Psicothema

Psicothema 2014, Vol. 26, No. 3, 391-394
doi: 10.7334/psicothema2013.309

ISSN 0214 - 9915 CODEN PSOTEG

www.psicothema.com

Sample size planning for multiple correlation: reply to Shieh (2013)

Douglas Bonett[1] and Thomas Wright[2]
[1] University of California and [2] Fordham University

Abstract

Background: Bonett and Wright (2011) proposed a simple and accurate sample size planning formula for estimating a squared multiple correlation with desired relative precision. Shieh (2013) incorrectly evaluated the accuracy of the Bonett-Wright formula. **Method:** To address a criticism of Shieh that the Bonett-Wright formula was not examined under a wider range of conditions, the accuracy of the Bonett-Wright sample size formula is evaluated under the additional conditions proposed by Shieh. A simple 2-step sample size formula for desired absolute precision is proposed and its accuracy is evaluated under the conditions proposed by Shieh. **Results:** The analyses indicate that the Bonett-Wright sample size formula for relative prediction and the new 2-step sample size formula for absolute precision are remarkably accurate. **Conclusions:** Simple sample size planning formulas for a squared multiple correlation are important tools in designing a multiple regression analysis where the primary goal is to obtain an acceptably accurate estimate of the squared multiple correlation. The computationally intensive and simulation-based methods proposed by Shieh are not necessary.

Keywords: multiple regression, relative precision, absolute precision.

Resumen

Determinación del tamaño de muestra para la correlación múltiple: respuesta a Shieh (2013). **Antecedentes:** Bonett y Wright (2011) propusieron una fórmula simple y precisa para calcular el tamaño de muestra necesario cuando el objetivo es estimar, con una cierta precisión relativa, el cuadrado del coeficiente de correlación múltiple. Shieh (2013) incorrectamente evaluó la fórmula de Bonett-Wright. **Método:** para responder a la crítica hecha por Shieh de que la fórmula de Bonett-Wright no había sido examinada para una amplia gama de condiciones, la fórmula es ahora evaluada bajo las condiciones adicionales propuestas por Shieh. También se propone una fórmula simple, en dos etapas, para calcular el tamaño de muestra necesario y su exactitud es evaluada bajo las condiciones propuestas por Shieh. **Resultados:** el análisis indica que la fórmula de Bonett-Wright bajo el criterio de predicción relativa y la nueva fórmula de dos etapas bajo el criterio de precisión absoluta son muy precisas. **Conclusiones:** es muy útil tener fórmulas que permitan planificar el tamaño de muestra en el contexto del análisis de regresión múltiple cuando el objetivo es estimar el cuadrado del coeficiente de correlación múltiple con una precisión aceptable. Los métodos de cómputo intensivo y por simulación propuestos por Shieh no son necesarios.

Palabras clave: regresión múltiple, precisión relativa, precisión absoluta.

Multiple regression is one of the most frequently used statistical methods in the social and behavioral sciences (Gordon, 2010, pp. 1-4). In most multiple regression analyses, a point estimate of the squared multiple correlation is reported and is often given primary emphasis in the interpretation of results. However, a point estimate of a squared multiple correlation that has been obtained from sample data will contain sampling error of unknown direction and magnitude. Consequently, it is important to supplement a squared multiple correlation point estimate with a confidence interval for the *population* squared multiple correlation. The reporting of effect sizes and confidence intervals are now "the minimum expectation for all APA journals" (*Publication Manual of the American Psychological Association*, 2010, p. 33).

