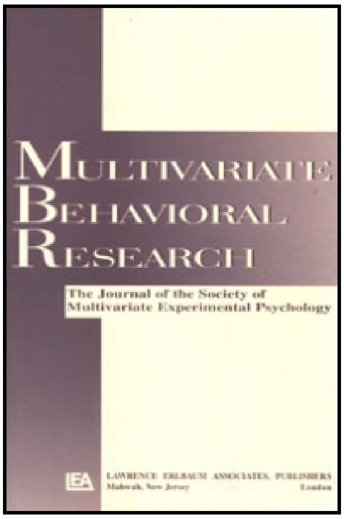

Multivariate Behavioral Research
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653673

Sample Size Calculation for Estimating or Testing a Nonzero Squared Multiple Correlation Coefficient

K. Krishnamoorthy [a]; Yanping Xia [a]
[a] University of Louisiana at Lafayette,

Online Publication Date: 01 July 2008

To cite this Article Krishnamoorthy, K. and Xia, Yanping(2008)'Sample Size Calculation for Estimating or Testing a Nonzero Squared Multiple Correlation Coefficient',Multivariate Behavioral Research,43:3,382 — 410

To link to this Article: DOI: 10.1080/00273170802285727

URL: http://dx.doi.org/10.1080/00273170802285727

Multivariate Behavioral Research, 43:382–410, 2008

ISSN: 0027-3171 print/1532-7906 online
DOI: 10.1080/00273170802285727

Sample Size Calculation for Estimating or Testing a Nonzero Squared Multiple Correlation Coefficient

K. Krishnamoorthy and Yanping Xia
University of Louisiana at Lafayette

The problems of hypothesis testing and interval estimation of the squared multiple correlation coefficient of a multivariate normal distribution are considered. It is shown that available one-sided tests are uniformly most powerful, and the one-sided confidence intervals are uniformly most accurate. An exact method of calculating sample size to carry out one-sided tests (null hypothesis may involve a nonzero value for the multiple correlation coefficient) to attain a specified power is given. Sample size calculation for computing confidence intervals for the squared multiple correlation coefficient with a specified expected width is also provided. Sample sizes for powers and confidence intervals are tabulated for a wide range of parameter configurations and dimensions. The results are illustrated using the empirical data from Timm (1975) that related scores from the Peabody Picture Vocabulary Test to four proficiency measures.

The multiple correlation coefficient is a commonly used measure of association between a random variable x_1 and a vector $\boldsymbol{x}_2$ of random variables. It is defined as the maximum correlation between x_1 and any linear combination of $\boldsymbol{x}_2$. This multiple correlation analysis is widely used in education and in social

Correspondence concerning this article should be addressed to K. Krishnamoorthy, Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA 70504. E-mail: krishna@louisiana.edu

and behavioral sciences. Correlation model and multiple linear regression model with random predictor variables involve this multiple correlation analysis. At the outset we point out that the following procedures in this article are applicable to estimate or test about the coefficient of determination in a multiple linear regression model only when the predictor variables are random (random model) and the response variable (dependent variable) and the predictor variables jointly follow a multivariate normal distribution. Methods for fitting a multivariate normal distribution for a given data set can be found in many multivariate textbooks; for example, see Johnson and Wichern (2002, p. 177) and Srivastava (2002, Section 3.5). Helland (1987) argued that the sample coefficient of determination in a linear regression model can be considered an estimator of a population parameter only when the predictor variables are random. That is, predictor variables are not fixed or controlled a priori; the measurements on the predictor variables and the response variable are obtained from a random sample of units. This random model commonly arises in behavioral and social sciences where levels of the multiple variables from each sampling unit cannot be controlled and are available only after the measurements are made. For more details on the differences between the correlation analysis with random predictor variables and the one with fixed predictor variables, see the article by Mendoza and Stafford (2001). For a specific situation where the multiple correlation analysis is used, see the example section.

To describe the problems that we address in this article, suppose we observe a set of p scores from each subject in a sample of n subjects from a population. Let R^2 denote the squared sample multiple correlation coefficient between the first variable and the set of remaining $p-1$ variables. As the formulas for computing R^2 in a random model or fixed level model (linear regression model) are the same, many standard software packages, such as Minitab, SAS, Matlab, and Mathematica, that provide regression analysis, can be used to compute R^2. Let ρ^2 denote the squared population multiple correlation coefficient. It is well known that R^2 is the maximum likelihood estimate of ρ^2, which by itself is of limited practical use; confidence intervals and hypothesis tests are warranted to assess the true value of ρ^2.

As mentioned earlier, finding a confidence interval for ρ^2 requires the assumption that the sample is from a p-variate normal population. Under this assumption and $n-p$ is even, Fisher (1928) derived an explicit expression for the cumulative distribution function (CDF) of R^2, that is, $P(R^2 \leq x|n, p, \rho^2)$, where P denotes the probability. Since then many alternative expressions for the CDF (for $n \geq p$) have been developed in the literature, and among them the one due to Gurland (1968) is simple to compute and easy to program in a computer language. Ding and Bargmann (1991) and Benton and Krishnamoorthy (2003) provided algorithms to compute the CDF due to Gurland. As available computing technologies allow us to compute the CDF accurately, and the CDF of R^2 depends only on ρ^2, inferences on ρ^2 can be made easily. For instance, if the hypotheses of interest are

$$H_0 : \rho^2 \leq \rho_0^2 \text{ vs. } H_a : \rho^2 > \rho_0^2, \tag{1}$$

where ρ_0^2 is a specified value, then the p value for testing the aforementioned hypotheses is given by $P(R^2 > r^2 | n, p, \rho_0^2)$, where r^2 is an observed value of R^2. The null hypothesis will be rejected if the p value is less than a predetermined nominal level α. Testing nil-null hypothesis (that is, $H_0 : \rho^2 = 0$) is widely documented in the literature and textbooks mainly because, in this case, percentiles of R^2 can be obtained from F percentiles that are tabulated in many standard textbooks. However, the results of testing nil-null hypothesis are less informative as the rejection of nil-null hypothesis merely indicates that the population ρ^2 is a positive number. On the other hand, the results of testing nonzero null hypothesis is more informative as the rejection of the null hypothesis in Equation (1) shows that the population ρ^2 is at least ρ_0^2.

One-sided confidence limits or confidence intervals are usually obtained by inverting the appropriate test procedures. For instance, by inverting the test of hypotheses in Equation (1), a lower confidence limit for ρ^2 can be obtained. Similarly, by inverting a test of two-sided alternative hypothesis, one can get a confidence interval for ρ^2. As will be seen later, the method of determining confidence limits turns out to be solving an equation involving the CDF of R^2 and the confidence coefficient. Table values for computing confidence limits for ρ^2 are given in the literature (Kramer, 1963; Lee, 1972). These tables, as pointed out by Helland (1987), are of limited use and not applicable for constructing upper limits. Helland provided an iterative method to compute one-sided limits for ρ^2. Steiger and Fouladi (1992) provided a program refer to as "R2," which computes confidence intervals for ρ^2 and powers for testing $\rho^2 = 0$. Mendoza and Stafford (2001) tabulated lower bounds for ρ^2 when R^2 takes values no more than .64 and values of p less than or equal to 17. These authors also provided *Mathematica* functions to compute confidence intervals for ρ^2. The PC calculator that accompanies the book by Krishnamoorthy (2006) computes confidence intervals and p values for testing a non-nil null hypothesis about ρ^2. This PC calculator is referred to as *StatCalc* and is freely available from http://www.ucs.louisiana.edu/∼kxk4695

The necessity and importance of power analysis in the areas of social, behavioral, and psychological sciences are well addressed in the literature; see, for example, the recent article by Maxwell (2004). In general, it is wise to determine the sample size prior to sampling in order to get significant results. Let us now suppose that a researcher decided to use the test procedures discussed in the preceding paragraph, and likes to determine the sample size required for the test to detect a specified difference between the true value of ρ^2 and

ρ_0^2 with a higher probability (power). For this purpose, many articles (e.g., Algina & Olejnik, 2003; Mendoza & Stafford, 2001; Steiger & Fouladi, 1992) provided sample size calculation for testing $H_0 : \rho^2 = 0$ vs. $H_a : \rho^2 > 0$. For the bivariate case, Cohen (1988) provided approximate sample sizes for testing nil-null hypothesis. As pointed out earlier, in many applications, testing $H_0 : \rho^2 = \rho_0^2$ vs. $H_a : \rho^2 > \rho_0^2$ is more meaningful than testing zero null hypothesis. For example, if an experimenter believes that the population ρ^2 is about 0.8 and likes to test $H_0 : \rho^2 \leq 0.7$ vs. $H_a : \rho^2 > 0.7$, then she or he may want to determine the sample size so that the test result will be significant with a higher probability. In view of this interest, we provide sample size calculation for testing one-sided hypotheses with a given input parameters ρ^2 and ρ_0^2 and a given nominal level and power. Furthermore, the sample size for a test procedure has to be determined so that the Type I error rates never exceed the nominal level, and the power of the test should be at least a predetermined value. As the test procedures we discussed earlier are exact, the Type I error rates are always within the nominal level regardless of values of ρ_0^2 being tested; however, the powers of the tests are affected if the true value of ρ^2 is misspecified.

