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 USEFUL STANDARD TABLES

JORGE L. MENDOZA AND KAREN L. STAFFORD
University of Oklahoma

In this article, the authors introduce a computer package written for Mathematica, the purpose of which is to perform a number of difficult iterative functions with respect to the squared multiple correlation coefficient under the fixed and random models. These functions include, among others, computation of confidence interval upper and lower bounds, power calculation, calculation of sample size required for a specified power level, and providing estimates of shrinkage in cross validating the squared multiple correlation under both the random and fixed models. Attention is given to some of the technical issues regarding the selection of, and working with, these two types of models as well as to issues concerning the construction of confidence intervals.

Much emphasis has been placed lately on the importance of using confidence intervals in data analysis in addition to or as opposed to performing null hypothesis testing. Some authors have gone so far as to suggest the replacement or banishment of hypothesis testing (e.g., Schmidt, 1996). In the replacement of hypothesis testing, it is argued that parameter estimates be accompanied by their margin of error—confidence intervals. Clearly, this recommendation is especially valid when sample size is either very large or very small. Large sample sizes can produce statistical tests that are overly sensitive and, thus, lead to the finding of statistically significant differences where only minuscule differences between parameters actually exist. On the other hand, small samples produce tests that are not very sensitive and lead to the detection of only large differences between the parameters. There are other reasons why these authors have argued against tests of hypotheses, but we

Educational and Psychological Measurement, Vol. 61 No. 4, August 2001 650-667
© 2001 Sage Publications

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/25578
Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

will not discuss them here. We are not taking a position regarding the merits or demerits of hypothesis testing. Others have already done this (e.g., Baril & Cannon, 1995; Cohen, 1994; Frick, 1996; Hagen, 1997). Instead, we will focus on some of the technical issues inherent in constructing confidence intervals. Confidence intervals on means are well understood and are generally simple to construct. Consequently, we will not discuss them in detail. Confidence intervals on correlations are more difficult to construct, especially when they involve the multiple correlation coefficient. We will focus on covering the procedure involved in obtaining a confidence interval on the squared multiple correlation.

Multiple correlation procedures are widely used in education and the social sciences. The underlying statistical models for these procedures are of two types. These models are referred to as fixed and random models or, respectively, correlation and regression models. A mathematical exposition of the two models is given in Sampson (1974). Mathematically, we can often view the fixed model as a conditional random regression model. The random model is more appropriate for nonexperimental situations in which the levels of the independent variables are not fixed a priori. It is common in education and other social sciences to have studies in which the levels of the multiple independent variables for each experimental unit cannot be controlled and are available only after the observations are made. This type of design clearly falls under the random model. In contrast, under the fixed model, the levels of the independent variables are fixed before data collection.

Ordinarily, when computing a multiple regression analysis in a statistical package like SAS or SPSS, the researcher need not worry whether data fit the assumptions of the random versus the fixed model. Tests of hypotheses and estimates of parameters are the same under both models (see Figure 1). One needs to be aware, however, of the model under which one is working when trying to establish the necessary sample size for a desired power level or when finding a confidence interval for the squared multiple correlation. In addition, although for large samples the results of power calculations under the two models are similar, researchers need to be aware of the model under which they are working when computing power. Lee (1972) gave tables for power calculations under the random model for the multiple correlation coefficient. These tables are, however, difficult to use. An easier set of tables is found in Gatsonis and Sampson (1989), but these tables only provide the sample size necessary to obtain a specified level of power. Sample size estimation and power calculations under the fixed model are given in Cohen (1988).

Power calculations, sample size determination, and confidence interval estimation are more complex under the random model than under the fixed model. Under the fixed model, the noncentral F distribution is used for calculations. This is a well-known distribution for which many tables and computer programs are available. Unfortunately, not many computer resources exist to aid the researcher working with the random model.

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

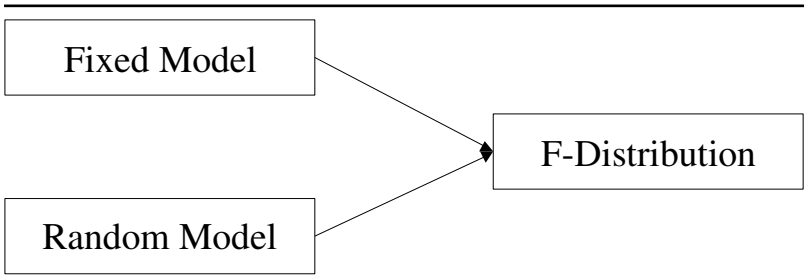


Figure 1. Testing the null hypothesis that $\rho^2 = 0$ under the fixed and random models.

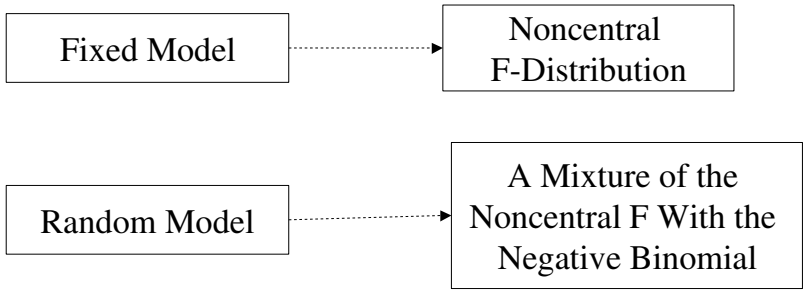


Figure 2. The appropriate distribution for constructing a confidence interval for ρ^2 under the fixed and random models.

