Multivariate Behavioral Research
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hmbr20

Sample Size Tables for Correlation Analysis with Applications in Partial Correlation and Multiple Regression Analysis

James Algina & Stephen Olejnik

Published online: 10 Jun 2010.

Multivariate Behavioral Research, *38* (3), 309-323

Sample Size Tables for Correlation Analysis with Applications in Partial Correlation and Multiple Regression Analysis

James Algina
University of Florida

Stephen Olejnik
University of Georgia

Tables for selecting sample size in correlation studies are presented. Some of the tables allow selection of sample size so that r (or r^2, depending on the statistic the researcher plans to interpret) will be within a target interval around the population parameter with probability .95. The intervals are ±.05, ±.10, ±.15, and ±.20 around the population parameter. Other tables allow selection of sample size to meet a target for power when conducting a .05 test of the null hypothesis that a correlation coefficient is zero. Applications of the tables in partial correlation and multiple regression analyses are discussed. SAS and SPSS computer programs are made available to permit researchers to select sample size for levels of accuracy, probabilities, and parameter values and for Type I error rates other than those used in constructing the tables.

For about 40 years, data analysts have been recommending to researchers in the behavioral sciences that an effect size measure should be reported in addition to a test for statistical significance (Cohen, 1965; Hays, 1963). In recent years there has been increased support for reporting effect sizes. Three examples of the increased support are as follows.

1. According to *The Publication Manual of the American Psychological Association* (2001) "it is almost always necessary to include some index of effect size or strength of relationship in your Results section." (p. 25).
2. The Editor of *Journal of Applied Psychology* requires an explanation for failing to report an effect size (Murphy, 1997).
3. The journal *Educational and Psychological Measurement* requires effect size estimates to be reported (Thompson, 1994).

The practice of reporting effect sizes has also received support from the APA Task Force on Statistical Inference (Wilkinson and the Task Force on Statistical Inference, 1999). The emphasis on reporting effect sizes and strength of relationship implies that researchers should plan studies not only to have sufficiently powerful hypothesis tests but also to have sufficiently accurate estimates of effect sizes.

One of the most widely used statistics in the social and behavioral sciences is the Pearson product moment correlation coefficient *r*. An interesting characteristic of this statistic is that it serves as an index of strength of linear association for two variables as well as the basis for testing the null hypothesis that the population correlation coefficient ρ is equal to some specified value, most often zero. Since the correlation coefficient is itself an effect size, it follows that a study, employing the correlation coefficient, should be planned so the correlation coefficient will be estimated with adequate accuracy and hypotheses about the correlation coefficient will be tested with sufficient power. For example, suppose a validity study is conducted in which the population correlation coefficient is thought to be .30. A sample size of 70 is used. In this study, power for a one-tailed test that the correlation coefficient is zero would be .82. And, there would be approximately a 95% chance that the sample correlation coefficient would be in the interval (.1345, .4655). Although the study has is good power, it would have insufficient accuracy if one regards a sample correlation coefficient as small as .14 or as large as .47 as misleading. In this case, the sample size of 70 is not large enough to ensure adequate estimation accuracy.

Estimating the correlation coefficient with adequate accuracy is important not only when estimation of the strength of association between pairs of variables in the study is the ultimate goal of the study, but also when the data will be used in additional analyses such as regression or structural equation modeling. Even in the latter types of studies, the strength of association between pairs of variables is typically of some interest.

The purpose of the present study is to present a table that will facilitate selection of sample sizes to ensure that correlation coefficients are estimated with adequate accuracy. Because some researchers use r^2 as an effect size to measure strength of association, we also present a table that will facilitate selection of sample sizes to ensure estimation accuracy for squared correlation coefficients. In addition we investigate the accuracy of two sample size selection methods that are approximate, but are very easy to apply. And for the sake of completeness, we present power tables, so that researchers can select a sample size that meets the twin goals of adequate power and accuracy.

