{i1}^2 + \sigma_{i2}^2)/2\right]^{1/2}}. \tag{2}$$

The fixed-effect methods assume that the m studies have been deliberately selected and that statistical inference applies only to the m study populations represented in the m studies.

The main purpose of averaging effect size estimate from several studies is to obtain an estimate of the average effect size that is more precise than an effect size estimate form a single study. The average standardized effect size is given by [Equation (3)] [2]

$$\delta = \frac{1}{m}\sum_{i=1}^{m}\delta_i. \tag{3}$$

In this section, we show how to compute an estimate $\hat{\delta}$ of this parameter and its variance from studies that used two independent groups. We can estimate the standardized mean difference δ from the study that used two independent groups as follows. Let X_1 and X_2 be the sample means of the two independent groups. And let S_1 and S_2 be the sample standard deviations of the two groups, and n_1 and n_2 be the sample sizes in the two groups.

2.1 *Glass's confidence interval*

Glass [1] proposed the confidence interval for effect size. The sample estimate of $\delta_{(i)}$ is the difference in sample means as follows [Equation (4)]

$$\hat{\delta}_{G(i)} = \frac{\bar{X}_{i1} - \bar{X}_{i2}}{S_i}. \tag{4}$$

where $\bar{X}_{i1}$ and $\bar{X}_{i2}$ are the sample mean of treatment and control group for studies i, (i = 1, 2..., m) and S_i is the sample standard deviation of control group.

An estimate of the variance of $\hat{\delta}_{G(i)}$ is given by [Equation (5)]

$$Var(\hat{\delta}_{G(i)}) = \frac{N}{n} + \frac{(\hat{\delta}_{G(i)})^2}{2N}. \tag{5}$$

The point estimator of δ is given by [Equation (6)]

$$\hat{\delta}_G = \frac{\sum_{i=1}^{m} w_i \hat{\delta}_{G(i)}}{\sum_{i=1}^{m} w_i}. \tag{6}$$

The approximate 100(1–α)% Glass's confidence interval for δ is given by [Equation (7)]

$$\hat{\delta}_G - z_{\alpha/2}\sqrt{\frac{1}{\sum_{i=1}^{m} w_i}} \le \delta \le \hat{\delta}_G + z_{\alpha/2}\sqrt{\frac{1}{\sum_{i=1}^{m} w_i}}, \tag{7}$$

where $w_i = \dfrac{1}{Var(\hat{\delta}_{G(i)})}$.

2.2 *Bonett's confidence interval*

Bonett [5] was proposed meta-analytic confidence intervals for standardized mean differences. The sample estimate of $\delta_{(i)}$ is given by [Equation (8)]

$$\hat{\delta}_{B(i)} = \frac{(\bar{X}_{i1} - \bar{X}_{i2})}{\left[(S_{i1}^2 + S_{i2}^2)/2\right]^{1/2}}. \tag{8}$$

where $\overline{X}_{i1}$ and $\overline{X}_{i2}$ are the sample mean of treatment and control group for studies i, (i = 1, 2..., m) S^2_{i1} and S^2_{i2} are the sample standard deviation of treatment and control group for studies i.

The Bonett's point estimator of δ is as follow [Equation (9)]

$$\hat{\delta}_B = \frac{\sum_{i=1}^{m} b_i \hat{\delta}_{B(i)}}{m} \quad (9)$$

where b is an approximate bias adjustment.

Let $b_i = 1 - \left[\frac{3}{4(n_{i1} + n_{i2}) - 9} \right]$

where n_{i1} and n_{i2} are sample sizes of the treatment and control group for studies i. An estimate of the variance of $\delta_{B(i)}$ is given by [Equation (10)]

$$Var(\hat{\delta}_{B(i)}) = \frac{\hat{\delta}^2_{B(i)} \left(\frac{S^4_{i1}}{df_{i1}} + \frac{S^4_{i2}}{df_{i2}} \right)}{8S^4_i} + \frac{\left(\frac{S^4_{i1}}{df_{i1}} + \frac{S^4_{i2}}{df_{i2}} \right)}{S^2_i} \quad (10)$$

where $S_i = \left(\frac{S^2_{i1} + S^2_{i2}}{2} \right)^{1/2}$

and $df_{i1} = n_{i1} - 1$, $df_{i2} = n_{i2} - 1$

The approximate 100(1–α)% Bonett's confidence interval for δ is given by [Equation (11)]

$$\left(\hat{\delta}_B \pm z_{\alpha/2} \left[\frac{\sum_{i=1}^{m} b_i^2 Var(\hat{\delta}_{B(i)})}{m^2} \right]^{1/2} \right) \quad (11)$$

where $\hat{\delta}_B$ is the Bonett's point estimator of δ and $Var(\hat{\delta}_{B(i)})$ is the estimate of variance of $\hat{\delta}_{B(i)}$ are given by (9) and (10), respectively and $z_{\alpha/2}$ is 100(1 – α/2) percentile of the standard normal distribution.

