Can a Regression Model with a Small R-squared Be Useful?

by KAREN GRACE-MARTIN

R^2 is such a lovely statistic, isn't it? Unlike so many of the others, it makes sense–the percentage of variance in Y accounted for by a model.

I mean, you can actually understand that. So can your grandmother. And the clinical audience you're writing the report for.

A big R^2 is always good and a small one is always bad, right?

Well, maybe.

I've seen a lot of people get upset about small R^2 values, or any small effect size, for that matter. I recently heard a comment that no regression model with an R^2 smaller than .7 should even be interpreted.

Now, there may be a context in which that rule makes sense, but as a general rule, no.

Just because effect size is small doesn't mean it's bad, unworthy of being interpreted, or useless. It's just small. Even small effect sizes can have scientific or clinical significance. It depends on your field.

For example, in a dissertation I helped a client with many years ago, the research question was about whether religiosity predicts physical health. (If you've been in any of my workshops, you'll recognize this example–it's a great data set. The model used frequency of religious attendance as an indicator of religiosity, and included a few personal and demographic control variables, including gender, poverty status, and depression levels, and a few others.

The model R^2 was about .04, although the model was significant.

It's easy to dismiss the model as being useless. You're only explaining 4% of the variation? Why bother?

But think about this. If you think about all of the things that might affect someone's health, do you really expect religious attendance to be a *major*

contributor?

Even though I'm not a health researcher, I can think of quite a few variables that I would expect to be much better predictors of health. Things like age, disease history, stress levels, family history of disease, job conditions.

And putting all of them into the model would indeed give better predicted values. If the *only* point of the model was prediction, my client's model *would* do a pretty bad job. (Perhaps the 70% comment came from someone who only runs prediction models).

But it wasn't. The point was to see if there was a small, but reliable relationship. And there was.

Do small effect sizes require larger samples to find significance? Sure. But this data set had over 5000 people. Not a problem.

Many researchers turned to using effect sizes because evaluating effects using p-values alone can be misleading. But effect sizes can be misleading too if you don't think about what they mean within the research context.

Sometimes being able to easily improve an outcome by 4% is clinically or scientifically important. Sometimes it's not even close enough. Sometimes it depends on how much time, effort, or money would be required to get a 4% improvement.

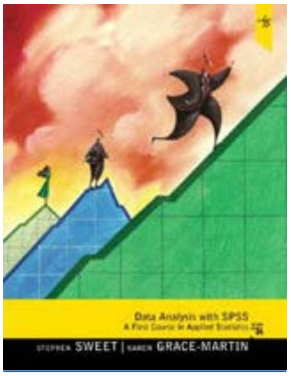
As much as we'd all love to have straight answers to what's big enough, that's not the job of any statistic. You've got to think about it and interpret accordingly.

Search

To search, type and hit enter

Connect

Read Our Book



Data Analysis with SPSS (4rd Edition)
by Stephen Sweet and Karen Grace-Martin

Statistical Resources by Topic

Missing Data

Books

Mixed and Multilevel Models

Categories

ANOVA (66)

Choosing a Statistical Test (8)

Confusing Statistical Terms (8)

Event History Analysis (7)

Factor Analysis (12)

Linear Regression (76)

Logistic Regression (42)

Missing Data (34)

Mixed and Multilevel Models (43)

Tagged as: effect size, R-squared

{ 21 comments... read them below or add one }

Serje Robidoux May 16, 2012 at 11:30 am

The counterargument to this position is that if you believe that religiosity is only a small piece of the puzzle, your model should include a whole lot of things that you think are more important as controls, and check whether the broader model with religiosity included is a better model than the one with only the big predictors. Otherwise you could be misattributing another health predictor to religiosity (e.g., hereditary health is probably a big predictor, and it may well be that people with unhealthy parents are more likely to seek a religious community too).

A model that only *improves* by small amounts can still be useful (say going from .7 to .74), but a model that, in it's entirety, only produces an R-sq of .04? I'd be worried that I haven't even begun to properly model the relationship.

I agree (strongly) with the point about interpreting the result within the context in which the research is being conducted, though.

REPLY

Karen May 16, 2012 at 3:26 pm

Hi Serje,

Yes, I see your point. I agree, it's always ideal to have more of the variation explained. And for an outcome that is generally well understood for the population being studied, there is a higher expectation of being able to explain most of the variation. You're absolutely correct that it would be better to model this hypothesis as an additional variation explained, and that not including the controls means you could be misattributing relationships.