Method

When the goal is to estimate ρ with sufficient accuracy, we need to find the smallest sample size *n* necessary for the sample correlation coefficient to fall into a prescribed interval with a prescribed probability:

$$prob[\max(-1, \rho - c) \leq r \leq \min(\rho + c, 1)] \geq p. \tag{1}$$

In Equation 1, $\max(-1, \rho - c)$ and $\min(\rho + c, 1)$ are the ends of an interval into which the researcher wants r to fall and p is the probability of that event. The quantity c operationalizes the researcher's view of adequate estimation accuracy. For example, if a researcher believes the population correlation coefficient is likely to be .45 and wants the sample correlation coefficient to be within $\pm.10$ with probability of at least .95 then Equation 1 becomes

$$prob[(.45 - .10) \leq r \leq (.45 + .10)] \geq .95$$

and the problem is to find n so that this interval is met. This type interval used in Equation 1 is widely used in survey sampling in order to select sample size (see, for example, Jaeger, 1984; Kish, 1965; or Sudman, 1978).

The following steps were used to find the required sample size for a given ρ and c. Beginning with a provisional sample size of $n = 3$, find the probability that r is in the interval in Equation 1. If the probability exceeds the target, the sample size is sufficient. If the provisional sample size is not sufficient, increase the provisional sample size by one and carry out the computations again. The probability that r is in the required interval was calculated by finding the area under the density function of r, under assumed bivariate normality, for the limits in Equation 1. The area can be found by using numerical integration methods. We wrote a Mathematica program to compute the areas. The program used the GaussKronrod method, which is the default in Mathematica. The GaussKronrod method is an adaptive Gauss quadrature method with error estimation based on evaluation at Kronrod points (Wolfram, 1999). The required sample size was found for all combinations of ρ from .00 to .95 and c = .05 to .20 both in steps of .05.

Because we used a simple search procedure, the numeric integration was time consuming for some of the combinations of ρ and c. While this problem could have been solved by using a more efficient search strategy, the numeric integration also had to be monitored carefully for problems in the integration. In some cases the value for the probability in Equation 1 increased as n increased and then declined again; in other cases the value of the probability increased by a large amount with small increases in n. In still other cases Mathematica produced various warnings indicating problems in the integration. All of these problems were solved by increasing the precision with which r was represented in the calculations and the precision of the computations in NIntegrate, the procedure used in Mathematica to carry out the integration. These changes increased the time required for the computations.

Because our procedure was time consuming and had to be monitored carefully and because some readers may be interested in sample sizes for values of c or p other than those we investigated, we investigated an approximate method for finding the required sample sizes. In our approximate method,[1] Harley's (1957) approximation to the distribution of r was used. Harley used the noncentral t distribution to approximate the distribution of

$$\tilde{r} = \frac{r}{\sqrt{1-r^2}}, \tag{2}$$

Let $ll = \max(-.9999, \rho - c)$ and $ul = \min(\rho + c, .9999)$. Since $\tilde{r}$ is a monotone function of r, for a given n, ρ, and c, the probability that r is in the required interval in Equation 1 can be computed as

$$prob\left(\tilde{l}l \le \tilde{r} \le \tilde{u}l\right)$$

where $\tilde{l}l$ and $\tilde{u}l$ are calculated by substituting ll and ul, respectively, for r in Equation 2. We prepared a SAS and a SPSS program[2] to compute the probability and find the required sample size.

When the goal is to estimate ρ^2 with sufficient accuracy, we need to find the sample size necessary for the sample squared correlation coefficient to fall into a prescribed interval with a prescribed probability:

$$prob[\max(\rho^2 - c,\ 0) \le r^2 \le \min(\rho^2 + c,\ 1)] \ge p. \tag{3}$$

Again for a combination of ρ and c we used the following steps: Beginning with the provisional sample size of $n = 3$, find the probability that r^2 is in the interval in Equation 3. If the probability exceeds the target then the sample size is sufficient. If the provisional sample size is not sufficient, increase the provisional sample size by one and carry out the computations again. The required sample size was found for all combinations of $0 \le \rho \le .95$ and

[1] We also used Fisher's Z approximation with the variance for Z that is reported in Stuart, Ord, and Arnold (1994). This variance is reported to be more accurate than the usual $1/(n - 3)$ approximation. The results were less accurate than those obtained by using Harley's (1957) approximation.