3 Proposed Confidence Intervals

In this section, the estimator and variance of the estimator of the effect size by using bootstrap method are considered. In addition, we adjust the Bonett's confidence interval for effect size by using the standard bootstrap method which is called Bonett-Bootstrap confidence interval.

Let $y_j = x_{j1}, x_{j2}, x_{jn}$; j = 1, 2 be a random sample from normal distribution with known mean μ_j and variance σ^2_j and $y^*_j = (x^*_{j1}, x^*_{j2}, ..., x^*_{jn})$; $j = 1, 2$ indicated n independence draws from $\hat{F}$, called a bootstrap sample. Because $\hat{F}$ is empirical distribution of the data, a bootstrap sample turns out to be the same as a random sample of size n drawn with replacement from the actual sample (x_{j1}, x_{j2}, x_{jn}).

The Monte Carlo algorithm proceeds in three steps as follows:

Step 1. Using a random number generator, independently draw a large number of bootstrap samples, is called $y^*_{(1)}$, $y^*_{(2)}$,..., $y^*_{(B)}$; where B is bootstrap replications.

Step 2. For each the bootstrap sample $y^*_{(b)}$, evaluate the sample estimate of δ, is called $\hat{\delta}^*_{(b)} = \hat{\delta}(y^*_{(b)})$, b = 1, 2..., B;

where $\hat{\delta}^*_{(b)} = \frac{(\overline{X}_{b1} - \overline{X}_{b2})}{\left[(S^2_{b1} + S^2_{b2}) / 2 \right]^{1/2}}$.

Step 3. To calculate the bootstrap sample estimate of $\delta_{(i)}$ is as follow

$$\hat{\delta}^*_{(i)} = \frac{1}{B} \sum_{b=1}^{B} \hat{\delta}^*_{(b)}$$

Consequently, the Bonett-Bootstrap estimator of effect size (δ) is given by [Equation (12)]

$$\hat{\delta}_{BNB} = \frac{1}{m} \sum_{i=1}^{m} b_i \hat{\delta}^*_{(i)} \quad (12)$$

where b is an approximate bias adjustment.

The sample variance of $\hat{\delta}^*_{(i)}$ is obtained from the following Equation (13):

$$Var(\hat{\delta}^*_{(i)}) = \frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\delta}^*_{(b)} - \hat{\delta}^*_{(i)} \right) \quad (13)$$

where B is the number of bootstrap resampling.

An approximate 100(1 – α)% confidence interval for effect size (δ) as follow Equation (14)

$$\left(\hat{\delta}_{BNB} \pm z_{\alpha/2} \left[\frac{\sum_{i=1}^{m} b_i^2 Var(\hat{\delta}^*_{(i)})}{m^2} \right]^{1/2} \right) \quad (14)$$

where $z_{\alpha/2}$ is the 100(1 – α/2) percentile of the standard normal distribution.

4 Simulation Study

Monte Carlo simulation was conducted using the R statistical software to investigate the estimated coverage probabilities and expected lengths of the proposed confidence interval and to compare with the existing confidence intervals.

We simulated data for two independent groups from two normal distributions and the values of effect size parameters (δ) are equal to 0.0, 0.5, 1.0, 1.5 and 2.0 respectively. In this study, we fixed the nominal confidence interval for effect size parameters equal to 95%. The coverage probability and expected length were estimated by means of the number of simulation study (M) is equal to 50,000. We consider in situation of the number of study in meta-analysis (m) is equal to 10 studies with varied sample sizes (n) are equal 15, 30 and 50 to represent small, medium and large sample sizes respectively. The number of bootstrap resampling (B) is equal to 1,000.

The estimated coverage probability is given by

$$\text{Coverage}_l = \begin{cases} 1 \;;\; Lower_l \;\le\; \delta \;\le\; Upper_l \\ 0 \;;\; Otherwise. \end{cases}$$

where $l = 1, 2, \ldots, m$.

$$\text{Coverage Probability} = \frac{\sum_{l=1}^{M} Coverage_l}{M}$$

where $Lower_l$ is the lower limit of confidence interval
$Upper_l$ is the upper limit of confidence interval
M is the number of simulation study.

The expected length are given by

$$Length_l = Upper_l - Lower_l$$

$$Expected\ Length = \frac{\sum_{l=1}^{M} Length_l}{M}$$

where $Lower_l$ is the lower limit of confidence interval
$Upper_l$ is the upper limit of confidence interval

The simulation results obtained for normal distribution and sample size n = 15 are shown in Table 1 and Table 2. In the Table 1 shows the coverage probabilities and Table 2 shows the expected lengths for 95% confidence interval obtained from each of three methods. It can be seen that proposed method is better than another because its coverage probability is closer to 95% confidence interval and the expected lengths is shorter than the old methods.