However, there are some outcome variables (many in sociology, for example) for wide populations that just won't ever be explained that much. So it's not a matter of another variable that's being left out of a model, but either so many competing variables each with a tiny effect that you can't include them all or just randomness. (And I realize these are often the same thing).

Now it's arguable that physical health isn't one of those, and I concede that's possible. But it's possible that it is in certain populations. For example, you may be able to control for 70% of the variation in physical health in a clinical population, but not in a national population.

This is also true in more exploratory situations. If an outcome is a new construct that isn't well known, it's likely that data won't have been

collected on every possible control. In this case, it's very possible that an effect of something like religiosity will later be explained away in another study. But that's interesting–this effect we thought we had? Turns out it's explained by X. If we never report the first small effect because we're waiting for a model that explains everything, we may never know what needs to be built into the model.

Again, it's the context.

REPLY

Nico Grobler May 17, 2012 at 5:17 pm

Is there a way to quantify the 'context' in which one has to interpret R2?

REPLY

Karen May 18, 2012 at 11:45 am

Hi Nico,

I'm not exactly sure what you mean by quantifying the context, but I would think the answer is 'no.' It's really about stopping and thinking about what information you really have.

Karen

REPLY

Nico Grobler May 18, 2012 at 3:53 pm

It seems therefore that there is no hard rule to follow but it boils down to experience.

Thanks.

REPLY

Karen May 21, 2012 at 11:40 am

Hi Nico,

That's true in almost all of statistics. Even hard rules like p<.05 indicating statistical significance aren't really hard.

So yes, experience always helps, especially in understanding your variables and research. But stopping and thinking about it helps at any level of experience.

Karen

REPLY

Tom Lloyd May 23, 2012 at 10:08 am

It also depends on the type of model you run. In my experience time series models often get higher R2 than others. And if the dependent variable varies in magnitude a lot then the R2 will tend to be higher too. So conversely a poor model can quite happily get quite a respectable looking R2. It's definitely not the whole story.

REPLY

Karen May 25, 2012 at 1:38 pm

Good point, Tom.

REPLY

PieroCampa May 23, 2012 at 12:31 pm

Dear Karen,
your tips are so useful, you are my virtual teacher in the hazardous world of data modeling.
Thanks!

REPLY

Julian May 23, 2012 at 7:30 pm

Great article – always nice when your own opinion is reinforced by someone who's actually qualified in the area

Quick question – I have at times with multiple linear regression come across a significant independent association (beta value) of an independent var (that I'm interested in) with the dependent var. However – due to a small effect size, the model itself is not significant. My question is – can you report a significant independent association of 2 variables from a non-significant model?

The other thing to consider is that if the association between those 2 variables is the only thing you're interested in (after controlling for other variables in the model), you could do a partial correlation. Correct me if Im wrong, but I believe this would give the same result as your multiple lin regression beta value (and the same P value), but you wouldn't have a model R2 or p value to report. Am I missing something?

REPLY

Karen May 25, 2012 at 1:37 pm

Hi Julian,

Thanks, glad it was helpful.

To answer your question, if I were in that situation–non significant model, but significant coefficient on a key predictor, I would dig into it more to understand what is going on in the data. Run correlations on the predictors, run the model with and without the key predictor, run a bunch of scatterplots, both of the raw variables and of residuals.

In many fields, I've seen it's the norm to ignore the overall model F and just report coefficients. So can you report it? Yes. Should you? Hmm, maybe not.

Yes, the partial correlation gives you a measure of the association. What you lose there is not just those statistics, but the conceptual idea that one variable is an outcome to be predicted and the ability to come up with predicted values. So are you really trying to describe a relationship or model data?

Karen

REPLY

Julian May 29, 2012 at 1:34 am

Thanks – of course you would always do the necessary background with scatterplots and checking that the findings are not driven by an outlier, etc.
I guess I am talking about describing a relationship rather than modelling data. For example – you identify a significant correlation between 2 variables and would like to see if this is independent of a potential confounder. In this context, my impression is that a significant coefficient is still of interest (assuming a pre-specified analysis) even if the overall model is not significant. After all, its not your fault if what you thought was a confounder actually wasn't, right?