Obtaining a confidence interval on the population squared multiple correlation ρ^2 requires that we identify a model. A confidence interval obtained under the fixed model is not the same as one obtained under the random model. Confidence intervals obtained under the fixed model are usually shorter than those obtained under the random model. The implication for us is that we must determine which model is appropriate before we can embark on computing the confidence interval for ρ^2. Figure 2 illustrates the appropriate distribution that must be used for constructing confidence intervals on ρ^2 under the random or fixed model.

The construction of a confidence interval on ρ^2 under either the random model or the fixed model is somewhat complex and requires the use of a computer or special tables. We will outline the procedures used to construct confidence intervals under both models. The construction of these intervals is similar in principle but different in execution. Constructing a confidence interval under the fixed model requires access to the noncentral *F* distribution. The experienced researcher can use this distribution, which is available in SAS, to obtain the interval. We have written and compiled a package in

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

Table 1
Functions Contained in MultipleR2 Mathematica Package

Function	Usage	Description
Random model functions		
probr	probr[n,p,rho2,x]	Computes the probability that R^2 is < x
confidence	confidence[n,p,R2,plevel]	Computes the lower or upper bound for a p-level confidence interval for ρ^2 given R^2
powerR2	powerR2[n,p,rho2,alpha]	Computes the probability of rejecting the null hypothesis that $\rho^2 = 0$ when ρ^2 = rho2
findsample	findsample[p,rho2,power, alpha]	For power and alpha, computes the sample size necessary to reject the null hypothesis that $\rho^2 = 0$ when ρ^2 = rho2
crossR2sample	crossR2sample[p,rho2, pcross,diff]	Gives the necessary sample to obtain a cross-validated R^2 that does not differ from rho2 by more than diff
crossR2 confidence	crossR2confidence[n,p, rho2,pcross]	For a given rho2, returns the one-sided confidence interval for the cross-validated R^2 at the p-cross level
Fixed model functions		
confidencefix	confidencefix[n,p,R2, plevel]	Computes the lower or upper bound for a p-level confidence interval for ρ^2 given R^2
probf	probf[n,p,fval]	Computes the probability that the F is < or = fval
powerR2fix	powerR2fix[n,p,rho2, alpha]	Computes the probability of rejecting the null hypothesis that $\rho^2 = 0$ when ρ^2 = rho2
invprobf	invprobf[n,p,upperp]	Inverse function for probf, returns the fval
stdnormal	stdnormal[z]	Computes the probability that a standard normal variable is < z
findsamplefix	findsamplefix[p,rho2, power,alpha]	For power and alpha, computes the sample size necessary to reject the null hypothesis that $\rho^2 = 0$ when ρ^2 = rho2
invnormal	invnormal[plevel]	Inverse of std normal

Mathematica 3.0, in which we have included a set of functions that give the confidence interval easily and directly under either the fixed or the random model. We will illustrate the use of these functions later in this article. Additional useful functions, including a function to compute power under both models, are included in the package. See the appendix for information about accessing the Mathematica package, which is available from the authors. Table 1 provides function, syntax, and output information for all of the available functions. Although it could be argued that better computational engines exist, we have written the package in Mathematica 3.0 because of the ease with which these functions can be incorporated into other calculations and plotting functions that exist in Mathematica. For users without access to Mathematica, we have created a set of confidence interval tables. (See Tables

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

Table 2
Lower Bounds for ^{2}, N = *20*

N_{20}	Number of Predictors									
R^2	2	3	4	5	6	8	10	12	14	16
.30	.002	0	0	0	0	0	0	0	0	0
.32	.012	0	0	0	0	0	0	0	0	0
.34	.023	0	0	0	0	0	0	0	0	0
.36	.036	0	0	0	0	0	0	0	0	0
.38	.048	.001	0	0	0	0	0	0	0	0
.40	.062	.015	0	0	0	0	0	0	0	0
.42	.078	.029	0	0	0	0	0	0	0	0
.44	.096	.046	0	0	0	0	0	0	0	0
.46	.111	.064	.010	0	0	0	0	0	0	0
.48	.131	.082	.026	0	0	0	0	0	0	0
.50	.148	.101	.046	0	0	0	0	0	0	0
.52	.170	.121	.069	.008	0	0	0	0	0	0
.54	.189	.143	.088	.029	0	0	0	0	0	0
.56	.214	.166	.113	.052	0	0	0	0	0	0
.58	.237	.190	.135	.077	.009	0	0	0	0	0
.60	.262	.215	.164	.103	.037	0	0	0	0	0
.62	.288	.242	.188	.130	.062	0	0	0	0	0
.64	.315	.27	.22	.16	.095	0	0	0	0	0

2-8.) Next, we briefly discuss the general principles involved in finding a confidence interval.

Confidence Intervals

The standard situation in setting a confidence interval on the mean of a normal distribution is a straightforward procedure. Also, in the standard situation, the processes involved in obtaining a confidence interval and testing a null hypothesis are very similar. The confidence interval is obtained by "inverting" the hypothesis test. Consider the null hypothesis Ho: $\mu = \mu_o$. The test of this null hypothesis is performed by checking to see whether the interval $\mu_o \pm t\, S/\sqrt{N}$ covers the observed sample mean $\bar{x}$. If the interval does not include the mean, we reject the null. Using this same interval with the sample mean instead of μ_o, we obtain the confidence interval on μ, $\bar{x} \pm t\, S/\sqrt{N}$. (The t in both situations refers to the appropriate cutoff point under the central t distribution, and S refers to the sample standard deviation.) Before leaving the subject of a confidence interval on the mean, it should be pointed out that because a regression weight β is a conditional mean, its estimation (point or interval) is the same under the fixed or random model (see Sampson, 1974).

