**Putting Confidence Intervals on *R2*, *η2,*or *R*[[1]](#footnote-1)©**

 Giving a confidence interval for an *R* or *R2* or *η2* is a lot more informative than just giving the sample value and a significance level. So, how does one compute a confidence interval for *R* or *R2 R2* or *η2*?

**Bivariate Correlation**

 **Benchmarks for ρ.** Again, context can be very important.

* .1 is small but not trivial
* .3 is medium
* .5 is large

**Confidence Interval for** **ρ, Correlation Analysis.** My colleagues and I (Wuensch, K. L., Castellow, W. A., & Moore, C. H. Effects of defendant attractiveness and type of crime on juridic judgment. *Journal of Social Behavior and Personality,* 1991, *6,* 713-724) asked mock jurors to rate the seriousness of a crime and also to recommend a sentence for a defendant who was convicted of that crime. The observed correlation between seriousness and sentence was .555, *n* = 318, *p* < .001. We treat both variables as random. Now we roll up our sleeves and prepare to do a bunch of tedious arithmetic.

 First we apply Fisher’s transformation to the observed value of *r*. I shall use Greek zeta to symbolize the transformed *r*.

. We compute the standard error as . We compute a 95% confidence interval for zeta:  This give us a confidence interval extending from .51557 to .73643 – but it is in transformed units, so we need to untransform it. . At the lower boundary, that gives us , and at the upper boundary 

 What a bunch of tedious arithmetic that involved. We need a computer program to do it for us. My [Conf-Interval-r.sas](http://core.ecu.edu/psyc/wuenschk/SAS/SAS-Programs.htm) program will do it all for you.

 Suppose that we obtain *r* = .8, *n* = 5. Using my SAS program, the 95% confidence interval runs from -0.28 to +0.99.

 Don’t have SAS? Try [the calculator at Vassar](http://vassarstats.net/rho.html). This calculator is appropriate only for zero-order *r*. For multiple *R2*, try the calculator provided by [Daniel Soper](https://www.danielsoper.com/statcalc/calculator.aspx?id=28).





 **Confidence Interval for** **ρ2, Regression Analysis.** If you consider your predictor variable to be fixed rather than random (that is, you arbitrarily chose the values of that variable, or used the entire population of possible values, rather than randomly sampling values from a population of values), then the confidence interval for ρ2 is computed somewhat differently. The SAS program [Conf-Interval-R2-Regr](http://core.ecu.edu/psyc/wuenschk/SAS/Conf-Interval-R2-Regr.sas) can be employed to construct such a confidence interval. Don’t have SAS, [do it with SPSS or R](http://daniellakens.blogspot.com/2014/06/calculating-confidence-intervals-for.html), or use Daniel Soper’s [online calculator](https://www.danielsoper.com/statcalc/calculator.aspx?id=28).

F= **20.91** ;

df\_num = **2** ;

df\_den = **48** ;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*;

ncp\_lower = MAX(**0**,fnonct (F,df\_num,df\_den,**.95**));

ncp\_upper = MAX(**0**,fnonct (F,df\_num,df\_den,**.05**));

eta\_squared = df\_num\*F/(df\_den + df\_num\*F);

eta2\_lower = ncp\_lower / (ncp\_lower + df\_num + df\_den + **1**);

eta2\_upper = ncp\_upper / (ncp\_upper + df\_num + df\_den + **1**);

output; **run**; **proc** **print**; var eta\_squared eta2\_lower eta2\_upper; **run**;

| **Obs** | **eta\_squared** | **eta2\_lower** | **eta2\_upper** |
| --- | --- | --- | --- |
| **1** | 0.46560 | 0.27247 | 0.57713 |

 If you use the program above with *F* = *t*2 for the test of a partial correlation, it will return the squared partial correlation coefficient and its confidence interval.

You can also use SAS PROC GLM to get confidence intervals for *R2*, *sr2*, and *pr2*.

**proc** **glm** data=Sage; model Cyberloafing = Conscientiousness Age / EFFECTSIZE ALPHA=**.1**;

| **Obs** | **eta\_squared** | **eta2\_lower** | **eta2\_upper** |
| --- | --- | --- | --- |
| **1** | 0.46560 | 0.27247 | 0.57713 |

The GLM Procedure

| **Proportion of Variation Accounted for** |
| --- |
| **Eta-Square** | 0.47 |
| **90% Confidence Limits** | (0.27,0.58) |

 Although the output says “Eta-Squared,” when the model is linear with no categorical predictors the values are for *R2*, *sr2* and *pr2*.