If I might ask a follow-up question, I've read of various guidelines regarding how many predictor variables can be included in a model. eg: 1 per 10 or 1 per 15 subjects in a dataset for linear regression (I'm in clinical research). If a model is 'over-fitted' (eg: 10 predictor variables for a sample of 20), how would that affect model significance?

REPLY

Spyros Makridakis January 12, 2013 at 8:00 am

Have you seen a scatter plot for even an R2 of 0.7. The dispersion of the data around the regression equation is so large that has a tiny predicted value (the reason is the predictive confidence interval is so large as to be of no practical value). I can ensure from my experience any R2<0.5 has very little predictive value beyond describing the model data.

Sorry for getting so late in this discussion but I am interested on the R2 values in medical studies and specifically in those dealing with hypertension research. Does any one knows their size, as no study mentions it? Also in regression it is extremely difficult to make sure that the model residuals are random, does anyone knows if this is done, again no study reports anything about testing the residuals?

REPLY

Karen January 16, 2013 at 10:16 am

Hi Spyros,

Agreed. A low R-squared means the model is useless for prediction. If that is the point of the model, it's no good.

I don't know anything specifically about hypertension studies and typical R-square values. Anyone else want to comment?

And it's a good point that most studies don't mention assumption testing, which is too bad. I assume it's because of space limitations in journals.

Karen

REPLY

Spyros Makridakis January 30, 2013 at 1:19 pm

Hi Karen,

Do you you think it is lack of space rather than the residuals are not random? I am not sure, but this (small R2 values) may explain the conflicting findings of the various studies.
Coming back to explaining the past versus predicting (a critical difference) is where the value of R2 is important. An R2 of .04 may explain the past data in a statistical significant manner and may have some value in doing so, but its predictive ability is practically zero when wanting to extrapolate beyond the available data.

REPLY

Andre May 1, 2013 at 9:45 am

Dear all,

I would like to add some complementary information about R2 and regression in general. First of all, I would recommand every researcher to explore the data with basic statistics and plots etc before undertaking a regression analysis and interpretating the results. Actually, it is quite rare to find linear relation in the nature (in social science as well) as the phenomena are most of the time very

complex. Instead of trying to prove linear relations even with low R2value and/or low p-value, it might be interesting to think about non-linear relations (polynomial, exp, logistical, etc…). This could be done by plotting the data. Most of the software suggest alternative tools to the linear regression. I hope it helps!

REPLY

Huong September 24, 2013 at 12:02 am

What if even after plotting the data, you still don't know what is going on? The analysis that I'm working on has R2=0.04, but the model fit has p-value<0.05 for either linear model or quadratic, cubic, exponential, logarithmic models. In that sense, I should pick the simplest one, right? But R2=0.04 can not imply linear relationship. So what kind of relationship do they have? And how do I find out? Thanks, Im glad I found this site and your reply!

REPLY

Karen September 25, 2013 at 10:51 am

Hi Huong,

Well, there may not be anything going on, or no discernible effects, anyway. Yes, start simple and see if you get an improvement in model fit with a more complicated model.

I would also suggest lots of graphing. Sometimes you can see the appropriate shape. Sometimes, not, though.

REPLY

Eduard May 2, 2013 at 3:20 pm

Good read, thanks!

I came across the same thing while doing economic research on capital gains tax for my thesis. I am not that experienced so it's nice to see my thoughts reinforced by someone much more credible than myself.

REPLY

rahayu June 16, 2013 at 10:13 pm

Hi!

I'm currently facing a similar experience – a very low R square for my model. I'm basically testing my model for causal- prediction and am using PLS methods for analysis. I need to somehow justify my results with some literature on this issue (low r square), but I find it difficult to find articles (journals) about this. Can anyone help?

Thanks!

REPLY

Luciano July 25, 2013 at 12:52 am

Err…I should say, if you feel the need to "somehow justify your (low R2) results" with "some" literature, you're taking a misled approach to this whole "science" thing. Sometimes hypothesis aren't confirmed by experiment. If that is what you inadvertently proved, it's your duty to report it as such.

REPLY

Leave a Comment

Name *

E-mail *

Website

Submit

PREVIOUS POST: Sample Size Estimates for Multilevel Randomized Trials

NEXT POST: Checking the Normality Assumption for an ANOVA Model

877-272-8096

Contact Us