[2] All SAS and SPSS programs referred to in this article can be downloaded at http://plaza.ufl.edu/algina/index.programs.html

$c = .05$ to $.20$ in steps of .05. We again used Mathematica to find the required the sample sizes by integrating the density function of r^2. The required sample size was found for all combinations of $0 \le \rho^2 \le .95$ and $c = .05$ to .20 to .20, both in steps of .05.

Although we did not have problems in the computations, the computations were time consuming for some combinations of ρ^2 and c and, consequently, we also used an approximate method to find the required sample sizes. The approximate method used Lee's (1971) F approximation to the distribution of R^2. Lee used the noncentral F distribution to approximate the distribution of

$$\tilde{\rho}^2 = \frac{\mathbf{R}^2}{1-\mathbf{R}^2},$$

where $\mathbf{R}^2$ is the population squared multiple correlation coefficient. When the number of independent variables is 1, Lee's method can be used to approximate the distribution of

$$\tilde{\rho}^2 = \frac{\rho^2}{1-\rho^2}. \tag{4}$$

Let $ll = \max(0, \rho^2 - c)$ and $ul = \min(\rho^2 + c, .9999)$. For a given n, ρ, and c, the probability that r is in the required interval is computed as

$$prob\left(\tilde{l}l \le \tilde{r} \le \tilde{u}l\right)$$

where $\tilde{l}l$ and $\tilde{u}l$ are calculated by substituting ll and ul, respectively, for ρ^2 in Equation 4. We prepared a SAS and a SPSS program to compute the probability and find the required sample size.

Results

The sample sizes obtained by integrating the density function of r are reported in Table 1. For $\rho \le .65$, results using Harley's (1957) approximation were in agreement with results in Table 1. Sample sizes computed using Harley's approximation were larger by one or two in 14 of

Table 1
Sample Sizes Necessary to Meet Target Accuracy for: p = .95: Distribution of r

	c			
ρ	.05	.10	.15	.20
0.00	1538	385	172	97
0.05	1530	383	171	97
0.10	1507	378	168	95
0.15	1469	368	164	93
0.20	1417	355	159	90
0.25	1352	339	151	86
0.30	1274	319	143	81
0.35	1185	297	133	75
0.40	1086	273	122	69
0.45	979	246	110	63
0.50	866	218	98	56
0.55	750	189	85	49
0.60	632	160	72	42
0.65	515	131	60	35
0.70	402	103	47	28
0.75	297	77	36	23
0.80	202	53	27	18
0.85	120	34	19	13
0.90	60	20	12	9
0.95	21	9	6	5

the 24 conditions in Table 1 with $\rho \geq .70$. Results produced by SAS and SPSS were in complete agreement.[3]

Inspection of the results in Table 1 indicates that smaller sample sizes are required as ρ get larger. This pattern is due to the decline in variability of the sampling distribution of r as ρ get larger. In large samples the standard error of r is

$$\frac{1-\rho^2}{\sqrt{n}}$$

and clearly declines as ρ increases.

[3] We used version 8.2 of SAS and version 11.5 of SPSS.

Table 2
Sample Sizes Necessary to Meet Target Accuracy for: p = .95: Distribution of r^2

	c			
ρ^2	.05	.10	.15	.20
.00	78	39	27	20
.05	276	87	46	30
.10	493	120	60	37
.15	662	163	69	41
.20	784	194	84	44
.25	862	214	93	51
.30	902	224	98	54
.35	907	226	99	55
.40	884	220	97	54
.45	836	209	92	52
.50	768	192	85	48
.55	685	171	76	43
.60	591	148	66	38
.65	491	124	56	32
.70	389	98	45	26
.75	290	74	35	21
.80	199	52	26	17
.85	121	33	18	12
.90	59	20	12	9
.95	21	9	6	5

Some researchers may be interested in sample size for values of p other than .95 and/or for other values of c. Since sample sizes based on Harley's (1957) approximation are quite accurate these can be obtained by using the SAS or SPSS programs based on Harley's method.

Table 2 presents required samples sizes based on the distribution of r^2. The results using Lee's (1971) approximation were in complete agreement except when ρ^2 was .90, target accuracy was $\pm$.20, and SAS was used. Then Lee's method resulted in $n = 8$ for this condition.