When planning a multiple regression analysis, it is important to obtain a sample size that is large enough to provide an acceptably narrow confidence interval for important population parameters such as the squared multiple correlation, denoted here as ρ^2. The textbook recommendations for planning a multiple regression analysis can be very misleading. For example, in a regression model with k predictor variables, Harris (1975) recommends a sample size of $n = 50 + k$ and Green (1991) recommends a sample size of $n = 50 + 8k$. Most sample size recommendations do not distinguish between the sample size required to test H_0: $\rho^2 = 0$ with desired power and the sample size required to obtain a $(1 - \alpha)$ confidence interval for ρ^2 with desired precision. Bonett and Wright (2011) showed that the approximate sample size required to obtain a $(1 - \alpha)$ confidence interval for ρ^2, with desired upper and lower interval estimates denoted as $\tilde{L}$ and $\tilde{U}$, can be expressed as

$$n = 16\tilde{\rho}^2\left[Z_{\alpha/2} / ln(\tilde{e})\right]^2 + k + 2 \tag{1}$$

where $\tilde{e} = (1 - \tilde{L})/(1 - \tilde{U})$ and $\tilde{\rho}^2 = (\tilde{L} + \tilde{U})/2$ is a planning value of ρ^2. Equation 1 clearly shows that the sample size requirement is not a multiplicative function of k, as suggested by Green (1991). Equation 1 also shows how higher levels of confidence and greater *relative precision* (smaller values of $\tilde{e}$) both play a fundamental

Received: November 14, 2013 • Accepted: February 11, 2014
Corresponding author: Douglas Bonett
Faculty of Psychology
University of California
95064 Santa Cruz (USA)
e-mail: dgbonett@ucsc.edu

role in the sample size requirement. The required sample size also depends on the planning value of ρ^2. However, the relation between the sample size requirement and ρ^2 needs to be qualified by the fact that larger values of $\tilde{e}$ are typically more appropriate with larger ρ^2 values. For instance, assuming $\rho^2 = .2$, the desired lower and upper limits might be $\tilde{L} = .15$ and $\tilde{U} = .25$ giving $\tilde{e} = (1 - .15)/(1 - .25) = 1.13$. With $\rho^2 = .9$ and desired lower and upper limits of .85 and .95 (also a width of .1), the desired relative precision is $\tilde{e} = (1 - .85)/(1 - .95) = 3$.

Shieh (2013) is highly critical of Equation 1, claiming that it is insufficiently accurate and that only computationally intensive simulation-based computer programs should be used to determine sample size requirements for the squared multiple correlation coefficient. Shieh also claimed that Bonett and Wright (2011) did not evaluate the accuracy of Equation 1 under a sufficiently wide range of conditions. Although Bonett and Wright evaluated the accuracy of Equation 1 for 42 realistic conditions ($\tilde{\rho}^2 = .05$ to .95, $\tilde{e} = 1.05$ to 3, $k = 2$ and 10), Shieh claimed that Equation 1 might be inaccurate under other conditions. Shieh examined the accuracy of Equation 1 for 27 additional conditions and concluded that Equation 1 is "not recommended for precise interval estimation of squared multiple correlation coefficient in multiple regression analysis" (p. 406) and "instead of the simplified formulas, it is prudent to consider a more sophisticated approach" (p. 406). However, Shieh's conclusion is based on an erroneous analysis of the accuracy of Equation 1. Specifically, Shieh incorrectly evaluated Equation 1 in terms of *absolute precision* $\tilde{w} = (\tilde{U} - \tilde{L})$ rather than *relative precision* $\tilde{e} = (1 - \tilde{L})/(1 - \tilde{U})$. Bonett and Wright emphasized the fact that Equation 1 was derived to approximate relative precision. Relative precision and absolute precision are two completely different sample size criteria, and a sample size approximation for desired relative precision, such as Equation 1, will not give the same result as a sample size approximation for desired absolute precision.

Bonett and Wright (2011) also proposed a simple method to approximate the sample size requirement to estimate ρ^2 with desired relative precision and specified "assurance". Assurance is the probability that an observed confidence interval will have desired or better precision. Shieh (2013) incorrectly evaluated the accuracy of the Bonett-Wright sample size method for specified assurance in terms of absolute precision rather than relative precision.