The inversion process is more complex in situations in which we are unable to find a function of the estimator that is independent of the parame-

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

Table 3
Lower Bounds for ^{2}, N = *30*

N_{30}	Number of Predictors									
R^2	2	3	4	5	6	8	10	12	14	16
.20	.001	0	0	0	0	0	0	0	0	0
.22	.009	0	0	0	0	0	0	0	0	0
.24	.019	0	0	0	0	0	0	0	0	0
.26	.030	.002	0	0	0	0	0	0	0	0
.28	.041	.014	0	0	0	0	0	0	0	0
.30	.053	.025	0	0	0	0	0	0	0	0
.32	.067	.04	.01	0	0	0	0	0	0	0
.34	.082	.053	.023	0	0	0	0	0	0	0
.36	.098	.070	.039	.005	0	0	0	0	0	0
.38	.115	.086	.053	.020	0	0	0	0	0	0
.40	.131	.103	.071	.037	.003	0	0	0	0	0
.42	.150	.121	.088	.055	.019	0	0	0	0	0
.44	.168	.139	.108	.075	.039	0	0	0	0	0
.46	.186	.158	.129	.095	.059	0	0	0	0	0
.48	.205	.18	.15	.116	.080	.001	0	0	0	0
.50	.230	.201	.171	.136	.101	.023	0	0	0	0
.52	.251	.223	.192	.162	.125	.046	0	0	0	0
.54	.274	.244	.217	.185	.151	.071	0	0	0	0
.56	.297	.271	.240	.21	.175	.098	.004	0	0	0
.58	.321	.294	.267	.235	.203	.126	.036	0	0	0
.60	.346	.318	.292	.262	.229	.154	.065	0	0	0
.62	.370	.346	.319	.290	.256	.184	.096	0	0	0
.64	.397	.372	.347	.32	.287	.217	.13	.025	0	0

ters of the sampling distribution. This is the case for the squared multiple correlation ρ^2. The confidence interval on ρ^2 is not a simple inverse of the procedure used to test the null hypothesis. The problems lie in that the variance of the sampling distribution of R^2 depends on ρ^2. The process of finding a confidence interval under this dependency requires working directly with the sampling distribution of the squared multiple correlation R^2. To explain briefly, we must identify the values $h_1$and h_2 that demarcate an area under the sampling distribution equal to the chosen $(1 - \alpha)$% level of confidence for each possible value of ρ^2 (see Mood & Grabill, 1963). The values are selected such that the sum of the conditional probabilities, $P(R^2 < h_1 \mid \rho^2)$ and $P(R^2 > h_2 \mid \rho^2)$, equals α. Without loss of generality, we focus on a specific n (sample size) and p (number of independent variables) to illustrate the procedure.

Let us consider a 90% confidence interval on ρ^2 when $p = 4$ and $n = 40$. To find the interval, we first find the values for h_1 and h_2 that make the sum of the conditional probabilities .10. The h_1 and h_2 are found for many consecutive values of ρ^2 in the interval (0,1). (Note that h_1 and h_2 yield the lower and upper bounds of R^2, respectively.) Plotting h_1 and h_2 for each ρ^2, we can see that h_1

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

Table 4
Lower Bounds for ^{2}, N = *40*

N_{40}	Number of Predictors									
R^2	2	3	4	5	6	8	10	12	14	16
.16	.004	0	0	0	0	0	0	0	0	0
.18	.013	0	0	0	0	0	0	0	0	0
.20	.023	.003	0	0	0	0	0	0	0	0
.22	.034	.013	0	0	0	0	0	0	0	0
.24	.046	.026	.004	0	0	0	0	0	0	0
.26	.058	.038	.017	0	0	0	0	0	0	0
.28	.072	.052	.030	.007	0	0	0	0	0	0
.30	.086	.066	.044	.021	0	0	0	0	0	0
.32	.102	.082	.06	.037	.012	0	0	0	0	0
.34	.119	.098	.075	.053	.029	0	0	0	0	0
.36	.135	.115	.092	.070	.045	0	0	0	0	0
.38	.154	.133	.109	.087	.062	.011	0	0	0	0
.40	.171	.15	.128	.106	.081	.029	0	0	0	0
.42	.190	.170	.147	.124	.101	.049	0	0	0	0
.44	.209	.189	.168	.146	.122	.070	.012	0	0	0
.46	.23	.210	.190	.167	.143	.091	.034	0	0	0
.48	.251	.232	.21	.187	.165	.114	.056	0	0	0
.50	.273	.253	.234	.210	.187	.136	.082	.015	0	0
.52	.296	.276	.255	.235	.211	.162	.105	.042	0	0
.54	.318	.299	.278	.257	.236	.187	.132	.069	0	0
.56	.341	.323	.304	.284	.262	.214	.159	.098	.026	0
.58	.364	.346	.328	.308	.287	.240	.188	.126	.058	0
.60	.389	.372	.353	.335	.314	.269	.217	.159	.089	.009
.62	.414	.397	.380	.360	.341	.297	.247	.188	.123	.043
.64	.44	.425	.407	.39	.37	.327	.28	.222	.157	.08

and h_2 define two curves on the ρ^2-R^2 plane. Given a specific value of R^2 (say R^{*2}), the confidence interval on ρ^2 is obtained by drawing a vertical line through R^{*2} parallel to the ρ^2 axis. This line intersects the two curves at points h^*_1 and h^*_2. When the points are projected onto the ρ^2 axis, we obtain the upper and lower bounds for the confidence interval. In Figure 3, we see that if the sample multiple correlation squared is .45, we can say with 90% confidence that ρ^2 lies in the interval between .18 and .57. We discuss this procedure at length in the next section.