Sample size calculation for constructing confidence intervals should be carried out with respect to following two criteria: (a) The coverage probability of the interval estimation procedure should be at least the nominal confidence level $1-\alpha$ regardless of values of the parameter, and (b) for a given input parameter, the expected width of the confidence interval should not be more than a specified value. Notice that if the input parameter is misspecified, then the calculated sample size according to the aforementioned criteria may not guarantee the expected width, but it does guarantee the coverage probability. We note here that our criteria for computing sample size are different from those used by Algina and Olejnik (2003) for the bivariate normal case. We show in the sequel that the sample size determined by Algina and Olejnik's method does not guarantee the coverage probability requirement.

In the following section, we provide some preliminary results and an expression for the CDF of R^2, which is used to outline inferential procedures for ρ^2. Then, we describe hypothesis tests and confidence intervals that can be obtained by inverting the tests. As we mentioned earlier, the inferential procedures for ρ^2 are not even mentioned in many standard textbooks in multivariate analysis, let alone their properties. We show that the one-sided tests are uniformly most powerful (UMP), and so the one-sided confidence intervals (which are obtained by inverting the UMP tests) are uniformly most accurate. We describe a method of computing sample sizes for the one-sided tests described earlier to attain a specified power and a method of calculating sample sizes for constructing confidence intervals with expected width not more than a specified value. Finally, we illustrate our sample size calculation using a numerical example and construction of confidence intervals and hypothesis testing using the empirical data from Timm (1975) that related scores from the Peabody Picture Vocabulary Test to four proficiency measures.

SOME PRELIMINARIES

Let R^2 denote the squared sample multiple correlation coefficient based on a sample of n vector observations each of dimension p. The CDF of R^2 can be expressed in many different forms. The one due to Gurland (1968) seems to be relatively easier to compute than the others and is given by

$$P(R^2 \leq x|n, p, \rho^2) = \sum_{i=0}^{\infty} P(Y = i) I_x \left(\frac{p-1}{2} + i, \frac{n-p}{2} \right), \tag{2}$$

where

$$I_x(a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x t^{a-1}(1-t)^{b-1} dt$$

is the incomplete beta function, and

$$P(Y = i) = \frac{\Gamma\left(\frac{n+1}{2} + i\right)}{\Gamma(i+1)\Gamma\left(\frac{n+1}{2}\right)} (\rho^2)^i (1-\rho^2)^{\frac{n+1}{2}}. \tag{3}$$

The expression in Equation (3), whether or not $(n+1)/2$ is an integer, can be regarded as the negative binomial probability mass function with success probability $(1-\rho^2)$, the number of successes $(n+1)/2$, and the number of failures i (see Muirhead 1982, p. 175). Ding and Bargmann (1991, *Applied Statistics* Algorithm AS 260) provided an algorithm and Fortran code to compute the CDF in Equation (2). Benton and Krishnamoorthy (2003) obtained an accurate and efficient algorithm by enhancing Algorithm AS 260. This algorithm is coded in Fortran and used to carry out all the computations in the sequel.

TEST AND CONFIDENCE INTERVAL AND THEIR PROPERTIES

The CDF of R^2 is decreasing in ρ^2 (see Appendix A), and so exact confidence limits for ρ^2 can be obtained using a method similar to the Clopper-Pearson approach for finding exact limits for a binomial proportion. For a more general result of obtaining confidence intervals using the CDF, see Casella and Berger (2002, Theorem 9.2.12). Application of these techniques for estimating ρ^2 dates back to Ezekiel and Fox (1959); however, as pointed out by Smithson (2001), until recently they have not been widely available due to their neglect in popular multivariate statistical textbooks and software packages. For example, the book by Anderson (1984, Section 4.2) explains the calculation of confidence intervals by pivoting the CDF of R^2 for the bivariate normal case but makes no mention of the general case.

Let r^2 be an observed value of R^2 based on a sample of size n. Suppose we are interested in testing Equation (1). For a given nominal level α, the test that rejects H_0 when the p value $P(R^2 \geq r^2|n, p, \rho_0^2) < \alpha$ is a size α test; that is, the maximum probability of rejecting the null hypothesis when it is true is α. As the probability density function (PDF) of R^2 has monotone likelihood ratio property (see Appendix A), this test is UMP in the class of tests of Equation (1) that are based on R^2. For testing $H_0 : \rho^2 \geq \rho_0^2$ vs. $H_a : \rho^2 < \rho_0^2$, the test that rejects H_0 when the p value $P(R^2 \leq r^2|n, p, \rho_0^2) < \alpha$ is a UMP test. It should be noted that, in general, a UMP test exists only when hypotheses are one-sided (see Casella & Berger, 2002, p. 391), and the existence of a UMP test for a two-sided hypothesis about ρ^2 has not been established.

A $1 - \alpha$ lower limit for ρ^2 is the minimum value of ρ_0^2 for which the null hypothesis in Equation (1) is accepted. That is, the minimum value of ρ_0^2 for which the p value is at least α, or equivalently, $\min\{\rho_0^2 : P(R^2 \geq r^2|n, p, \rho_0^2) \geq \alpha\}$. For a given n, p and r^2, $P(R^2 \geq r^2|n, p, \rho^2)$ is increasing in ρ^2 (see Appendix A), and so this minimum, say, ρ_L^2 is the solution of the equation

$$P(R^2 \geq r^2|n, p, \rho_L^2) = \alpha. \tag{4}$$

Using similar arguments, one can obtain an upper limit by inverting a left-tail test. Specifically, a $1 - \alpha$ upper limit ρ_U^2 for ρ^2 is the solution of the equation

$$P(R^2 \leq r^2|n, p, \rho_U^2) = \alpha. \tag{5}$$

Two-sided limits can be obtained by replacing α in the previous equations by $\alpha/2$ and then solving them for ρ_L^2 and ρ_U^2.

These one-sided confidence limits given earlier are uniformly most accurate (the confidence interval that has the minimum probability of covering false values of ρ^2) because they are obtained by inverting the one-sided UMP tests (see Casella & Berger, 2002, Theorem 9.3.5).

SAMPLE SIZE CALCULATION FOR POWER

Suppose it is desired to test (right-tail test)

$$H_0 : \rho^2 \le \rho_0^2 \ \text{ vs. } \ H_a : \rho^2 > \rho_0^2, \tag{6}$$

where ρ_0^2 is a specified value, at a nominal level α. The sample size to attain a power of β at a specified value ρ_1^2 under H_a is the least value of n for which

$$P(R^2 \ge k_n | n, p, \rho_0^2) = \alpha \ \text{ and } \ \text{power} = P(R^2 \ge k_n | n, p, \rho_1^2) \ge \beta. \tag{7}$$

Notice that the critical value of k_n is to be determined so that the Type I error rate is α and the power is at least β.

The sample size n that satisfies the probability requirements in Equation (7) was computed as follows. For a given ρ_0^2, p, α, and an initial small trial value of n, we found the critical value k_n that satisfies $P(R^2 \ge k_n | n, p, \rho_0^2) = \alpha$ using a root finding method and then computed the power in Equation (7). By incrementing the value of n and repeating the computation, we found the least value of n for which the power is at least β.

The sample sizes for testing Equation (6) at the level 0.05 and power of at least 0.80 were reported for values of $p = 2(1)10$ in Table 1. The number below the sample size is the critical value k_n that satisfies both conditions in Equation (7). For example, when $p = 4$ one wants to test $H_0 : \rho^2 \le 0.8$ vs. $H_a : \rho^2 > 0.8$ at $\alpha = 0.05$. Suppose she or he also believes that the true value of ρ_1^2 is 0.9. Then the required sample size to attain a power of not less than 0.80 is 48, and the critical value is 0.8808. Furthermore, these values satisfy the equations in Equation (7). That is,

$$P(R^2 \ge .8808 | 48, 4, .8) = 0.05 \ \text{ and } \ \text{power} = P(R^2 \ge .8808 | 48, 4, .9) = 0.8032.$$

If a sample of 48 observations produces a value of R^2 exceeding 0.8808, then $H_0 : \rho^2 \le 0.8$ will be rejected at level 0.05.