In the last few years, the use of open-source `R` statistical functions has increased dramatically in the social and behavioral sciences. Researchers can now compute a confidence interval for a squared multiple correlation using a simple `R` command (see Kelley, 2007). With this readily available resource, we can now recommend a simple 2-step sample size formula that will closely approximate the sample size needed to estimate a squared multiple correlation with desired *absolute precision*. Any future assessment of our sample size formulas should examine Equation 1 with respect to *relative precision*, and the proposed 2-step procedure should be examined with respect to *absolute precision*.

The approximate sample size requirement to estimate ρ^2 with desired absolute precision $\tilde{w} = (\tilde{U} - \tilde{L})$ can be obtained in two computational steps. First, compute the following step-1 sample size approximation

$$n_1 = 16\tilde{\rho}^2\left(1-\tilde{\rho}^2\right)^2 \left(Z_{\alpha/2} / \tilde{w}\right)^2 + k + 2 \tag{2}$$

Using a result from Bonett and Wright (2000), define a step-2 sample size approximation as

$$n_2 = \left(n_1 - k\right)\left(w_1 / \tilde{w}\right)^2 + k \tag{3}$$

where w_1 is the width of a confidence interval for ρ^2 based on a sample of size n_1. To obtain w_1, use the `ci.R2` function in the "MBESS" `R` package with the sample size set to n_1 and the sample squared multiple correlation set to its expected value. The expected value is approximately $1 + (n - k)(\tilde{\rho}^2 - 1)/(n - 1)$.

To illustrate the computation of Equations 2 and 3, suppose a researcher is planning a multiple regression analysis with $k = 4$ predictor variables and wants to compute a 95% confidence interval for ρ^2. The researcher believes that the population squared multiple correlation is about .4 and would like the width of the 95% confidence interval to be about .3. Applying Equation 2 gives $n_1 = 16(.4)(.6)^2(1.96/.3)^2 + 6 = 104.3$, which would be rounded up to 105. If the population squared multiple correlation is .4, then the expected value of the estimated squared multiple correlation would be about $1 + (105 - 4)(.4 - 1)/(105 - 1) = .417$ in a sample of size 105. In `R`, enter the command

```
ci.R2(R2=.417, N=105, K=4, conf.level=.95,
        Random.Predictors=T)
```

which displays a lower limit of .246 and an upper limit of .539 corresponding to a width of $w_1 = .293$. Computing Equation 3 gives a step-2 approximation of $n_2 = (105 - 4)\left(\frac{.293}{.3}\right)^2 + 4 = 100.3$, which would be rounded up to 101. It is interesting to note that Equations 2 and 3 also can be used to approximate the sample size required to estimate η^2 in a fixed-x regression model or in an ANOVA with desired absolute precision by setting `Random.Predictors=F` in the `ci.R2` function.

Unlike Equation 1, where $\tilde{\rho}^2$ can be replaced with its one-sided upper prediction limit to approximate the specified "assurance" (see Bonett & Wright, 2011), Equation 2 is maximized at $\tilde{\rho}^2 = 1/3$. To approximate the specified assurance with absolute precision, two-sided prediction limits for $\tilde{\rho}^2$ are needed and $\tilde{\rho}^2$ is replaced with the value within the prediction limits that is closest to 1/3. Bonett and Wright (2011) gave a simple approximation for the upper one-sided prediction limit, but we now recommend using an exact one-sided prediction limit that can be easily obtained using the `ci.R2` function.