One-Sided Confidence Interval on ^{2}

As we previously mentioned, the computation of a confidence interval on ρ^2 is complex because we cannot find a function of the sample multiple correlation R^2 (and ρ^2) that is independent of the parameters of the sampling distribution. When we cannot find such a function, we must work directly with the

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

Table 5
Lower Bounds for ^{2}, N = *50*

N_{50}	Number of Predictors									
R^2	2	3	4	5	6	8	10	12	14	16
.14	.008	0	0	0	0	0	0	0	0	0
.16	.017	.002	0	0	0	0	0	0	0	0
.18	.028	.012	0	0	0	0	0	0	0	0
.20	.039	.023	.007	0	0	0	0	0	0	0
.22	.051	.036	.018	.002	0	0	0	0	0	0
.24	.065	.048	.031	.015	0	0	0	0	0	0
.26	.079	.062	.046	.028	.010	0	0	0	0	0
.28	.094	.078	.061	.043	.025	0	0	0	0	0
.30	.110	.093	.077	.058	.039	.002	0	0	0	0
.32	.127	.11	.092	.075	.057	.017	0	0	0	0
.34	.143	.127	.110	.092	.074	.034	0	0	0	0
.36	.161	.144	.127	.109	.092	.053	.011	0	0	0
.38	.179	.163	.146	.129	.111	.071	.029	0	0	0
.40	.198	.182	.165	.148	.131	.092	.05	.003	0	0
.42	.218	.203	.185	.168	.150	.111	.070	.024	0	0
.44	.238	.223	.206	.189	.171	.134	.092	.048	0	0
.46	.258	.244	.228	.210	.194	.156	.115	.070	.021	0
.48	.281	.264	.249	.235	.215	.18	.138	.093	.045	0
.50	.300	.287	.271	.255	.238	.203	.162	.119	.070	.015
.52	.325	.308	.294	.278	.262	.227	.186	.144	.097	.044
.54	.345	.333	.318	.301	.286	.251	.213	.170	.124	.071
.56	.369	.356	.341	.325	.310	.277	.240	.199	.153	.102
.58	.394	.380	.367	.351	.335	.303	.267	.226	.183	.131
.60	.417	.405	.391	.377	.363	.330	.295	.257	.213	.164
.62	.443	.429	.416	.403	.389	.358	.324	.285	.244	.196
.64	.467	.455	.442	.43	.415	.387	.352	.317	.277	.23

sampling distribution. Specifically, we work with the integral of the sampling distribution of R^2 for a given value of ρ^2. In a one-sided $(1 - \alpha)\%$ confidence interval, we must solve for h such that

$$P(R^2 \geq h) = \int_h^1 g(R^2; \rho^2) dR^2 = \alpha. \quad (1)$$

The function $g(R^2; \rho^2)$ represents the sampling distribution of R^2 under the fixed or random model. If h were only a function of ρ^2, we could construct a figure similar to Figure 3 and find the one-sided $(1 - \alpha)\%$ confidence interval. Unfortunately, the sampling distribution of R^2 also depends on n and p. So the approach taken in Figure 3 is not practical because h is a function of ρ^2, n, p, and α. We are not working with a single curve but with a multidimensional surface.

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

Table 6
Lower Bounds for ^{2}, N = *75*

N_{75}	Number of Predictors									
R^2	2	3	4	5	6	8	10	12	14	16
.10	.007	0	0	0	0	0	0	0	0	0
.12	.017	.007	0	0	0	0	0	0	0	0
.14	.028	.018	.007	0	0	0	0	0	0	0
.16	.04	.03	.018	.008	0	0	0	0	0	0
.18	.053	.042	.031	.020	.009	0	0	0	0	0
.20	.067	.056	.045	.034	.022	0	0	0	0	0
.22	.080	.070	.059	.048	.036	.012	0	0	0	0
.24	.096	.086	.075	.063	.051	.028	.002	0	0	0
.26	.112	.101	.090	.079	.067	.043	.018	0	0	0
.28	.129	.118	.107	.096	.084	.060	.035	.008	0	0
.30	.146	.135	.124	.112	.100	.077	.051	.025	0	0
.32	.163	.152	.142	.131	.12	.095	.07	.043	.016	0
.34	.181	.171	.160	.148	.138	.114	.088	.062	.034	.009
.36	.201	.189	.18	.168	.157	.133	.109	.082	.054	.026
.38	.219	.209	.198	.188	.176	.154	.129	.103	.075	.047
.40	.239	.229	.218	.207	.196	.175	.15	.125	.096	.068
.42	.259	.249	.239	.229	.218	.195	.172	.146	.119	.091
.44	.280	.269	.259	.249	.238	.216	.194	.168	.142	.113
.46	.300	.291	.282	.271	.260	.238	.215	.192	.165	.138
.48	.322	.313	.303	.292	.283	.265	.24	.215	.189	.163
.50	.343	.333	.326	.316	.306	.285	.263	.240	.214	.187
.52	.365	.357	.347	.339	.329	.308	.286	.264	.239	.215
.54	.388	.379	.371	.361	.352	.333	.312	.288	.265	.240
.56	.411	.402	.393	.385	.376	.356	.336	.315	.293	.269
.58	.435	.425	.418	.410	.401	.382	.362	.342	.319	.296
.60	.458	.45	.442	.433	.425	.407	.389	.367	.346	.323
.62	.481	.474	.467	.458	.450	.433	.415	.395	.375	.353
.64	.506	.498	.492	.483	.476	.46	.442	.423	.403	.382