Suppose one is interested in testing (left-tail test)

$$H_0 : \rho^2 \ge \rho_0^2 \ \text{ vs. } \ H_a : \rho^2 < \rho_0^2, \tag{8}$$

then the required sample sizes were reported in the lower triangular part of Table 1. For example, when $p = 4$ one wants to test $H_0 : \rho^2 \ge 0.9$ vs. $H_a : \rho^2 < 0.9$ at $\alpha = 0.05$. Suppose the true value of ρ^2 is $\rho_1^2 = 0.8$. Then the required sample size to attain a power of 0.80 is 48, and the critical value is 0.8514. We again note that these values satisfy equations

$$P(R^2 \le k_n | n, p, \rho_0^2) = \alpha \ \text{ and } \ \text{power} = P(R^2 \le k_n | n, p, \rho_1^2) \ge \beta.$$

TABLE 1
Sample Sizes n and Critical Values k_n for Testing H_0: $\rho^2 \le \rho_0^2$ vs. $H_a: \rho^2 > \rho_0^2$ (Entries Above the Diagonal) $H_0: \rho^2 \ge \rho_0^2$ vs. $H_a: \rho^2 < \rho_0^2$ (Entries Below the Diagonal) With Power $\ge$0.80 and $\alpha = 0.05$

$p = 2$		ρ_1^2									
ρ_0^2		.1	.2	.3	.4	.5	.6	.7	.8	.9	.95
.0	n	76	36	23	17	13	10	8	6	5	4
	k_n	0.0509	0.1083	0.1708	0.2325	0.3057	0.3993	0.4995	0.6584	0.7715	0.9025
.1	n	—	264	77	38	23	15	11	8	6	5
	k_n		0.1642	0.2287	0.2957	0.3667	0.4503	0.5296	0.6328	0.7480	0.8295
.2	n	264	—	347	91	41	23	14	9	6	5
	k_n	0.1317		0.2660	0.3337	0.4059	0.4836	0.5753	0.6830	0.8076	0.8698
.3	n	77	347	—	366	90	38	20	12	7	5
	k_n	0.1628	0.2336		0.3669	0.4368	0.5131	0.5972	0.6876	0.8108	0.8997
.4	n	38	91	366	—	337	78	31	15	8	6
	k_n	0.1973	0.2686	0.3346		0.4677	0.5401	0.6209	0.7151	0.8255	0.8846
.5	n	23	41	90	337	—	276	60	22	10	7
	k_n	0.2383	0.3074	0.3724	0.4353		0.5685	0.6436	0.7308	0.8312	0.8867
.6	n	15	23	38	78	276	—	197	39	13	8
	k_n	0.2847	0.3547	0.4155	0.4757	0.5362		0.6696	0.7492	0.8439	0.8992
.7	n	10	14	20	31	60	197	—	116	20	9
	k_n	0.3368	0.4110	0.4703	0.5243	0.5801	0.6374		0.7713	0.8572	0.9186
.8	n	8	9	12	15	22	39	116	—	47	14
	k_n	0.4594	0.4907	0.5522	0.5891	0.6374	0.6865	0.7395		0.8751	0.9251
.9	n	5	6	7	8	10	13	19	47	—	51
	k_n	0.5742	0.6340	0.6727	0.6998	0.7357	0.7672	0.7998	0.8446		0.9376
.95	n	4	5	5	6	6	7	9	14	51	—
	k_n	0.7071	0.7699	0.7699	0.8048	0.8048	0.8269	0.8535	0.8836	0.9226	

(continued)

TABLE 1
(*Continued*)

$p = 3$		ρ_1^2									
ρ_0^2		.1	.2	.3	.4	.5	.6	.7	.8	.9	.95
.0	n	93	45	29	21	16	12	10	8	6	6
	k_n	0.0644	0.1329	0.2058	0.2831	0.3693	0.4861	0.5751	0.6983	0.8643	0.8643
.1	n	—	267	80	40	25	17	12	9	7	6
	k_n		0.1674	0.2386	0.3162	0.3966	0.4880	0.5973	0.7099	0.8218	0.8904
.2	n	267	—	349	93	43	24	16	11	7	6
	k_n	0.1350		0.2681	0.3409	0.4197	0.5111	0.5987	0.7024	0.8570	0.9115
.3	n	80	349	—	368	91	39	21	13	8	6
	k_n	0.1742	0.2358		0.3686	0.4434	0.5272	0.6204	0.7192	0.8469	0.9289
.4	n	40	93	368	—	339	79	32	16	9	7
	k_n	0.2194	0.2770	0.3364		0.4692	0.5462	0.6339	0.7362	0.8516	0.9091
.5	n	25	43	91	339	—	277	61	23	11	7
	k_n	0.2758	0.3264	0.3794	0.4371		0.5701	0.6496	0.7435	0.8496	0.9293
.6	n	17	24	39	79	277	—	198	40	14	8
	k_n	0.3441	0.3846	0.4317	0.4826	0.5380		0.6713	0.7556	0.8561	0.9286
.7	n	12	15	21	32	61	198	—	117	20	10
	k_n	0.4293	0.4599	0.5002	0.5411	0.5876	0.6393		0.7733	0.8678	0.9276
.8	n	9	10	12	16	23	40	117	—	47	15
	k_n	0.5500	0.5650	0.5892	0.6225	0.6567	0.6955	0.7419		0.8784	0.9301
.9	n	6	7	8	9	10	13	20	47	—	51
	k_n	0.7114	0.7277	0.7419	0.7537	0.7635	0.7852	0.8128	0.8477		0.9391
.95	n	5	6	6	6	7	8	10	15	51	—
	k_n	0.8373	0.8468	0.8468	0.8468	0.8564	0.8646	0.8769	0.8946	0.9240	

(*continued*)

TABLE 1
(*Continued*)

$p = 4$		ρ_1^2									
ρ_0^2		.1	.2	.3	.4	.5	.6	.7	.8	.9	.95
.0	n	105	51	33	23	18	14	11	9	8	7
	k_n	0.0741	0.1517	0.2328	0.3306	0.4174	0.5266	0.6507	0.7645	0.8318	0.9027
.1	n	—	271	82	42	26	18	14	10	8	7
	k_n		0.1705	0.2489	0.3342	0.4293	0.5307	0.6182	0.7591	0.8613	0.9189
.2	n	271	—	351	94	44	26	17	12	8	7
	k_n	0.1383		0.2702	0.3488	0.4354	0.5275	0.6304	0.7384	0.8859	0.9328
.3	n	83	351	—	369	92	41	22	14	9	7
	k_n	0.1849	0.2380		0.3704	0.4499	0.5371	0.6410	0.7450	0.8714	0.9448
.4	n	42	94	369	—	340	80	33	17	10	8
	k_n	0.2395	0.2846	0.3382		0.4708	0.5521	0.6460	0.7543	0.8710	0.9250
.5	n	26	44	92	340	—	278	62	24	12	8
	k_n	0.3049	0.3421	0.3863	0.4388		0.5716	0.6554	0.7550	0.8644	0.9410
.6	n	18	26	40	80	278	—	199	41	14	9
	k_n	0.3857	0.4149	0.4469	0.4894	0.5396		0.6729	0.7616	0.8756	0.9382
.7	n	13	16	22	33	62	199	—	117	21	11
	k_n	0.4852	0.5015	0.5270	0.5568	0.5948	0.6412		0.7755	0.8742	0.9349
.8	n	10	11	13	17	24	40	117	—	48	16
	k_n	0.6151	0.6208	0.6321	0.6510	0.6740	0.7028	0.7440		0.8808	0.9345
.9	n	7	8	9	10	11	14	21	48	—	52
	k_n	0.7838	0.7845	0.7877	0.7916	0.7957	0.8067	0.8243	0.8514		0.9402
.95	n	6	7	7	7	8	9	11	15	52	—
	k_n	0.8900	0.8866	0.8866	0.8866	0.8873	0.8893	0.8941	0.9021	0.9257	

(*continued*)

TABLE 1
(*Continued*)