Review of the results in Table 2 indicates the required sample size increases at first with increases in ρ^2 and then declines with further increases in ρ^2, reflecting changes in variability of the sampling distribution of r^2 as ρ^2 get larger. In large samples the standard error of r^2 is

$$\frac{2\rho\left(1-\rho^2\right)}{\sqrt{n}}$$

provided $\rho \neq 0$. The standard error increases as ρ^2 increases to 1/3 and then decreases again as ρ^2 increases.

It should be noted that the required sample size could be quite different depending on whether one plans to estimate ρ or ρ^2. This is because a particular level of accuracy selected for estimating ρ (e.g., $\pm$.10) implies a different level of accuracy for estimating ρ^2. For example if one believes ρ is .15, plans to estimate ρ and wants $\pm$.10 accuracy, the required sample size is 368 (see Table 1) to have a .95 probability that r is in the interval .05 to .25. If one plans to estimate ρ^2, and wants $\pm$.10 accuracy, a sample size between 39 and 87 is required (see Table 2) to have a .95 probability that r^2 is in the interval $[\max(.15^2 - .10, 0) \leq r^2 \leq \min(.15^2 + .10, 1)] = (.00 \leq r^2 \leq .125)$. The exact sample size is 63. With a sample size of 368, there would be, approximately, a .95 probability that r^2 is in the interval $[\max(.15^2 - .032, 0) \leq r^2 \leq \min(.15^2 + .032, 1)] = (0 \leq r^2 \leq .0545)$. That is, ρ^2 would be estimated with much better than $\pm$.10 accuracy.

Partial Correlation Coefficients and Semi-Partial Correlation Coefficients

The distribution of a partial correlation coefficient is the distribution of a correlation coefficient with a sample size reduced by the number of control variables (cf. Stuart, Ord, & Arnold, 1994). It follows that our tables can be used to select sample sizes for partial correlation coefficients and squared partial correlation coefficients. For example, suppose a researcher believes that the population partial correlation coefficient, with four variables controlled, is .30 and wants $\pm$.15 accuracy. The relevant entry in Table 1 is 143. Adding four, the required sample size would be 147. It should be noted that the SAS and SPSS programs compute the actual sample size required, not the sample size minus the number of control variables.

Algina, Moulder, and Moser (2002) developed a sample size table for accurate estimation of a squared semi-partial correlation coefficient. These coefficients are also known as the increase in R^2 and are widely used as effect sizes in multiple regression studies. Algina et al. developed their table by using large sample analytic results due to Alf and Graf (1999) and simulation methods, and reported results for combinations of ρ_r^2 (the squared multiple correlation coefficient for the reduced model) in the range .00 to .60 and $\rho_f^2 - \rho_r^2$ (where ρ_f^2 refers to the full model) in the range .00 to .30. The quantity $\Delta\rho^2 = \rho_f^2 - \rho_r^2$ is the population squared semi-partial correlation coefficient. The results indicated that the largest sample sizes are required when $\rho_r^2 = 0.00$. Algina et al. showed that the required sample size did not depend strongly on the number of predictors and reported the maximum required sample over a range of predictors from two to ten. Table 3, by contrast, shows required sample sizes for two, six, and ten predictors. The results in Table 3 are for $\rho_r^2 = 0$. Comparing the results in Table 3 for a value of ρ_f^2 to the results in Table 3 for the same value of ρ^2, we find that the sample sizes in Table 2 are quite similar to those in Table 2. Thus for the case of $\rho_r^2 = 0$, the tables in this article can be used in place of the table in Algina et al. without much loss in accuracy. Since the required sample sizes decrease as ρ_r^2 increases, using the results in Table 3 will ensure adequate accuracy for the squared correlation coefficients as well as for the squared semi-partial correlation coefficients. This is quite useful because it is common in multiple regression studies to interpret both the squared correlation coefficients and squared semi-partial correlation coefficients.