Table 1: The estimated coverage probabilities of 95% confidence intervals for effect size in normal distribution for n = 15

		Coverage Probabilities		
n	***δ***	**Glass**	**Bonett**	**Bootstap**
15	0.0	0.9470	0.9620	0.9560
	0.5	0.9450	0.9540	0.9560
	1.0	0.9420	0.9560	0.9500
	1.5	0.9090	0.9490	0.9580
	2.0	0.9190	0.9530	0.9543

Table 2: The estimated expected lengths of 95% confidence intervals for effect size in normal distribution for n = 15

		Expected Lengths		
n	***δ***	**Glass**	**Bonett**	**Bootstap**
15	0.0	0.4569	0.4602	0.4556
	0.5	0.4643	0.4682	0.4552
	1.0	0.4861	0.4915	0.4027
	1.5	0.5191	0.5272	0.5043
	2.0	0.5631	0.5753	0.5001

The simulation results obtained for normal distribution and sample size n = 30 are shown in Table 3 and Table 4. In the Table 3 shows the coverage probabilities and Table 4 shows the expected lengths for 95% confidence interval obtained from each of three methods. It can be seen that the proposed method is better than another, because in terms of minimum expected lengths and coverage probability is closer to 95% confidence interval.

Table 3: The estimated coverage probabilities of 95% confidence intervals for effect size in normal distribution for n = 30

		Coverage Probabilities		
n	***δ***	**Glass**	**Bonett**	**Bootstap**
30	0.0	0.9480	0.9530	0.9550
	0.5	0.9510	0.9560	0.9570
	1.0	0.9490	0.9490	0.9501
	1.5	0.9140	0.9510	0.9520
	2.0	0.9130	0.9500	0.9510

Table 4: The estimated expected lengths of 95% confidence intervals for effect size in normal distribution for n = 30

		Expected Lengths		
n	***δ***	**Glass**	**Bonett**	**Bootstap**
30	0.0	0.3215	0.3227	0.3203
	0.5	0.3264	0.3279	0.3100
	1.0	0.3411	0.3432	0.3406
	1.5	0.3646	0.3676	0.3631
	2.0	0.3951	0.3993	0.3072

The simulation results obtained for normal distribution and sample size n = 50 are shown in Table 5 and Table 6. In the Table 5 shows the coverage probabilities and Table 6 shows the expected lengths for 95% confidence interval obtained from each of three methods. Similarly, the proposed confidence interval is shorter than another and the estimated coverage probability for the proposed method is close to the nominal confidence level 95% for all cases.

Table 5: The estimated coverage probabilities of 95% confidence intervals for effect size in normal distribution for n = 50

		Coverage Probabilities		
n	***δ***	**Glass**	**Bonett**	**Bootstap**
50	0.0	0.9550	0.9550	0.9530
	0.5	0.9400	0.9470	0.9580
	1.0	0.9230	0.9480	0.9530
	1.5	0.9290	0.9550	0.9520
	2.0	0.9240	0.9570	0.9530

Table 6: The estimated expected lengths of 95% confidence intervals for effect size in normal distribution for n = 50

		Expected Lengths		
n	***δ***	**Glass**	**Bonett**	**Bootstap**
50	0.0	0.2486	0.2492	0.2499
	0.5	0.2524	0.2531	0.2541
	1.0	0.2639	0.2648	0.2663
	1.5	0.2817	0.2831	0.2856
	2.0	0.3049	0.3069	0.3105

5 Conclusions

The standardized mean difference is the mean difference between groups in standard score form such as the ratio of the difference between the treatment and control means to the standard deviation. The confidence interval for the effect size base on a bootstrap estimator has been developed. The proposed confidence intervals are compared with Glass's and Bonett's confidence intervals through a Monte Carlo simulation study. The Bonett-Bootstrap confidence interval is better than another because its coverage probability is closer to the desired value (95% confidence interval) and its expected length is shorter than another for all situations.

Acknowledgments

The author would like to express my gratitude to Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Thailand for a supporting research grant.

References

[1] G. V. Glass, "Primary secondary and meta-analysis of research," *Educational Researcher*, vol. 5, no. 10, pp. 3–5, Nov. 1976.

[2] L. V. Hedges, "Fixed and random effects models in meta-analysis," *Psychological Methods*, vol. 3, pp. 486–504, 1998.

[3] C. F. Bond, W. L. Wiitala, and F. D. Richard, "Meta-analysis of raw mean differences," *Psychological Methods*, vol. 8, pp. 406–418, 2003.

[4] D. G. Bonett, "Confidence intervals for standardized linear contrasts of means," *Psychological Methods*, vol. 13, pp. 99–109, 2008.

[5] D. G. Bonett, "Meta-analytic interval estimation for standardized and unstandardized mean differences," *Psychological Methods*, vol. 14, pp. 225–238, 2009.

[6] B. Efron and R. Tibshirani, "Bootstrap methods for standard errors, confidence intervals and other measure of statistical accuracy," *Statistical Science*, vol. 1, pp. 54–57, 1986.

[7] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ: Erlbaum, 1988.