$p=5$		ρ_1^2									
ρ_0^2		.1	.2	.3	.4	.5	.6	.7	.8	.9	.95
.0	n	115	56	36	26	20	16	13	11	9	8
	k_n	0.0819	0.1669	0.2569	0.3511	0.4490	0.5497	0.6574	0.7514	0.8646	0.9240
.1	n	—	274	85	44	28	20	15	11	9	8
	k_n		0.1735	0.2574	0.3501	0.4488	0.5489	0.6564	0.7937	0.8863	0.9355
.2	n	274	—	353	96	46	27	18	13	9	8
	k_n	0.1414		0.2723	0.3555	0.4466	0.5493	0.6575	0.7665	0.9051	0.9458
.3	n	85	353	—	371	94	42	23	15	10	8
	k_n	0.1944	0.2402		0.3720	0.4553	0.5495	0.6596	0.7665	0.8892	0.9549
.4	n	44	96	371	—	341	81	34	18	11	9
	k_n	0.2578	0.2925	0.3401		0.4724	0.5579	0.6574	0.7702	0.8859	0.9362
.5	n	28	46	93	341	—	279	63	25	13	9
	k_n	0.3335	0.3583	0.3930	0.4404		0.5732	0.6610	0.7655	0.8765	0.9494
.6	n	20	27	41	81	279	—	200	42	15	10
	k_n	0.4244	0.4388	0.4614	0.4960	0.5413		0.6746	0.7674	0.8847	0.9455
.7	n	14	18	23	34	62	200	—	118	22	12
	k_n	0.5314	0.5395	0.5511	0.5715	0.6010	0.6430		0.7773	0.8800	0.9408
.8	n	11	12	14	18	24	41	118	—	49	17
	k_n	0.6642	0.6642	0.6670	0.6756	0.6880	0.7111	0.7463		0.8832	0.9384
.9	n	8	9	9	10	12	15	21	48	—	52
	k_n	0.8281	0.8222	0.8222	0.8201	0.8204	0.8244	0.8332	0.8544		0.9417
.95	n	7	8	8	8	9	10	11	16	52	—
	k_n	0.9177	0.9105	0.9105	0.9105	0.9076	0.9067	0.9067	0.9105	0.9271	

(*continued*)

TABLE 1
(*Continued*)

$p = 6$ ρ_0^2		ρ_1^2 .1	.2	.3	.4	.5	.6	.7	.8	.9	.95
.0	n	124	60	39	28	22	17	14	12	10	9
	k_n	0.0885	0.1810	0.2749	0.3769	0.4713	0.5929	0.6974	0.7852	0.8866	0.9376
.1	n	—	277	87	46	30	21	16	13	10	9
	k_n		0.1766	0.2664	0.3643	0.4648	0.5790	0.6875	0.7817	0.9036	0.9465
.2	n	278	—	355	98	47	28	19	14	11	9
	k_n	0.1446		0.2744	0.3618	0.4604	0.5692	0.6809	0.7892	0.8828	0.9546
.3	n	88	355	—	372	95	43	24	16	11	9
	k_n	0.2039	0.2423		0.3738	0.4615	0.5612	0.6763	0.7847	0.9026	0.9619
.4	n	46	98	372	—	342	82	35	19	12	10
	k_n	0.2746	0.3000	0.3418		0.4740	0.5635	0.6680	0.7841	0.8978	0.9445
.5	n	30	47	95	342	—	280	63	26	13	10
	k_n	0.3582	0.3723	0.3999	0.4421		0.5747	0.6679	0.7751	0.9019	0.9557
.6	n	21	28	43	82	280	—	201	42	16	11
	k_n	0.4553	0.4608	0.4761	0.5024	0.5430		0.6762	0.7750	0.8925	0.9513
.7	n	16	19	24	35	63	200	—	119	23	13
	k_n	0.5688	0.5685	0.5728	0.5853	0.6079	0.6448		0.7792	0.8853	0.9457
.8	n	12	13	15	19	25	42	119	—	49	17
	k_n	0.7024	0.6989	0.6959	0.6970	0.7030	0.7189	0.7486		0.8863	0.9458
.9	n	9	10	10	11	13	15	22	49	—	53
	k_n	0.8577	0.8490	0.8490	0.8441	0.8399	0.8390	0.8428	0.8578		0.9427
.95	n	8	9	9	9	10	10	12	17	53	—
	k_n	0.9346	0.9263	0.9263	0.9263	0.9219	0.9219	0.9182	0.9177	0.9287	

(*continued*)

TABLE 1
(*Continued*)

$p = 7$ ρ_0^2		ρ_1^2 .1	.2	.3	.4	.5	.6	.7	.8	.9	.95
.0	n	132	64	41	30	23	19	15	13	11	10
	k_n	0.0944	0.1924	0.2958	0.3974	0.5069	0.5997	0.7287	0.8107	0.9024	0.9470
.1	n	—	280	90	48	31	22	17	14	11	10
	k_n		0.1795	0.2736	0.3769	0.4876	0.6051	0.7134	0.8039	0.9163	0.9542
.2	n	281	—	357	99	48	29	20	15	12	10
	k_n	0.1475		0.2764	0.3690	0.4736	0.5873	0.7012	0.8078	0.8961	0.9609
.3	n	90	357	—	373	96	44	26	17	12	10
	k_n	0.2126	0.2444		0.3755	0.4675	0.5723	0.6817	0.8003	0.9132	0.9670
.4	n	48	99	373	—	343	83	36	20	13	11
	k_n	0.2900	0.3069	0.3436		0.4755	0.5690	0.6780	0.7965	0.9074	0.9509
.5	n	31	48	96	343	—	281	64	27	14	11
	k_n	0.3798	0.3858	0.4063	0.4437		0.5762	0.6732	0.7841	0.9100	0.9606
.6	n	22	29	44	83	281	—	201	43	17	12
	k_n	0.4830	0.4812	0.4888	0.5086	0.5447		0.6780	0.7803	0.8994	0.9560
.7	n	17	20	25	36	64	201	—	119	23	14
	k_n	0.5996	0.5941	0.5926	0.5982	0.6145	0.6466		0.7813	0.8944	0.9499
.8	n	13	14	16	20	26	43	119	—	50	18
	k_n	0.7330	0.7272	0.7204	0.7158	0.7167	0.7263	0.7507		0.8885	0.9488
.9	n	10	11	11	12	14	16	23	50	—	54
	k_n	0.8788	0.8690	0.8690	0.8627	0.8557	0.8525	0.8514	0.8611		0.9437
.95	n	9	9	10	10	10	11	13	17	53	—
	k_n	0.9459	0.9459	0.9375	0.9375	0.9375	0.9325	0.9272	0.9240	0.9300	

(*continued*)

TABLE 1
(*Continued*)

$p = 8$		ρ_1^2									
ρ_0^2		.1	.2	.3	.4	.5	.6	.7	.8	.9	.95
.0	n	140	68	44	32	25	20	17	14	12	11
	k_n	0.0993	0.2018	0.3069	0.4140	0.5184	0.6296	0.7192	0.8307	0.9143	0.9540
.1	n	—	283	92	50	33	24	18	15	12	11
	k_n		0.1824	0.2817	0.3883	0.4987	0.6111	0.7351	0.8220	0.9260	0.9600
.2	n	284	—	359	101	50	31	21	16	13	11
	k_n	0.1505		0.2784	0.3749	0.4821	0.5947	0.7191	0.8234	0.9067	0.9657
.3	n	92	359	—	375	97	45	27	18	13	11
	k_n	0.2209	0.2465		0.3771	0.4734	0.5829	0.6953	0.8137	0.9217	0.9709
.4	n	50	101	375	—	344	84	37	21	14	12
	k_n	0.3043	0.3140	0.3454		0.4771	0.5744	0.6875	0.8075	0.9153	0.9560
.5	n	33	50	97	344	—	282	65	28	15	12
	k_n	0.3994	0.3988	0.4125	0.4453		0.5777	0.6784	0.7923	0.9169	0.9645
.6	n	23	30	45	84	282	—	202	44	18	13
	k_n	0.5081	0.5001	0.5010	0.5147	0.5463		0.6796	0.7854	0.9054	0.9598
.7	n	18	21	26	37	65	202	—	120	24	15
	k_n	0.6263	0.6169	0.6107	0.6103	0.6209	0.6484		0.7831	0.8989	0.9535
.8	n	14	15	17	20	27	43	120	—	51	19
	k_n	0.7579	0.7507	0.7412	0.7338	0.7291	0.7330	0.7529		0.8906	0.9515
.9	n	11	12	12	13	14	17	23	50	—	54
	k_n	0.8946	0.8844	0.8844	0.8773	0.8723	0.8640	0.8593	0.8640		0.9451
.95	n	10	10	11	11	11	12	14	18	54	—
	k_n	0.9539	0.9539	0.9459	0.9459	0.9459	0.9406	0.9345	0.9296	0.9315	

(*continued*)

TABLE 1
(*Continued*)