Power

We believe it is important to plan correlation studies to have adequate estimation accuracy. Another important goal is to have adequate power. The sample size required to achieve a target for power can also be computed from the distribution of *r*. An appropriate sample size table can be found, for example, in Kraemer (1985). However, to make the article self contained for purposes of illustrating use of such tables in correlation and multiple regression studies, Tables 4 and 5 contain the required sample sizes for two- and one-tailed tests respectively using $\alpha = .05$. With a few exceptions the results were computed using Mathematica. In the exceptional cases, which are marked in bold in Tables 4 and 5, the integration would not converge and Harley's (1957) approximation, implemented in SAS, was used instead. We also compared results using Harley's approximation to all other results obtained using Mathematica. In the only discrepancy, Harley's procedure indicated a sample size that was too large by one.

As noted earlier, the distribution of a partial correlation coefficient is the distribution of a correlation coefficient with a sample size reduced by the number of control variables. Therefore Tables 4 and 5 can be used to select sample sizes to achieve adequate power for tests on the partial correlation coefficient.

In multiple regression analysis with random predictors, the power function for testing the hypothesis that a regression coefficient for an independent variable is equal to zero is equal to the power function for testing

Table 3
Sample Sizes Necessary to Meet Target Accuracy for p = .95: Distribution of $\Delta R^2 = R_f^2 - R_r^2$

			c			
k	ρ_r^2	ρ_f^2	±.05	±.10	±.15	±.20
2	0.00	0.00	70	36	24	18
2	0.00	0.05	278	75	40	25
2	0.00	0.10	498	125	56	32
2	0.00	0.15	667	167	75	42
2	0.00	0.20	787	197	88	49
2	0.00	0.25	865	217	97	55
2	0.00	0.30	904	226	101	57
6	0.00	0.00	71	35	24	18
6	0.00	0.05	278	70	33	23
6	0.00	0.10	498	125	56	32
6	0.00	0.15	667	167	70	38
6	0.00	0.20	787	197	86	38
6	0.00	0.25	865	217	97	54
6	0.00	0.30	904	226	101	57
10	0.00	0.00	69	35	24	19
10	0.00	0.05	278	70	31	19
10	0.00	0.10	498	121	55	31
10	0.00	0.15	667	167	59	27
10	0.00	0.20	787	197	88	24
10	0.00	0.25	865	217	97	53
10	0.00	0.30	904	226	101	57

the hypothesis that the partial correlation between that variable and the dependent variable, with the other independent variables partialled out, is zero. This follows from the fact that the t statistics for testing these two hypotheses are equal. Thus Tables 4 and 5 can be used to conduct power analyses in multiple regression analysis. To conduct such a power analysis one must specify the squared multiple correlation for all of the variables in a regression model ρ_f^2 and the increase in the squared multiple correlation coefficient expected to be associated with a particular variable ($\Delta\rho^2$). Then the table is entered with the positive square root of the squared partial correlation coefficient:

Table 4
Sample Size Required to Achieve Target Power for a Specified Correlation Coefficient (ρ) at α = .05: Two-Tailed Test

	Power				
ρ	.50	.60	.70	.80	.90
.05	1537	1959	2467	3137	4198
.10	384	489	616	782	1046
.15	171	217	273	346	462
.20	96	122	153	193	258
.25	62	78	97	123	164
.30	43	54	67	**84**	**112**
.35	31	39	49	**61**	**81**
.40	24	30	37	46	**61**
.45	19	23	29	36	47
.50	15	19	23	29	37
.55	13	15	19	23	30
.60	11	13	15	19	24
.65	9	11	13	16	20
.70	8	9	11	13	17
.75	7	8	9	11	14
.80	6	7	8	9	11
.85	5	6	7	8	9
.90	5	5	6	6	8

$$\sqrt{\frac{\Delta\rho^2}{1-\left(\rho_f^2-\Delta\rho^2\right)}}.$$

The required sample size is the value in the table plus $q = k - 1$ where k is the number of predictors in the model. For example, suppose there will be a total of four independent variables in a model and a literature review suggests that the $\rho_f^2 = .50$. If a particular variable is expected to uniquely account for 10% of the variance ($\Delta\rho^2 = .10$), the relevant partial correlation is

$$\sqrt{\frac{.10}{1-(.50-.10)}}=.408.$$

Table 5
Sample Size Required to Achieve Target Power for a Specified Correlation Coefficient (ρ) at α = .05: One-Tailed Test