Accuracy of sample size formulas

Shieh (2013) claimed that the evaluation of Equation 1 performed by Bonett and Wright (2011) under 42 conditions was not sufficiently detailed, and he examined Equation 1 under 27 additional conditions. Unfortunately, Shieh inappropriately evaluated the accuracy of Equation 1 in terms of absolute precision rather than relative precision. We will examine Equation 1 in terms of relative precision and Equation 3 in terms of relative precision for 24 of the 27 conditions proposed by Shieh. Shieh defined relative precision as $\tilde{e} = (1 - \tilde{\rho}^2 + \tilde{w}/2)/(1 - \tilde{\rho}^2 - \tilde{w}/2)$ where $\tilde{w}$ is the desired absolute precision. However, in three of the 27 conditions $\tilde{e}$ is undefined because $1 - \tilde{\rho}^2 - \tilde{w}/2$ is non-positive.

In two conditions where Shieh had used $\tilde{\rho}^2 = .9$, we used $\tilde{\rho}^2 = .85$ (for $\tilde{w} = .2$) or $\tilde{\rho}^2 = .75$ (for $\tilde{w} = .4$) so that $1 - \tilde{\rho}^2 - \tilde{w}/2$ would be a positive value. The accuracy of Equation 1 under 24 conditions not considered by Bonett and Wright is summarized in Table 1. Note that the expected relative precision (denoted as $\hat{e}$ in Table 1) for the sample size approximation given by Equation 1 is very close to the desired relative precision ($\tilde{e}$) in all 24 conditions. The accuracy of Equation 1 described in Table 1 is very similar to the results reported by Bonett and Wright for 42 other conditions. Note also that the expected absolute precision (denoted as $\hat{w}$ in Table 1) for the sample size approximation given by Equation 3 is very close to the desired absolute precision ($\tilde{w}$) in all 24 conditions.

Discussion

As can be seen in Table 1, Equations 1 and 3 are remarkably accurate. In fact, they are more accurate than necessary because these sample size formulas require a planning value for the population squared multiple correlation and researchers are usually unable to accurately specify this value. A minor misspecification of $\tilde{\rho}^2$ will result in a far greater sample size planning error than the approximation errors in Equations 1 and 3. For example, with a planning value of $\tilde{\rho}^2 = .1$ and a desired confidence interval width of $\tilde{w} = .2$, the required sample size is about 135, but with a planning value of $\tilde{\rho}^2 = .2$, the sample size requirement jumps to 202. When planning a multiple regression analysis, we recommend that researchers compute Equations 1 and 2 using several different plausible values of $\tilde{\rho}^2$ and use the largest sample size value.

Table 1
Accuracy of sample size formulas

$\tilde{L}$	$\tilde{U}$	$\tilde{e}$	$\tilde{w}$	n	n_2	$\hat{e}$	$\hat{w}$
0	.20	1.25	0.2	131	135	1.27	0.200
.10	.30	1.29	0.2	202	202	1.29	0.200
.20	.40	1.33	0.2	230	229	1.33	0.200
.30	.50	1.40	0.2	225	225	1.40	0.200
.40	.60	1.50	0.2	194	197	1.50	0.200
.50	.70	1.67	0.2	149	154	1.66	0.199
.60	.80	2.00	0.2	97	104	1.93	0.200
.70	.90	3.00	0.2	48	60	2.88	0.196
.75	.95	5.00	0.2	28	40	4.23	0.197
0	.30	1.43	0.3	80	80	1.45	0.299
.05	.35	1.46	0.3	93	91	1.47	0.302
.15	.45	1.54	0.3	105	104	1.55	0.299
.25	.65	1.67	0.3	102	102	1.66	0.300
.35	.65	1.86	0.3	88	90	1.83	0.300
.45	.75	2.20	0.3	67	72	2.15	0.299
.55	.85	3.00	0.3	43	50	2.87	0.302
.65	.95	7.00	0.3	20	32	5.94	0.291
0	.40	1.67	0.4	55	53	1.69	0.400
.10	.50	1.80	0.4	61	60	1.79	0.399
.20	.60	2.00	0.4	59	59	1.97	0.400
.30	.70	2.33	0.4	50	53	2.28	0.399
.40	.80	3.00	0.4	38	43	2.85	0.400
.50	.90	5.00	0.4	24	32	4.38	0.397
.55	.95	9.00	0.4	17	27	6.80	0.392