$p = 9$ ρ_0^2		ρ_1^2 .1	.2	.3	.4	.5	.6	.7	.8	.9	.95
.0	n	146	71	46	34	26	21	18	15	13	12
	k_n	0.1049	0.2125	0.3224	0.4279	0.5453	0.6551	0.7417	0.8468	0.9236	0.9593
.1	n	—	287	94	51	34	25	20	16	13	12
	k_n		0.1850	0.2893	0.4033	0.5176	0.6316	0.7297	0.8371	0.9337	0.9645
.2	n	287	—	361	103	51	32	22	17	14	12
	k_n	0.1533		0.2804	0.3805	0.4938	0.6099	0.7350	0.8366	0.9154	0.9694
.3	n	95	361	—	376	99	46	28	19	14	12
	k_n	0.2289	0.2486		0.3788	0.4779	0.5930	0.7079	0.8255	0.9286	0.9740
.4	n	52	103	376	—	345	85	38	22	15	13
	k_n	0.3175	0.3208	0.3471		0.4786	0.5796	0.6964	0.8174	0.9221	0.9601
.5	n	34	51	98	345	—	283	66	28	16	13
	k_n	0.4177	0.4109	0.4186	0.4469		0.5792	0.6834	0.8063	0.9228	0.9677
.6	n	25	31	46	85	282	—	203	45	19	14
	k_n	0.5280	0.5177	0.5125	0.5207	0.5479		0.6811	0.7902	0.9107	0.9631
.7	n	19	22	27	38	66	203	—	121	25	16
	k_n	0.6497	0.6373	0.6272	0.6217	0.6271	0.6502		0.7849	0.9030	0.9566
.8	n	15	16	18	21	28	44	121	—	52	20
	k_n	0.7787	0.7706	0.7592	0.7493	0.7406	0.7399	0.7551		0.8926	0.9540
.9	n	12	13	13	14	15	18	24	51	—	55
	k_n	0.9068	0.8966	0.8966	0.8892	0.8837	0.8738	0.8666	0.8670		0.9461
.95	n	11	11	12	12	12	13	15	19	54	—
	k_n	0.9599	0.9599	0.9523	0.9523	0.9523	0.9470	0.9404	0.9345	0.9329	

(*continued*)

TABLE 1
(*Continued*)

$p = 10$		ρ_1^2									
ρ_0^2		.1	.2	.3	.4	.5	.6	.7	.8	.9	.95
.0	n	153	74	48	35	28	22	19	16	14	13
	k_n	0.1091	0.2221	0.3361	0.4510	0.5512	0.6771	0.7607	0.8601	0.9310	0.9636
.1	n	—	290	96	53	36	26	21	17	14	13
	k_n		0.1877	0.2966	0.4127	0.5255	0.6499	0.7466	0.8497	0.9400	0.9681
.2	n	291	—	363	104	52	33	24	18	15	13
	k_n	0.1562		0.2823	0.3871	0.5050	0.6240	0.7330	0.8480	0.9225	0.9724
.3	n	97	363	—	378	100	47	29	20	15	13
	k_n	0.2365	0.2507		0.3804	0.4835	0.6026	0.7194	0.8359	0.9345	0.9765
.4	n	53	104	378	—	346	86	39	23	16	14
	k_n	0.3302	0.3272	0.3488		0.4801	0.5847	0.7048	0.8263	0.9278	0.9635
.5	n	35	52	100	346	—	284	67	29	17	14
	k_n	0.4350	0.4226	0.4246	0.4485		0.5807	0.6882	0.8134	0.9279	0.9704
.6	n	26	33	47	86	283	—	204	46	20	15
	k_n	0.5480	0.5326	0.5236	0.5265	0.5495		0.6827	0.7948	0.9155	0.9658
.7	n	20	23	28	39	67	203	—	122	26	16
	k_n	0.6704	0.6556	0.6424	0.6325	0.6330	0.6519		0.7866	0.9068	0.9670
.8	n	16	17	19	22	29	45	121	—	52	21
	k_n	0.7962	0.7875	0.7749	0.7632	0.7511	0.7464	0.7571		0.8956	0.9562
.9	n	13	14	14	15	16	19	25	52	—	56
	k_n	0.9165	0.9065	0.9065	0.8990	0.8933	0.8824	0.8732	0.8700		0.9470
.95	n	12	12	13	13	13	14	15	20	55	—
	k_n	0.9645	0.9645	0.9573	0.9573	0.9573	0.9522	0.9484	0.9387	0.9343	

Specifically,

$$P(R^2 \leq .8514|48, 4, .9) = 0.05 \text{ and power } =$$

$$P(R^2 \leq .8514|48, 4, .8) = 0.8051.$$

Thus, if a sample of 48 observations produces a value of R^2 that is less than 0.8514, then $H_0 : \rho^2 \geq 0.9$ will be rejected at level 0.05.

It is interesting to note that the right-tail test when $(\rho_0^2, \rho_1^2) = (.8, .9)$ and the left-tail test when $(\rho_0^2, \rho_1^2) = (.9, .8)$ require the same sample size of 48 to attain a power of at least 0.80. We observe from Table 1 that, in most cases, the sample size required for the right-tail test at $(\rho_0^2 = x, \rho_1^2 = y)$ is the same as the one for the left-tail test at $(\rho_0^2 = y, \rho_1^2 = x)$ even though the distribution of R^2 is asymmetric. In other situations the sample size for the right-tail test at $(\rho_0^2 = x, \rho_1^2 = y)$ is typically one unit more than the sample size required for the left-tail test at $(\rho_0^2 = y, \rho_1^2 = x)$.

In the next section we provide methods of computing the sample size required to construct confidence intervals with expected width not more than a specified value.

SAMPLE SIZE CALCULATION FOR CONSTRUCTING CONFIDENCE INTERVALS

In general, for a given p, ρ^2, and confidence coefficient $1 - \alpha$, the sample size is to be determined so that the expected width of a random confidence interval should not exceed a prespecified quantity, say w. Notice that a $1 - \alpha$ confidence interval (ρ_L^2, ρ_U^2) for ρ^2 is determined by

$$P(R^2 \geq r^2|n, p, \rho_L^2) = \frac{\alpha}{2} \quad \text{and} \quad P(R^2 \leq r^2|n, p, \rho_U^2) = \frac{\alpha}{2}, \tag{9}$$

where r^2 is an observed value of R^2. It is clear from Equation (9) that ρ_L^2 and ρ_U^2 are implicitly functions of r^2. We usually refer to the interval $(\rho_L^2(r^2), \rho_U^2(r^2))$ is an observed value or a realization of the random interval $(\rho_L^2(R^2), \rho_U^2(R^2))$. Furthermore, the minimum (with respect to ρ^2) coverage probability of the random interval $(\rho_L^2(R^2), \rho_U^2(R^2))$ is $1 - \alpha$. For a given ρ^2 and a specified width w, we need to determine the least value of n for which

$$E(\rho_U^2(R^2) - \rho_L^2(R^2)) = \int_0^1 (\rho_U^2(x) - \rho_L^2(x)) f(x|n, p, \rho^2) dx \leq w, \tag{10}$$

where f denotes the PDF of R^2. Determining the smallest value of n using Equation (10) is very difficult, and so we resorted to finding the value of n using a different approach, which uses Algina and Olejnik's (2003) method to find an initial search value of n and Monte Carlo simulation to estimate the expected width of $(\rho_L^2(R^2), \rho_U^2(R^2))$.

For the bivariate normal case, Algina and Olejnik's (2003) computed the least sample size required for the interval of the form $R^2 \pm c$ that would contain the true parameter ρ^2 with probability of at least $1-\alpha$. Specifically, for a given ρ^2, c, and $1-\alpha$, these authors were interested in finding the smallest value of n for which

$$P\left[\max\{0, \rho^2 - c\} \le R^2 \le \min\{\rho^2 + c, 1\} \middle| n, p, \rho^2\right] \ge 1 - \alpha. \qquad (11)$$

Notice that the value of n that satisfies Equation (11) guarantees that the random interval

$$\left(\max\{0, R^2 - c\}, \min\{R^2 + c, 1\}\right) \qquad (12)$$

would contain ρ^2 with probability at least $1-\alpha$. However, it should be pointed out that the inequality in Equation (11) holds only at the specified value of the parameter. The interval Equation (12) is not a $1-\alpha$ confidence interval because its coverage probability depends on the value ρ^2, and its minimum (with respect to ρ^2) coverage probability could be well below the nominal level $1-\alpha$. For example, when $p = 2$, $\rho^2 = .70$, from Table 2 of Algina and Olejnik (2003), we found a sample of 98 observations was required to estimate ρ^2 within $\pm.1$. Indeed, we calculated

$$P(\rho^2 - .1 \le R^2 \le \rho^2 + .1 | n = 98, p = 2, \rho^2 = .7) = 0.95,$$

or equivalently,

$$P(R^2 - .1 \le \rho^2 \le R^2 + .1 | n = 98, p = 2, \rho^2 = .7) = 0.95.$$

Suppose that ρ^2 is misspecified while its actual value is 0.6, then

$$P(R^2 - .1 \le \rho^2 \le R^2 + .1 | n = 98, p = 2, \rho^2 = .6) = 0.90;$$

if the actual value of ρ^2 is 0.5, then this probability is 0.84. Thus, we see that the coverage probability of the interval of the form $R^2 \pm c$ decreases as ρ^2 decreases. On the other hand, the random interval $(\rho_L^2(R^2), \rho_U^2(R^2))$ has coverage probability $1-\alpha$ regardless of values of ρ^2.