	Power				
ρ	.50	.60	.70	.80	.90
.05	1083	1441	1881	2471	3422
.10	271	360	470	616	853
.15	121	160	208	273	377
.20	68	90	117	153	211
.25	44	58	75	97	134
.30	31	40	51	67	**92**
.35	23	29	38	49	**67**
.40	18	23	29	37	50
.45	14	18	22	29	39
.50	12	14	18	23	31
.55	10	12	15	19	25
.60	8	10	12	15	20
.65	7	9	10	13	17
.70	6	7	9	11	14
.75	6	6	7	9	12
.80	5	6	6	8	10
.85	4	5	6	6	8
.90	4	4	5	5	7

For a two-tailed test and a target power of .80, we find 46 for a .40 partial correlation coefficient and 36 for a .45 partial correlation coefficient in Table 4. Interpolating we find approximately 44 for .408. Adding $q = k - 1 = 3$ the required sample size is 47.

Discussion

Pearson product moment correlation coefficients or their squares are widely used and frequently reported in the social and behavioral sciences. Because correlation coefficients and their square serve as measures of strength of association, researchers who report these coefficients should plan their studies so that estimated coefficients have sufficient accuracy. To help researchers accomplish this goal, the purpose of the current investigation was to provide sample size tables for four accuracy criteria for r: $prob[\max(-1, \rho - c) \leq r \leq \min(\rho + c, 1)] \geq p$, where $c = .05, .10, .15$, and $.20$ and for four accuracy criteria for r^2: $prob[\max(\rho^2 - c, 0) \leq r^2 \leq \min(\rho^2 + c, 1)] \geq p$ with the same values for c. That is, in our tables we enumerated sample sizes that will ensure, under bivariate normality, that the difference between the sample and population coefficients will be no larger than some small value (i.e., .05, .10, .15, or .20) approximately 95 percent of the time. In addition, tables were presented to select sample sizes for adequate power for a test of H_0: $\rho = 0$ with $\alpha = .05$. Both the accuracy and the power tables can be used to select sample size in partial correlation and multiple regression analyses.

SAS and SPSS computer programs were prepared to permit researcher to select sample size for levels of accuracy, probabilities, and parameter values other than those used in constructing the accuracy tables. In addition a program for power analysis was prepared that permits sample size to be selected for Type I error rates other than .05.

To use the accuracy table or program for estimating ρ, one has to specify a value for ρ. Thus, use of these tools may appear circular because if one knows the value of ρ there is no reason to do the study. To plan a study to have adequate accuracy, a researcher overcomes this objection by specifying a range of population values into which s/he believes ρ is likely to fall. Sample size is then selected by using the value of ρ that results in the largest sample size. If the researcher has no a priori notion about the magnitude of ρ, $\rho = 0$ should be used. Similar comments apply to the table or program for estimating ρ^2. In this case, if the researcher has no a priori notion about ρ^2, the $\rho^2 = 1/3$ should be used. The required sample sizes are 909, 226, 95, and 55 for $c = .05, .10, .15$, and $.20$, respectively. If the researcher is planning the study to have adequate power, to use the table or program s/he would specify the minimum value for ρ that would be substantively interesting.

The approach presented in this article does not exhaust the possibilities for setting sample sizes to achieve estimation accuracy in correlation studies. For example, one possibility is to select the sample size so the length of the confidence interval computed in the study will be sufficiently narrow. A problem with this approach is that the length of a confidence interval for a correlation or squared correlation depends on the sample value for the statistic of interest. Thus one would have to know the sample value for the statistic of interest before conducting the study. This problem can be overcome by selecting the sample size assuming the value for r (or r^2 if that is the statistic of interest) that will result in the widest confident interval. For r this value is zero and the necessary sample sizes for the width of a 95% confidence interval to be less than .10, .20, .30, and .40 respectively are 1537, 384, 171 and 96. These sample sizes are very similar to those obtained to ensure that there is a .95 probability that r is within $\pm c$ of $\rho = 0$ where $c = .05, .10, .15$, and .20 (see Table 1). If the objective is a narrow confidence interval for ρ^2 then sample size should be selected assuming $r^2 = 1/3$ and the necessary sample sizes are 908, 225, 99, and 54 for lengths of .10, .20, .30, and .40 respectively. These sample sizes are very similar those obtained to ensure that there is a .95 probability that r^2 is within $\pm c$ of $\rho^2 = 1/3$ where $c = .05, .10, .15$, and .20. Thus selecting sample sizes based on the length of a confidence interval results in sample sizes very similar to those obtained if one has no a priori notion of the size of the population correlation coefficient.