Note: n is the sample size approximation for desired relative precision, n_2 is the sample size approximation for desired absolute precision, $\hat{e}$ is the relative precision that would be expected with a sample of size n (from Equation 1), and $\hat{w}$ is the absolute precision that would be expected with a sample of size n_2 (from Equation 3)

Equations 1 and 2 are sufficiently accurate for classroom discussions of sample size requirements for multiple correlation. These sample size formulas have many pedagogical advantages over the popular rules of thumb – specifically, Equations 1 and 2 clearly describe how the sample size requirement depends on the desired level confidence, the desired precision, and the planning value of the population squared multiple correlation. Equations 1 or 2 also could be used in the classroom to easily illustrate how the sample size requirement changes for different levels of confidence and different levels of precision. In addition, Equation 2 might be preferred in classroom discussions because it does not require the computation of $ln(\tilde{e})$. Another advantage of Equation 2 is that $\tilde{\rho}^2$ can be set to 1/3 to give a conservatively large sample size requirement in applications where the researcher has little prior knowledge regarding a plausible value of the squared multiple correlation.

Many of the conditions examined by Shieh (2103) correspond to confidence intervals that are uselessly wide. Although Table 1 includes these wide interval conditions to assess the accuracy of Equations 1 and 3 under diverse conditions, researchers should specify desired lower and upper limits that will provide an accurate description of ρ^2. For instance, the conditions in Table 1 where $\tilde{L} = 0$ would not be useful in practice because the lower limit suggests that the predictor variables are completely unrelated to the response variables, but the upper limit suggests that there is a potentially important relation between the predictor variables and the response variables. As can be seen in Table 1 for the $\tilde{w} = .2$ conditions, large sample sizes are needed to accurately estimate ρ^2 for values of ρ^2 that are most common in the social and behavioral sciences. Most of the $\tilde{w} = .4$ conditions may not be useful in practice because multiple courses of action could be implied within such a wide range of values (see Mathews, 2010, p. 3). When planning a study to estimate ρ^2, the confidence interval should be sufficiently narrow to provide an unambiguous assessment of the strength of the relation between the predictor variables and the response variable.

Many published estimates ρ^2 of have used sample sizes, perhaps based on misleading rules of thumb, that were too small to produce an informatively narrow confidence interval for ρ^2. The results summarized in Bonett and Wright (2011) and in Table 1 suggest that Equations 1 and 3 are very simple and useful sample size planning tools that can assist researchers in designing multiple regression studies that will provide informative confidence intervals for ρ^2. The computationally intensive simulation-based sample size approaches recommended by Shieh (2013), which are unnecessarily complicated, are "black box" procedures that lack the pedagogical benefits of Equations 1 and 2.

References

American Psychological Association (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC. Author.

Bonett, D.G., & Wright, T.A. (2000). Sample size requirements for Pearson, Kendall, and Spearman correlations. *Psychometrika*, *65*, 23-28.

Bonett, D.G., & Wright, T.A. (2011). Sample size requirements for multiple regression interval estimation. *Journal of Organizational Behavior, 32*, 822-830.

Gordon, R.A. (2010). *Regression analysis in the social sciences*. New York: Routledge.

Green, S.B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*, *26*, 499-510.

Harris, R.J. (1975). *A primer of multivariate statistics*. New York: Academic Press.

Kelley, K. (2007). Methods for the behavioral, educational, and social science: An R package. *Behavior Research Methods*, *39*, 979-984.

Mathews, P. (2010). *Sample size calculations: Practical methods for engineers and scientists*. Fairport Harbor, OH: Mathews, Malnar and Bailey.

Shieh, G. (2013). Sample size requirements for interval estimation of the strength of association effect sizes in multiple regression analysis. *Psicothema, 25*, 402-407.