Even though there is no apparent relation between the values of n that satisfy Equation (10) and Equation (11), surprisingly we found that the least value of n that satisfies the latter (with $c = w/2$) is close to the one that satisfies the former in cases where large sample is required. Using this fact, we computed the smallest sample size so that the left-hand side of Equation (10) is close to w as follows.

For a given p, ρ^2, $1-\alpha$, and w,

1. Determine the smallest n that satisfies Equation (11) with $c = w/2$; call this value n_0.
2. Using the n_0 in Step 1 for n, estimate the expected width of the interval $(\rho_L^2(R^2), \rho_U^2(R^2))$ using Monte Carlo simulation consisting of 10,000 runs.
3. If the expected width in Step 2 is less (or greater) than w, then search for a value n smaller (or greater) than n_0 until the estimated expected width is close to but not more than w.

The value n_0 in Step 1 can be obtained as the smallest value of n for which

$$P\left(R^2 \leq \min\{\rho^2 + c, 1\} \middle| n, p, \rho^2\right) - P\left(R^2 \leq \max\{0, \rho^2 - c, \} \middle| n, p, \rho^2\right) \geq 1-\alpha. \tag{13}$$

We used the algorithm for computing the CDF of R^2 by Benton and Krishnamoorthy (2003) to compute the CDFs in Equation (13). Notice that Equation (11) and Equation (13) are equivalent. For a given ρ^2, p, c, and confidence level $1-\alpha$, the value of n was searched, starting with a small initial value of n, until the left-hand side of Equation (13) exceeds $1-\alpha$.

The Monte Carlo estimate of the expected width in Equation (10) was obtained as follows. For a given n, ρ^2, p, and confidence coefficient $1-\alpha$, we generated 10,000 values of R^2 (see Appendix B for an efficient way of generating R^2). Treating each generated value as r^2, we found the $1-\alpha$ confidence interval using Equation (9). The average width of these 10,000 confidence intervals was used as a Monte Carlo estimate of w in Equation (10).

For $p = 2(1)10$, 15, and 20, we presented the required sample sizes for constructing 95% confidence intervals with expected widths no more than .4, .3, .2, .1, and .05 in Table 2. For example, if a researcher expects that the population ρ^2 is around 0.80 when the correlation analysis involves $p = 6$ variables, and she or he wants to construct a 95% confidence interval for ρ^2 with expected width not exceeding 0.1, then the required sample size is 206. We also observe that the sample size required to estimate small ρ^2 with a specified expected width is much larger than the one required for large ρ^2. For instance, when $p = 4$ and $\rho^2 = 0.3$, a sample of 906 observations is needed to construct confidence interval for ρ^2 with expected width not more than 0.1; when $\rho^2 = 0.7$, only a sample of 393 observations is needed. Furthermore, we see that the sample size increases with increasing p. So, for the values of p not listed in the table, an interpolation can be used. For instance, if one wants to determine the sample

TABLE 2
Sample Sizes Required for 95% Confidence Intervals
With Expected Widths at Most *w*

	ρ^2				$p = 2$				
w	.2	.3	.4	.5	.6	.7	.8	.9	.95
.4	37	48	50	45	37	27	17	9	6
.3	77	93	94	84	66	46	27	13	8
.2	189	222	217	192	148	100	55	21	11
.1	784	902	880	768	591	389	199	62	23
.05	3144	3612	3536	3073	2361	1550	789	226	67
					$p = 3$				
.4	40	50	51	47	38	28	18	10	7
.3	80	95	95	85	67	47	28	14	9
.2	192	224	218	193	149	101	56	22	12
.1	788	903	882	769	592	392	204	64	24
.05	3148	3613	3537	3076	2362	1553	793	232	69
					$p = 4$				
.4	42	52	53	48	39	29	19	11	9
.3	82	96	96	86	68	48	29	15	10
.2	195	226	221	194	150	102	57	23	13
.1	795	906	885	770	593	393	204	66	25
.05	3155	3616	3540	3078	2363	1554	793	232	70
					$p = 5$				
.4	45	54	54	49	40	30	20	12	10
.3	84	98	98	87	69	49	30	16	11
.2	197	227	224	195	151	103	58	24	14
.1	803	911	888	771	593	394	205	67	26
.05	3164	3621	3543	3079	2364	1556	795	232	71
					$p = 6$				
.4	46	55	56	50	41	31	21	13	10
.3	87	100	99	88	70	50	31	16	12
.2	200	229	225	196	152	104	59	25	14
.1	815	917	891	772	595	394	206	68	27
.05	3175	3627	3546	3080	2366	1557	796	233	72
					$p = 7$				
.4	48	57	57	52	42	32	22	14	11
.3	89	103	101	90	71	51	32	17	13
.2	204	232	227	197	153	104	60	26	15
.1	828	924	896	774	596	395	207	69	27
.05	3189	3635	3551	3081	2367	1558	797	234	72

(*continued*)

TABLE 2
(*Continued*)

	ρ^2				$p = 8$				
w	.2	.3	.4	.5	.6	.7	.8	.9	.95
.4	50	59	59	53	43	33	23	15	12
.3	92	104	103	91	72	51	32	18	14
.2	206	234	229	201	154	105	60	26	16
.1	841	933	902	778	597	396	208	70	28
.05	3205	3645	3557	3083	2368	1560	798	235	72
					$p = 9$				
.4	51	60	60	54	44	34	24	16	13
.3	94	107	104	92	73	52	33	19	15
.2	209	236	231	202	155	106	61	27	17
.1	843	944	909	781	598	397	209	71	29
.05	3224	3656	3565	3088	2369	1561	799	236	73
					$p = 10$				
.4	53	62	61	55	45	35	25	17	14
.3	96	108	105	93	73	53	34	20	16
.2	211	238	233	203	156	107	62	28	18
.1	845	944	909	782	598	397	209	71	30
.05	3224	3656	3565	3088	2369	1561	799	236	73
					$p = 15$				
.4	62	70	68	61	50	40	29	22	18
.3	106	116	112	98	79	57	39	25	19
.2	223	246	239	209	163	113	65	32	21
.1	857	974	952	828	632	416	216	72	33
.05	3375	3753	3633	3137	2405	1581	808	237	77
					$p = 20$				
.4	80	79	75	66	56	45	37	26	23
.3	117	122	117	103	84	62	42	30	24
.2	228	248	240	209	165	115	69	37	26
.1	857	974	952	832	643	428	226	77	38
.05	3545	3870	3718	3202	2456	1621	839	255	80

size for constructing a 95% confidence interval with expected width not more than 0.2, when $\rho^2 = 0.8$ and $p = 16$, we can interpolate the sample size using the sample sizes given for $p = 15$ and 20 as

$$n = \frac{c_1 \times 65 + c_2 \times 69}{c_1 + c_2} \simeq 67,$$

where the weights $c_1 = 1/15$ and $c_2 = 1/20$. For this sample size, we estimated the expected width of 95% confidence intervals as 0.197, which is very close to 0.2. If p is close to 20, then we should choose $c_1 = 1/20$ and $c_2 = 1/15$.

ILLUSTRATIONS

We now illustrate the methods in two scenarios. The first one concerns calculating sample sizes prior to collecting samples, and in the second one we illustrate the inferential procedures using the empirical data from Timm (1975) that relate scores from the Peabody Picture Vocabulary Test to four proficiency measures.

Suppose that a researcher is interested in computing a 95% confidence interval (CI) for a squared population multiple correlation coefficient ρ^2 so that the width of the CI is at most 0.1. Let us assume that the correlation study involves $p = 5$ variables, and the researcher, based on her or his own experience, believes that the true value of ρ^2 is around 0.8. Then the required sample size (from Table 2 under heading $p = 5$ and $w = 0.1$) is 205.