Similarly, the approach presented in this article does not exhaust the possibilities for setting sample sizes to achieve power in correlation studies because there are additional hypotheses that may be of interest to test. For example, a researcher might be interested in testing hypotheses such as

$$\begin{aligned} H_0&: \rho^2 \geq \rho_0^2 \\ H_1&: \rho^2 < \rho_0^2 \end{aligned} \tag{5}$$

In Equation 5 the researcher is interested in presenting evidence that the squared correlation coefficient is smaller than ρ_0^2. This pair of hypotheses has the same structure as the hypotheses for test of close fit in structural equation modeling. Or the researcher might be interested in testing the hypothesis

$$\begin{aligned} H_0&: \rho^2 \leq \rho_0^2 \\ H_1&: \rho^2 > \rho_0^2 \end{aligned} \tag{6}$$

Here the researcher wants to present evidence that the squared correlation coefficient is larger than ρ_0^2 and the structure of the hypotheses is the same as the hypotheses for test of not close fit in structural equation modeling. Testing hypotheses like those in Equations 5 and 6 is one way of addressing the problems critics have pointed out with nil null hypotheses (e.g., H_0: $\rho^2 = 0$). Power analyses for hypotheses like these can be carried out using a program like Mathematica or Lee's approximation to the distribution of r^2. However, just as with power analyses for testing a nil null hypothesis, the sample sizes that emerge from power analyses for tests of close fit and for tests of not close fit do not guarantee sufficient accuracy of estimation.

References

Alf, E. F. & Graf, R. G. (1999). Asymptotic confidence limits for the difference between two squared multiple correlations: A simplified approach. *Psychological Methods*, *4*, 70-75.

Algina, J. Moulder, B. C., & Moser, B.K. (2002). Sample size requirements for accurate estimation of squared semi-partial correlation coefficients. *Multivariate Behavioral Research, 37,* 37-57.

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (4th Ed.). Washington, DC: Author.

Cohen, J. (1965). Some statistical issues in psychological research. In B.B. Wolman (Ed.), *Handbook of clinical psychology* (pp. 95-121). New York: Academic Press.

Harley, B. I. (1957). Relationship between the distributions of non-central *t* and of a transformed correlation coefficient. *Biometrika*, *44*, 219-224.

Hays, W. L. (1963). *Statistics*. New York: Holt, Rinehart and Winston.

Jaeger, R. M. (1984). *Sampling in education and the social sciences*. New York: Longman.

Kish, L. (1965). *Survey sampling*. New York, John Wiley.

Kraemer, H. C. (1985). A strategy to teach the concept and application of power of statistical tests. *Journal of Educational Statistics*. *10*, 173-195

Lee, Y. S. (1971). Some results on the sampling distribution of the multiple correlation coefficient. *Journal of the Royal Statistical Society*, *B*, *33*, 117-130.

Murphy, K. R. (1997). Editorial. *Journal of Applied Psychology*, *82*, 3-5.

Stuart, A., Ord, J. K., & Arnold, S. (1994). *The advanced theory of statistics* (6th ed., Vol. 2A). New York: Oxford University Press.

Sudman, S. (1978). *Applied sampling*. New York: Academic Press.

Thompson, B. (1994). Guidelines for authors. *Educational and Psychological Measurement*, *54*, 837-847.

Wilkinson, L. & the Task Force on Statistical Inference (1999). Statistical methods in psychology journals. *American Psychologist*, *54*, 594-604.

Wolfram, S. (1999). *The Mathematica book*, (4th Ed.). Champaign, IL: Wolfram Media.

Accepted February, 2003.