One practical solution to the problem is to write a computer program that can integrate the sampling distribution of R^2. The sampling distribution of R^2 is a function of the noncentral F distribution under the fixed regression model. In contrast, the sampling distribution of R^2 under the random model is

$$P(R^2 < h) = \left\{ B\left(1, \frac{p}{2}, \frac{n-p-1}{2}\right) \right\}^{-1} (1-\rho^2)^{\frac{n-1}{2}} \sum_{k=0}^{\infty} C_k. \tag{2}$$

C_k is also a function of the incomplete beta function B,

$$C_k = \frac{\{((n-1)/2)_k\}^2 \rho^{2k} B\left(h, \frac{p}{2}+k, \frac{n-p-1}{2}\right)}{(p/2)_k \, k!},$$

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

Table 7
Lower Bounds for ^{2}, N = *100*

N_{100}	Number of Predictors									
R^2	2	3	4	5	6	8	10	12	14	16
.08	.008	0	0	0	0	0	0	0	0	0
.10	.017	.010	.002	0	0	0	0	0	0	0
.12	.029	.021	.013	.005	0	0	0	0	0	0
.14	.041	.033	.025	.017	.009	0	0	0	0	0
.16	.055	.046	.038	.030	.021	.005	0	0	0	0
.18	.068	.060	.052	.044	.035	.018	0	0	0	0
.20	.083	.075	.067	.059	.050	.032	.014	0	0	0
.22	.099	.091	.082	.074	.066	.048	.030	.012	0	0
.24	.115	.107	.099	.090	.082	.064	.046	.028	.012	0
.26	.132	.123	.115	.107	.099	.081	.063	.044	.027	.020
.28	.149	.142	.133	.124	.117	.099	.080	.062	.043	.031
.30	.167	.159	.151	.142	.134	.117	.099	.080	.062	.045
.32	.186	.177	.170	.161	.153	.136	.118	.100	.081	.062
.34	.204	.196	.188	.180	.172	.155	.138	.119	.100	.081
.36	.223	.216	.208	.199	.192	.175	.157	.140	.120	.101
.38	.243	.236	.228	.219	.212	.195	.178	.160	.142	.123
.40	.262	.256	.248	.240	.232	.217	.200	.182	.164	.145
.42	.283	.275	.269	.260	.253	.237	.221	.203	.185	.167
.44	.304	.297	.289	.281	.275	.259	.242	.226	.207	.190
.46	.325	.318	.310	.303	.296	.281	.265	.249	.231	.213
.48	.345	.339	.331	.325	.318	.303	.288	.271	.255	.238
.50	.367	.361	.354	.347	.340	.326	.310	.294	.279	.261
.52	.390	.383	.376	.369	.363	.349	.335	.318	.303	.286
.54	.411	.405	.398	.392	.386	.373	.358	.343	.329	.312
.56	.434	.428	.422	.415	.409	.397	.382	.368	.354	.337
.58	.457	.450	.446	.439	.433	.421	.407	.394	.379	.364
.60	.480	.474	.468	.464	.457	.445	.432	.419	.405	.391
.62	.503	.498	.492	.488	.481	.469	.457	.445	.432	.417
.64	.527	.522	.517	.512	.506	.495	.483	.471	.458	.445

and $(l)_0 = 1, (l)_k = (l)(l+1)\ldots(l+k-1)$. (See Gatsonis & Sampson, 1989, for more details.) We present here documentation for a program in Mathematica 3.0 to numerically integrate this sampling distribution. This program is part of a larger package that we have developed to compute confidence intervals and power.

The Mathematica 3.0 program is based on the following algorithm. For a manifest value of R^2, say R^{*2}, the program sets $h = R^{*2}$ and $\rho^2 = R^{*2}$ and computes the $P(R^2 \geq R^{*2})$. If this probability is not equal to α, the value of ρ^2 is changed, and the probability is computed again. The program searches through different values of ρ^2 until it finds a ρ^2 such that

$$P(R^2 \geq R^{*2} \mid \rho^2) = \alpha.$$

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

Table 8
Lower Bounds for ^{2}, N = *200*

N_{200}	Number of Predictors									
R^2	2	3	4	5	6	8	10	12	14	16
.06	.013	.009	.006	.002	0	0	0	0	0	0
.08	.025	.021	.017	.013	.009	.001	0	0	0	0
.10	.038	.034	.030	.026	.022	.014	.006	0	0	0
.12	.052	.048	.045	.040	.036	.028	.019	.015	.002	0
.14	.067	.063	.060	.055	.051	.042	.034	.027	.017	.007
.16	.083	.080	.075	.071	.065	.058	.050	.041	.039	.023
.18	.099	.096	.092	.087	.083	.075	.066	.057	.050	.040
.20	.117	.113	.108	.104	.100	.092	.083	.075	.066	.064
.22	.134	.130	.126	.122	.117	.110	.101	.091	.083	.076
.24	.151	.148	.144	.139	.135	.127	.119	.110	.101	.092
.26	.170	.166	.162	.158	.154	.146	.137	.128	.119	.111
.28	.188	.184	.180	.177	.172	.165	.156	.147	.139	.130
.30	.207	.203	.199	.195	.192	.183	.175	.167	.159	.150
.32	.226	.222	.218	.215	.211	.203	.195	.187	.178	.170
.34	.245	.241	.238	.235	.231	.223	.215	.207	.199	.191
.36	.265	.261	.258	.254	.251	.243	.235	.227	.219	.211
.38	.285	.282	.278	.274	.271	.264	.256	.248	.240	.233
.40	.305	.302	.298	.295	.291	.284	.276	.269	.261	.253
.42	.326	.323	.319	.315	.312	.305	.298	.290	.283	.275
.44	.347	.343	.340	.336	.333	.326	.319	.312	.305	.297
.46	.367	.364	.361	.357	.354	.347	.341	.334	.327	.319
.48	.389	.385	.382	.379	.375	.369	.362	.356	.349	.342
.50	.410	.407	.404	.400	.397	.391	.384	.378	.372	.365
.52	.431	.428	.425	.422	.419	.413	.407	.401	.395	.387
.54	.453	.450	.447	.445	.441	.435	.430	.423	.417	.411
.56	.474	.472	.469	.467	.463	.458	.452	.447	.440	.434
.58	.497	.495	.491	.489	.487	.481	.475	.470	.464	.458
.60	.519	.516	.514	.512	.509	.503	.499	.493	.487	.482
.62	.541	.539	.537	.534	.532	.527	.521	.517	.511	.506
.64	.565	.562	.560	.557	.555	.550	.546	.541	.535	.530