In addition, the researcher wants to test $H_0 : \rho^2 \leq 0.7$ vs. $H_a : \rho^2 > 0.7$ at level of significance 0.05 and likes to find the sample size for rejection of the null hypothesis with probability of at least 0.80. For this test, the sample size can be obtained from Table 1 ($p = 5$, $\rho_1^2 = 0.80$, $\rho_0^2 = 0.70$) and is 118.

If the null hypothesis is rejected, then this test just provides evidence to conclude that the true value of ρ^2 exceeds 0.7. On the other hand, a lower confidence limit for ρ^2 may be more informative than the result of the one-sided hypothesis test. For instance, the researcher wants to find a 95% lower limit ρ_L^2 for ρ^2 such that the expected value of $(\rho^2 - \rho_L^2)$ is not more than 0.05. As the left endpoint of a 90% CI is a 95% one-sided lower confidence limit for ρ^2, an approximate sample size for this purpose can be obtained from Table 3. Noting that $p = 5$, $\rho^2 = 0.8$. and using $w = 0.1$, we found the required sample size from Table 3 as 147. We should emphasize that the expected width of a 90% CI based on a sample of 147 measurements is at most 0.1, but there is no guarantee that $E(\rho^2 - \rho_L^2)$ is around 0.05 because the CI is not necessarily centered at r^2.

We now illustrate the calculation of confidence intervals, one-sided limits, and p values using the example given in Timm (1975) and Rencher (1998). The original data set contains seven variables, three dependent variables and four explanatory variables (representing test scores) observed for a sample of 37 students. We here consider only five variables, namely, Peabody Picture Vocabulary Test (PPVT) and four learning proficiency tests *named* (N), *still* (S), *named action* (NA) and *sentence still* (SS). The data are given in Timm (1975, Table 4.7.1) and analyzed in Rencher (1998) using *canonical correlation model*

TABLE 3
Sample Sizes Required for 90% Confidence Intervals
With Expected Widths at Most w

	ρ^2				$p = 2$				
w	.2	.3	.4	.5	.6	.7	.8	.9	.95
.4	26	33	35	32	26	20	13	7	5
.3	52	64	66	59	47	33	20	10	7
.2	130	153	152	134	105	71	40	16	9
.1	546	630	619	540	417	277	144	47	18
.05	2209	2538	2490	2163	1662	1095	559	164	50
					$p = 3$				
.4	28	35	36	33	27	20	13	9	6
.3	56	67	67	60	48	34	20	11	7
.2	132	155	154	135	106	72	40	17	10
.1	549	632	621	541	417	277	145	47	19
.05	2212	2545	2491	2163	1665	1096	561	165	50
					$p = 4$				
.4	31	37	38	34	29	21	15	10	7
.3	58	68	68	61	49	35	22	12	8
.2	135	157	155	136	107	73	41	18	11
.1	551	634	622	542	419	277	145	48	20
.05	2212	2545	2491	2163	1665	1096	561	165	51
					$p = 5$				
.4	32	39	39	36	30	22	16	10	8
.3	60	70	70	62	50	36	23	13	9
.2	138	159	156	137	107	74	42	19	11
.1	554	635	624	543	419	278	147	48	20
.05	2216	2547	2493	2164	1665	1096	563	166	52
					$p = 6$				
.4	34	40	41	37	31	24	17	11	9
.3	63	72	71	63	51	37	24	14	10
.2	140	160	157	138	108	75	43	20	12
.1	557	637	625	544	420	278	147	49	21
.05	2218	2548	2496	2167	1665	1098	563	167	52
					$p = 7$				
.4	37	42	42	38	32	24	18	14	10
.3	64	73	72	64	52	38	24	15	11
.2	143	162	159	139	110	76	43	20	13
.1	559	639	626	545	421	280	148	50	22
.05	2220	2548	2496	2167	1667	1098	562	167	53

(*continued*)

TABLE 3
(*Continued*)

w	ρ^2 .2	.3	.4	.5	.6	.7	.8	.9	.95
					$p = 8$				
.4	39	43	43	39	33	26	19	13	11
.3	67	75	74	66	53	38	25	16	12
.2	145	163	160	140	111	76	44	21	14
.1	561	641	628	547	422	280	148	51	22
.05	2222	2551	2497	2167	1667	1099	564	168	54
					$p = 9$				
.4	41	45	45	41	34	27	20	14	13
.3	69	77	75	67	54	40	26	17	13
.2	147	165	161	141	111	77	46	22	15
.1	563	642	628	547	422	281	148	51	23
.05	2227	2551	2501	2167	1668	1099	565	168	54
					$p = 10$				
.4	45	47	46	42	35	27	21	16	14
.3	71	78	76	68	55	41	27	17	14
.2	150	167	163	142	112	78	46	23	16
.1	566	644	630	548	423	282	149	52	24
.05	2228	2552	2501	2171	1669	1101	565	169	55
					$p = 15$				
.4	76	63	56	49	41	34	30	21	19
.3	96	88	83	73	60	45	32	24	19
.2	161	175	169	148	117	82	50	29	22
.1	578	652	636	553	428	286	153	55	29
.05	2242	2563	2506	2175	1672	1105	568	172	58
					$p = 20$				
.4	129	99	81	69	60	53	38	26	24
.3	148	115	97	82	68	57	50	29	24
.2	193	184	175	153	121	86	55	41	26
.1	589	660	642	558	433	290	156	59	42
.05	2253	2572	2513	2182	1677	1109	571	174	62

under normality. Note that the random regression model (with PPVT score as dependent variable and the four proficiency test scores as predictor variables) discussed in the introduction can be used to relate the PPVT scores to the proficiency test scores. This is because, as mentioned in the introduction, the predictor variables cannot be fixed or controlled a priori.

We now compute the squared multiple correlation coefficient R^2 using the following correlation matrix of these varibales taken from Rencher (1998, p. 330).

	PPVT	N	S	NA	SS
PPVT	1	–	–	–	–
N	.44	1	–	–	–
S	.27	.40	1	–	–
NA	.67	.65	.65	1	–
SS	.59	.67	.43	.80	1

Let $\boldsymbol{r}$ denote this correlation matrix and partition $\boldsymbol{r}$ as $\begin{pmatrix} 1 & \boldsymbol{r}'_{21} \\ \boldsymbol{r}_{21} & \boldsymbol{r}_{22} \end{pmatrix}$. The squared multiple correlation coefficient between PPVT and (N, S, NA, SS) is given by

$$r^2 = \boldsymbol{r}'_{21}\boldsymbol{r}_{22}^{-1}\boldsymbol{r}_{21} = 0.499.$$

To find a 95% CI for ρ^2, we need to solve the equations in Equation (9). That is, we need to determine the values of ρ_L^2 and ρ_U^2 so that

$$P(R^2 \geq 0.499|37, 5, \rho_L^2) = 0.025 \quad \text{and} \quad P(R^2 \leq 0.499|37, 5, \rho_U^2) = 0.025. \tag{14}$$

Entering 37 for n, 5 for number of variates, and 0.499 for observed r^2 in the *StatCalc* calculator mentioned earlier, we obtained $\rho_L^2 = 0.174$ and $\rho_U^2 = 0.668$. Thus, (0.174, 0.668) is a 95% confidence interval for ρ^2. Note this CI can also be obtained using the "R2" program due to Steiger and Fouladi (1992).

Suppose one is interested in testing $H_0 : \rho^2 \leq .3$ vs. $H_a : \rho^2 > .3$, where ρ^2 is the true squared multiple correlation coefficient between PPVT and (N, S, NA, SS). Then the p value for this test is given by

$$P(R^2 > 0.499|n = 37, p = 5, \rho_0^2 = .3) = 0.147.$$

Thus, at the level of significance 0.05, we cannot conclude that the true ρ^2 is larger than 0.3. This p value can be obtained using the *StatCalc* or the "R2" program. As the calculated p value is greater than 0.05, we cannot conclude that the true value of ρ^2 is greater than 0.3.

CONCLUDING REMARKS

We described available exact inferential procedures for a squared multiple correlation coefficient ρ^2. We also showed that the one-sided tests are UMP and the one-sided confidence intervals are uniformly most accurate. For a given sample size, number of predictors, and squared sample multiple correlation coefficient, hypothesis tests or interval estimation can be carried out using available software mentioned earlier. Also, we provided calculation of sample sizes for estimating ρ^2 within a given precision and testing ρ^2 with a power of at least 0.80. In particular, we provided sample size calculation for testing (a) $H_0 : \rho^2 \geq \rho_0^2$ vs. $H_a : \rho^2 < \rho_0^2$ and for testing (b) $H_0 : \rho^2 \leq \rho_0^2$ vs. $H_a : \rho^2 > \rho_0^2$. This type of hypotheses involving nonzero null values arises in test of fit in structural equation modeling. Specifically, as pointed out by Algina and Olejnik (2003), the pair of hypotheses in (a) has the same structure as those for test of close fit in structural equation modeling, and the ones in (b) have the same structure as the hypotheses for test of not close fit.