Type III (unique) Effect Sizes

| **Source** | **DF** | **Partial Variation Accounted For** |
| --- | --- | --- |
| **Semipartial Eta-Square** | **Conservative90% Confidence Limits** | **Partial Eta-Square** | **90% Confidence Limits** |
| **Conscientiousness** | 1 | 0.2521 | 0.0916 | 0.3993 | 0.3205 | 0.1423 | 0.4580 |
| **Age** | 1 | 0.1486 | 0.0272 | 0.2955 | 0.2176 | 0.0649 | 0.3630 |

 **Confidence Interval for** **ρ2, Correlation Analysis.** Use R2, which is available for free, from James H. Steiger and Rachel T. Fouladi. You can download the program and the manual [here](http://www.statpower.net/Software.html#R2). Unzip the files and put them in the directory/folder R2. Navigate to the R2 directory and run (double click) the file R2.EXE. A window will appear with R2 in white on a black background. Hit any key to continue. Enter the letter O to get the Options drop down menu. Enter the letter C to enter the confidence interval routine. Enter the letter N to bring up the sample size data entry window. Enter 318 and hit the enter key. Enter the letter K to bring up the number of variables data entry window. Enter 2 and hit the enter key. Enter the letter R to enter the *R2* data entry window. Enter .308 (that is .555 squared) and hit the enter key. Enter the letter C to bring up the confidence level data entry window. Enter .95 and hit the enter key. The window should now look like this:



 Enter G to begin computing. Hit any key to display the results.



 As you can see, we get a confidence interval for *R2* that extends from .267 to .610.

 Hit any key to continue, F to display the File drop-down menu, and X to exit the program.

 **Bad News:** R2 will not run on Windows 7 Home Premium, which does not support DOS. It ran on XP just fine. It might run on Windows 7 Pro.

 **Good News:** You can get a free DOS emulator, and R2 works just fine within the virtual DOS machine it creates. See my document [DOSBox](http://core.ecu.edu/psyc/wuenschk/docs30/DOSBox.docx).

 **More Good News:** The programs that assume a regression model rather than a correlation model give you confidence intervals that differ little from those given by R2.

 The R2 program will not handle sample sizes greater than 5,000. In that case you can use the approximation procedure which is programmed into my SAS program [Conf-Interval-R2-Regr-LargeN.sas](http://core.ecu.edu/psyc/wuenschk/SAS/Conf-Interval-R2-Regr-LargeN.sas). This program assumes a regression model (fixed predictors) rather than a correlation model (random predictors), but in my experience the confidence intervals computed by R2 differ very little from those computed with my large *N* SAS program when sample size is large (in the thousands).

 **What Confidence Coefficient Should I Employ?** When dealing with *R2*, if you want your confidence interval to correspond to the traditional test of significance, you should employ a confidence coefficient of (1 - 2α). For example, for the usual .05 criterion of statistical significance, use a 90% confidence interval, not 95%. This is illustrated below.

 Suppose you obtain *r* = .26 from *n* = 62 pairs of scores. If you compute *t* to test the null that rho is zero in the population, you obtain *t*(60) = 2.089. If you compute *F*, you obtain *F*(1, 60) = 4.35. The *p* value is .041. At the .04 level, the correlation is significant. When you put a 95% confidence interval about *r* you obtain .01, .48. Zero is not included in the confidence interval. Now let us put a confidence interval about the *r2* (.0676) using Steiger & Fouladi’s R2.





 Oh my, the confidence interval includes zero, even though the *p* level is .04. Now lets try a 90% interval.





 That is more like it. Note that the lower limit from the 90% interval is the same as the “lower bound” from the 90% interval.

**Unstandardized Slopes**

 Standard computer programs will give you a confidence interval for the unstandardized slope for predicting the criterion variable from the predictor variable. For example, in SAS: **proc** **reg**; a: model ar = misanth / **CLB**; In SPSS, in the Linear Regression Statistics window, just remember to ask for confidence intervals.



**Multiple Correlation and Regression.**

 The programs mentioned above can be used to put confidence intervals on multiple *R2* too. They can also be employed to put a confidence intervals on squared partial correlation coefficients for single variables or blocks of variables. I would prefer to use the squared semipartial correlation coefficient (the amount by which *R2* increases when a variable or block of variables is added to an existing model). To convert *pr2* to *sr2* do this:  where A is the previous set of variables, B is the new set of variables, and Y is the outcome variable. For an example, see my document [Multiple Regression with SPSS](http://core.ecu.edu/psyc/wuenschk/MV/MultReg/Multreg-SPSS.pdf) .

**CI for Every Correlation in a Matrix**

 If you desire to obtain a CI for every element of a correlation matrix, directly from the raw data, check this out: [An SPSS Macro to Compute Confidence Intervals for Pearson’s Correlation](https://sites.google.com/a/lakeheadu.ca/bweaver/Home/statistics/spss/my-spss-page/rhoci).

Karl L. Wuensch, October, 2019.

* [Fair Use of this Document](http://core.ecu.edu/psyc/wuenschk/Fair-Use.htm)
* [SPSS Confidence Interval on *R2*](http://core.ecu.edu/psyc/wuenschk/SPSS/CI-R2-SPSS.docx)
* [Confidence Intervals on Partial *R*2 and on Standardized Beta Weights](http://core.ecu.edu/psyc/wuenschk/docs30/CI_Semipartial.docx)
* [Proc GLM CI Less Conservative](http://core.ecu.edu/psyc/wuenschk/docs30/CI_R2_Reg2_GLM.docx)
* [Return to Wuensch’s Statistics Lesson Page](http://core.ecu.edu/psyc/wuenschk/StatsLessons.htm)
1. © Copyright 2019, Karl L. Wuensch - All rights reserved. [↑](#footnote-ref-1)