The ρ^2 identified in this manner is the lower bound of the one-sided $(1 - \alpha)\%$ confidence interval. The search is conducted with an interval-halving algorithm similar to the one given in MacCallum, Browne, and Sugawara (1996). (All other search procedures in our Mathematica Package are based also on this algorithm.)

The Mathematica Package

The Mathematica Package contains a set of functions for computing power, constructing a confidence interval, and estimating the level of prediction under either the random or fixed model. Table 1 lists these functions and

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

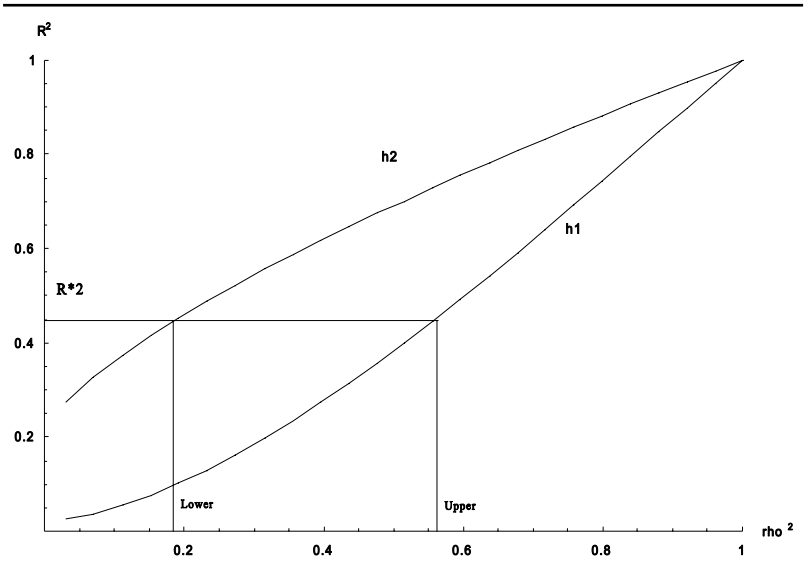


Figure 3. Lower bounds (h_2) and upper bounds (h_1) for the squared multiple correlation coefficient for sample size = 40 and number of independent variables = 4.

gives a short definition of each function. We did not use Mathematica's noncentral *F* distribution. Instead, we created our own because we found the Mathematica function a bit unstable when we tried to integrate it under some extreme conditions. Our noncentral *F* probability routine resembles the one given in SAS 6.11.

One of the functions in the package computes the confidence interval under the random model, whereas another function computes the confidence interval under the fixed model. The package also includes functions for power and sample size calculations under both models. The functions in this Mathematica package were checked against existing tables for accuracy. The functions always were extremely close to the table values. In addition, we took precautions with the functions to stay within the accuracy limitations of a 32-bit processor. We feel confident in the accuracy of our functions. We have used the random confidence function in the package to create Tables 2 through 8. Researchers without access to Mathematica can use these tables to find a confidence interval for ρ^2 under the random model.

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

A Confidence Interval on the Cross-Validated Multiple Correlation

Sometimes in accessing an obtained sample regression b, it is desirable to estimate how well it will predict in future samples. This is a situation likely to be encountered in many situations. The researcher generates a regression equation in a validation study and would like to know how well the obtained regression equation is likely to predict in future samples. Not exactly the same question, but a related one, is how well the obtained regression equation would predict if it were to be applied to the population. We call this measure the predictive precision of the equation, and it is given by

$$\rho_c^2(b) = \frac{(b'\Sigma_{yx})^2}{b'\Sigma_{xx}b\sigma_y^2}. \tag{3}$$

The vector Σ_{xy} contains the population covariances between the independent and dependent variables. Similarly, Σ_{xx} is the variance-covariance matrix of the independent variables. The predictive precision given in Equation 3 is a random variable prior to data collection but a fixed parameter once we obtain the regression equation (see Mendoza, 1977). The ρ_c^2 ranges from 0 to ρ^2.

Under the assumption of multivariate normality, Park and Dudycha (1974) have shown that the sampling distribution of ρ_c^2 follows the noncentral F distribution,

$$\rho_c^2 = \frac{\rho^2}{1+\frac{(p-1)}{F_{1,p-1,\lambda}}},$$

with noncentrality parameter

$$\lambda = \sqrt{\frac{(n-p-2)\rho^2}{(1-\rho^2)}}.$$

For a given ρ^2, we can find the probability that

$$P\left(\rho_c^2 > \rho_{c_L}^2 \middle| \rho^2\right) = 1-a. \tag{4}$$

By working backward, we can identify a lower bound for ρ_c^2 that will satisfy the probability statement in Equation 4. We have created a function "crossR2confidence" that for a given ρ^2, n, p, and a specified confidence level, returns a $(1-\alpha)$ lower bound for the level of predictive precision ρ_c^2. For the usual situation in which we do not know ρ^2, we could use the sample multiple correlation R^2, or, if we wish to be conservative, we could use the lower bound of ρ^2, which can be obtained with our random confidence interval function. By using the sample multiple R^2 in the function "crossR2confidence," we ob-

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

tain an estimate of the "shrinkage"—the difference between the returned value and the sample R^2.