We should emphasize that all the inferential procedures and sample size calculations were developed under the assumption of multivariate normality. The multivariate normality is an idealized situation that is often not achieved in practice. Robustness of these inferential procedures to normality violation is not known. However, we believe that the inferential procedures may perform satisfactorily if the underlying distribution is an elliptical distribution (Muirhead, 1982, Theorem 5.2.8). The multivariate normal distribution is a member of the family of elliptical distributions, and the distribution of R^2 is asymptotically normal whether the sample is from a normal distribution or from an elliptical distribution. On the basis of this result, we expect that the inferential procedures based on the normality assumption may work reasonably well if the sample is from an elliptical distribution.

ACKNOWLEDGMENT

We are grateful to the reviewers and the editor for providing some useful comments and suggestions. The detailed comments by the editor improved the presentation of an earlier version of the article.

REFERENCES

Algina, J., & Olejnik, S. (2003). Sample size tables for correlation analysis with applications in partial correlation and multiple regression analysis. *Multivariate Behavioral Research, 38*, 309–323.

Anderson, T. W. (1984). *An introduction to multivariate statistical analysis*. New York: Wiley.

Benton, D., & Krishnamoorthy, K. (2003). Computing discrete mixtures of continuous distributions: Noncentral chi-square, noncentral t and the distribution of the square of the sample multiple correlation coefficient. *Computational Statistics and Data Analysis, 43*, 249–267.

Casella, G., & Berger, R. L. (2002). *Statistical inference*. Pacific Grove, CA: Duxbury.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlebaum.

Ding, C. G., & Bargmann, R. E. (1991). Evaluation of the distribution of the square of the sample multiple-correlation coefficient. *Applied Statistics, 40*, 195–236.

Ezekiel, M., & Fox, K. A. (1959). *Methods of correlation and regression analysis*. New York: Wiley.
Fisher, R. A. (1928). The general sampling distribution of the multiple correlation coefficient. *Proceedings of the Royal Society of London A, 121*, 654–673.
Gurland, J. (1968). A relatively simple form of the distribution of the multiple correlation coefficient. *Journal of the Royal Statistical Society. Series B, 30*, 276–283.
Helland, I. S. (1987). On the interpretation and use of R^2 in regression analysis. *Biometrics, 43*, 61–69.
Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis.* Upper Saddle River, NJ: Prentice Hall.
Kramer, K. H. (1963). Tables for constructing confidence limits on the multiple correlation coefficient. *Journal of the American Statistical Association, 58*, 1082–1085.
Krishnamoorthy, K. (2006). *Handbook of statistical distributions with applications*. New York: Chapman & Hall/CRC.
Lee, Y. S. (1972). Tables of upper percentage points of the multiple correlation coefficient. *Biometrika, 59*, 175–189.
Maxwell, S. E. (2004). The persistence of underpowered studies in psychological research: Causes, consequences, and remedies. *Psychological Methods, 9*, 147–163.
Mendoza, J. L., & Stafford, K. L. (2001). Confidence intervals, power calculation, and sample size estimation for the squared multiple correlation coefficient under the fixed and random regression models: A computer program and standard tables. *Educational and Psychological Measurement, 61*, 650–667.
Muirhead, R. J. (1982). *aspects of multivariate statistical theory*. New York: Wiley.
Rencher, A. C. (1998). *Multivariate statistical inference and applications*. New York: Wiley.
Smithson, M. (2001). Correct confidence intervals for various regression effect sizes and parameters: The importance of noncentral distribution in computing intervals: Confidence interval for effect sizes. *Education and Psychology Measurement, 61*, 605–632.
Srivastava, M. S. (2002). *Methods of multivariate statistics*. New York: Wiley.
Steiger, J. H., & Fouladi, R. T. (1992). R2: A computer program for interval estimation, power calculation, and hypothesis testing for the squared multiple correlation. *Behavior Research Methods, Instruments, and Computers, 4*, 581–582.
Timm, N. H. (1975). *Multivariate analysis with applications in education and psychology*. Belmont, CA: Wadsworth.

APPENDIX A

The probability density function of R^2 (see Muirhead, 1982, page 172) corresponding to the CDF in Equation (2) is given by

$$f(R^2|n, p, \rho^2) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-p+1}{2}\right)\Gamma\left(\frac{p-1}{2}\right)} (R^2)^{\frac{p-3}{2}} (1-R^2)^{\frac{n-p-1}{2}} (1-\rho^2)^{\frac{n}{2}} \cdot F\left(\frac{n}{2},\ \frac{n}{2};\ \frac{p-1}{2};\ \rho^2 R^2\right).$$

To show that the density function $f(R^2|n, p, \rho^2)$ has a monotone likelihood ratio, it suffices to prove

$$\frac{\partial^2 \log f(R^2|n, p, \rho^2)}{\partial\rho^2 \partial R^2} \geq 0$$

for all ρ^2 and R^2. It can be readily verified that

$$\frac{\partial^2 \log f(R^2|n, p, \rho^2)}{\partial\rho^2 \partial R^2} = \frac{\partial^2 \log F\left(\frac{n}{2},\ \frac{n}{2};\ \frac{p-1}{2};\ \rho^2 R^2\right)}{\partial\rho^2 \partial R^2}.$$

Write

$$g(\xi) = F\left(\frac{n}{2},\ \frac{n}{2};\ \frac{p-1}{2};\ \xi\right) = \sum_{i=0}^{\infty} \delta_i \xi^i,$$

where δ_i $(i = 0, 1, \ldots)$ are positive constants (see Anderson, 1984, page 113). Then we have

$$\frac{\partial^2 \log f(R^2|n, p, \rho^2)}{\partial\rho^2 \partial R^2} = \frac{\partial^2 \log g(\rho^2 R^2)}{\partial\rho^2 \partial R^2} = H(\rho^2 R^2),$$

where

$$H(\xi) = \frac{(g'(\xi) + \xi g''(\xi))g(\xi) - \xi(g'(\xi))^2}{g^2(\xi)}.$$

It is easy to see that

$$(g'(\xi) + \xi g''(\xi))g(\xi) - \xi(g'(\xi))^2 = \frac{1}{2}\sum_{i,j=0}^{\infty} (i-j)^2 \delta_i \delta_j \xi^{i+j-1}.$$

Hence,

$$\frac{\partial^2 \log f(R^2|n, p, \rho^2)}{\partial\rho^2 \partial R^2} = H(\rho^2 R^2) \geq 0.$$

Because the PDF R^2 has monotone likelihood ratio, the distribution function of R^2 is monotone in ρ^2. More specifically, for a given x, n, and p, $P(R^2 \leq x|n, p, \rho^2)$ is a decreasing function of ρ^2 (see Casella & Berger, 2002, p. 406).

APPENDIX B

Pseudo random numbers on R^2 can be generated using the following algorithm, which is based on the distributional result (see Muirhead, 1982, Theorem 5.2.4) that

$$\frac{R^2}{1-R^2} \sim \frac{\chi^2_{p-1}\left(\frac{\rho^2}{1-\rho^2}\chi^2_{n-1}\right)}{\chi^2_{n-p}},$$

where all the chi-square random variables are independent.

Algorithm B1
For a given $n,\ p$, and ρ^2,
Set $\eta = \frac{\rho^2}{1-\rho^2}$
Generate a χ^2_{n-1} variate
Generate a normal variate X with mean $\sqrt{\eta\chi^2_{n-1}}$ and variance 1
Generate a χ^2_{p-2} variate
Set $V = \chi^2_{p-2} + X^2$ so that $V \sim \chi^2_{p-1}\left(\frac{\rho^2}{1-\rho^2}\chi^2_{n-1}\right)$
Generate a χ^2_{n-p} variate
Set $R^2 = \frac{V}{V+\chi^2_{n-p}}$
Note that if $p = 2$, then V should be set to X^2, which follows $\chi^2_1\left(\frac{\rho^2}{1-\rho^2}\chi^2_{n-1}\right)$. Notice further that to generate a variate R^2, this algorithm generates three independent chi-square variates and one normal variate regardless of the values of the dimension p and the sample size n.