A Practical Example

Allen, Cipielewski, and Stanovich (1992) investigated the relationship between the amount of time that 63 elementary school children spent reading and a number of indicator variables. Two independent variables contributed significantly to the prediction of reading time and were retained in their model with $R = .56$, $R^2 = .31$. Assuming the random model, a one-sided confidence interval is easily established with the Mathematica Package. Using "confidence[63,2,.31,.95]" in the Mathematica package, we find that the lower bound for a 95% confidence interval on ρ^2 is .14; that is, we can say with 95% confidence that ρ^2 is between .14 and 1. We could also get an idea of the magnitude of the lower bound by looking at Table 5. Finally, we wish to obtain an estimate of the shrinkage we would expect to see in a cross-validation study. Using the estimate of ρ^2 obtained from the sample, submitted to the function in the form crossR2confidence[63,2,.31,.95], we find that the cross-validated lower bound returned by the function is .265. Therefore, the estimate of shrinkage is .045(.31 – .265).

What of a similar study conducted with, again, two independent variables, the same observed R^2 of .31, but now with a sample size of only 30? The decrease in estimation precision is seen by the lengthening of the confidence interval, the lower bound of which is now returned as .06. As we would expect, although the level of power is still good at .84, the value returned is considerably less. The increased imprecision is also reflected in the estimate of the cross-validated lower bound, which is decreased to .21, thus increasing the shrinkage estimate to .10. Next, we consider the calculation of power.

In the process of planning an experiment, a researcher is interested in detecting a medium effect size ($\rho^2 = .13$; Cohen, 1992) with a sample of 50 individuals and three predictors. To calculate power, the researcher enters "powerR2[50,3,.13,.05]," and Mathematica returns .56. If the researcher is not content with power of .56 and, instead, wants power to be .85, the researcher may enter "findsample[3,.13,.85,.05]," and Mathematica returns 90. In this situation, the researcher needs 90 subjects to detect a medium effect with power of .85.

Other Features

The Mathematica Package can also be used to make a few observations regarding the effect of number of predictors and sample size on the size of the confidence interval. To illustrate the relation between p and the size of the confidence interval, consider Figure 4. By increasing the number of predictors from two to six (while keeping R^2 at about .4), the width of the confidence

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

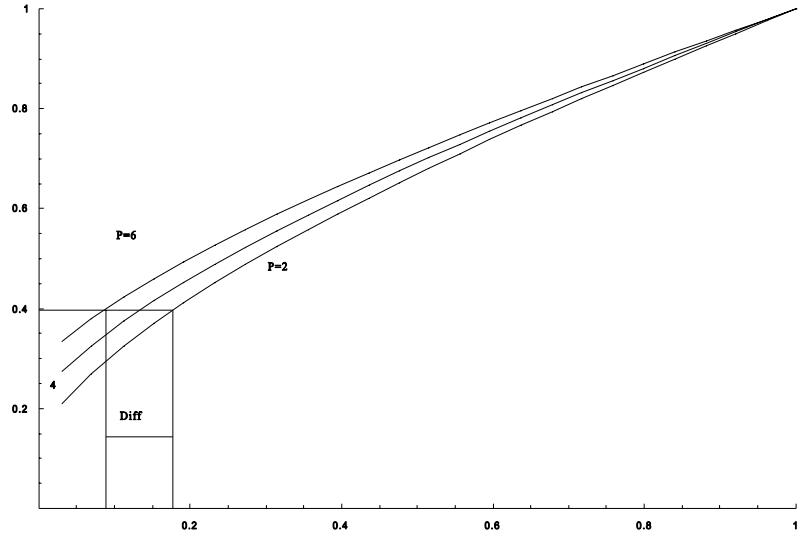


Figure 4. The effect of increasing p on the width of the confidence interval.

interval on ρ^2 is increased from .82 (1 – .18) to .92 (1 – .08). As p is increased, the size of the confidence interval is also increased. We would have to have an increase in R^2 of approximately .08 just to keep the size of the confidence interval at the same level. So, unless we could get an increase greater than .08 in R^2, we would decrease precision by increasing the number of predictors from two to six.

Our routine could be used to obtain this kind of information. We begin by typing "confidence[41,2,.4,.95]" in Mathematica after loading the package, and Mathematica returns the value of .175. The same operation for "confidence[41,6,.4,.95]" yields a value of .0875. We can see the decrease in precision as we moved from 2 to 6. By entering "confidence[41,2,.48,.95]" (= .1725), we see that it would take at least an increase of .08 in the R^2 to get back to the original level of precision. The evaluation between number of predictors and the size of the confidence interval provides yet another tool in trying to identify the "best" regression model. The package also may be used to illustrate the effect of p on power. We can show that for a given difference, as p increases, power decreases.

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

Program Accuracy

As a final word about accuracy, we have checked and compared our answers to many of the tables available in the literature, and they have always agreed. We have also taken great care in ensuring that the numerical algorithm stay in proper bounds. However, when n and p get large and R^2 gets small, the numbers in some of the numerical algorithms may get too large for the accuracy of some microcomputers. For the random functions up to an n of 200, we are confident that the results are accurate. For n greater than 200, the result should be accurate as long as p is not too large (>16) and R^2 is not too small (<.10). A researcher dealing with large n and p values combined with small values of R^2 should take a few precautions to make sure that the functions are returning the appropriate values. Because the algorithm involving the fixed functions are simpler than those involving the random functions, both functions should be executed to make sure that the values are close when n and p are both large. In addition, we can step down in sample size or p and rerun the function to ascertain that the value obtained earlier is within appropriate bounds. A final alternative is to pursue an approximation. Lee (1972) gives a normal approximation for R when n and p are large.

Program Syntax

Again, further information including the syntax necessary to use the package can be found in the appendix and in Table 1. We hope it is apparent that by using this package, a large amount of previously difficult-to-acquire information can now be obtained quickly and easily. Note that the package can be used with either the random or fixed model. The functions that deal with the fixed model have the word *fix* attached to them. We hope that other researchers will find this package to be a useful tool.

Appendix
Using the Package (multiple R^2) Functions in Mathematica 3.0

The Mathematica Package within which all the functions are located is named MultipleR2. Here, we offer instructions for accessing the package from both a floppy disk and a fixed disk. Furthermore, the names of each function within the package, the syntax for using most functions, and a brief description of the output of these functions is given. It should be noted that after loading the package, additional syntax and output information is available by typing a question mark (?) followed by the function name. Users of the package also should be aware that the package was written and compiled using Release 2, November 1997 or later. When using Release 1, April 1997, some apparent computational difficulties have occurred in a number of the functions when certain parameter values are used. Wolfram Research has been contacted regarding these problems and is working on a solution to the problem.

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

Before you begin, it is important to know that Mathematica commands are case sensitive. You must enter the commands and arguments exactly as they appear in this text. Also, to submit commands and arguments to Mathematica for evaluation, the shift key and the enter key must be pressed simultaneously.

Accessing the Package

If you are accessing the package from a floppy disk, at the Mathematica edit screen, with the disk in drive A, submit the following string:

<<A:\MultipleR2.m

All the functions within the MultipleR2 package will now be available after simultaneously pressing the SHIFT and ENTER keys.

If you are accessing the package from a fixed disk, we assume the platform under which you are running Mathematica 3.0 is Windows-95, 98, or NT. Although Mathematica runs essentially the same under other platforms, the file structure under other platforms may differ. If you wish to use the package from a hard drive, you must first copy the package into a subdirectory (folder) under Mathematica 3.0. Although it can be copied into an existing folder, we recommend creating a dedicated folder under Mathematica 3.0 for personal programs. This folder we call "myprograms." Once the package is copied into the desired folder, type the following to retrieve the package:

<<Foldername`MultipleR2'

where *foldername* is replaced with the name of the folder under which you have copied the package, and the ` is the grave accent symbol found in the upper-left portion of your keyboard. The functions within the package will now be available to you.

Using the MultipleR2 Functions

In this section are most of the functions available in the MultipleR2 package together with a description of the output of each. The function names are written in the correct and required case. As noted in this article, for all functions, N represents sample size, and p represents the number of predictors in the model.

The probr function (probr[N,p,r2,X]) returns the probability that $R^2 < X$ under the random regression model in which X is the parameter value against which the estimate from the sample is compared.

The confidence function (confidence[N,p,r2,plevel]) returns the upper or lower bound for a p-level confidence interval for ρ^2 given an observed R^2 under the random regression model.

The crossR2sample function (crossR2sample[p,rho2,pcross,diff]) returns the necessary sample size to obtain a cross-validated R^2 that does not differ from ρ^2 by more than "diff" with probability pcross.

The findsample function (findsample[p,Rho2,power,alpha]) for the given power and α computes the sample size necessary to reject the hypothesis that $\rho^2 = 0$ when $\rho^2 = R^2$ under the random regression model.

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016

The powerR2 function (powerR2[N,p,Rho2,alpha]) computes the probability of rejecting the hypothesis that $\rho^2 = 0$ when $\rho^2 = R^2$.

The crossR2confidence function (crossR2confidence[N,p,Rho2,pcross]) returns, for a given R^2, a one-sided confidence interval for the cross-validated R^2 at the pcross level.

In addition to the functions listed above, several other useful functions are included in the package. Furthermore, functions that return the fixed model equivalents for the random model values are included in the package.

References

Allen, L., Cipielewski, J., & Stanovich, K. (1992). Multiple indicators of children's reading habits and attitudes: Construct validity and construct correlates. *Journal of Educational Psychology*, *84*, 489-503.

Baril, G. L., & Cannon, J. T. (1995). What is the probability that null hypothesis testing is meaningless? *American Psychologist*, *50*, 1098-1099.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*, 997-1003.

Frick, R. W. (1996). The appropriate use of null hypothesis testing. *Psychological Methods*, *1*, 479-490.

Gatsonis, C., & Sampson, A. (1989). Multiple correlation: Exact power and sample size calculations. *Psychological Bulletin*, *106*, 516-524.

Hagen, R. L. (1997). In praise of the null hypothesis statistical test. *American Psychologist*, *52*, 15-24.

Lee, Y. S. (1972). Tables of upper percentage points of the multiple correlation coefficient. *Biometrika*, *59*, 175-189.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*, 130-149.

Mendoza, J. L. (1977). A note on the level of predictive precision of a fitted linear equation. *Psychometrika*, *42*, 145-147.

Mood, A., & Grabill, F. (1963). *Introduction to the theory of statistics*. New York: McGraw-Hill.

Park, C. N., & Dudycha, A. L. (1974). Cross validation approach to sample size determination for regression models. *Journal of the American Statistical Association*, *69*, 214-218.

Sampson, A. (1974). A tale of two regressions. *Journal of the American Statistical Association*, *69*, 682-689.

Schmidt, F. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods*, *1*, 115-129.

Downloaded from epm.sagepub.com at UNIV OF OKLAHOMA on January 20, 2016