Statistics By Jim

Making statistics intuitive

Basics   Hypothesis Testing   Regression   ANOVA   Fun   Glossary   Blog   My Store

The Difference between Linear and Nonlinear Regression Models

By Jim Frost — 32 Comments

The difference between linear and nonlinear regression models isn't as straightforward as it sounds. You'd think that linear equations produce straight lines and nonlinear equations model curvature. Unfortunately, that's *not* correct. Both types of models can fit curves to your data—so that's not the defining characteristic. In this post, I'll teach you how to identify linear and nonlinear regression models.

The difference between nonlinear and linear is the "non." OK, that sounds like a joke, but, honestly, that's the easiest way to understand the difference. First, I'll define what linear regression is, and then everything else must be nonlinear regression. I'll include examples of both linear and nonlinear regression models.

Linear Regression Equations

A linear regression model follows a very particular form. In statistics, a regression model is linear when all terms in the model are one of the following:

- The constant
- A parameter multiplied by an independent variable (IV)

Then, you build the equation by only adding the terms together. These rules limit the form to just one type:

Dependent variable = constant + parameter * IV + … + parameter * IV

Statisticians say that this type of regression equation is linear in the parameters.
However, it is possible to model curvature with this type of model. While the function

must be linear in the parameters, you can raise an independent variable by an exponent to fit a curve. For example, if you square an independent variable, the model can follow a U-shaped curve.

While the independent variable is squared, the model is still linear in the parameters. Linear models can also contain log terms and inverse terms to follow different kinds of

curves and yet continue to be linear in the parameters.

The regression example below models the relationship between body mass index (BMI) and body fat percent. In a different blog post, I use this model to show how to make predictions with regression analysis. It is a linear model that uses a quadratic (squared) term to model the curved relationship.

Nonlinear Regression Equations

I showed how linear regression models have one basic configuration. Now, we'll focus on the "non" in nonlinear! If a regression equation doesn't follow the rules for a linear

model, then it must be a nonlinear model. It's that simple! A nonlinear model is literally not linear.

The added flexibility opens the door to a huge number of possible forms. Consequently, nonlinear regression can fit an enormous variety of curves. However, because there are so many candidates, you may need to conduct some research to determine which functional form provides the best fit for your data.

Below, I present a handful of examples that illustrate the diversity of nonlinear regression models. Keep in mind that each function can fit a variety of shapes, and there are many nonlinear functions. Also, notice how nonlinear regression equations are not comprised of only addition and multiplication! In the table, thetas are the parameters, and Xs are the independent variables.
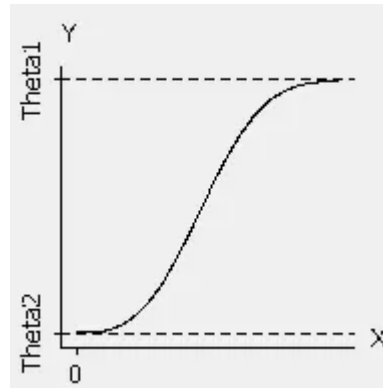
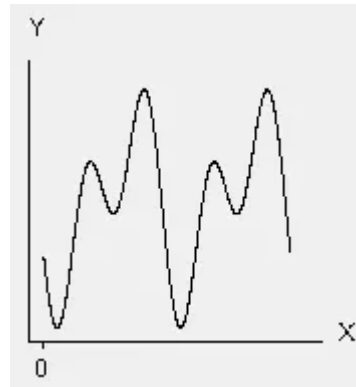| Nonlinear equation	Example form

Power:
$$\theta_1 * X^{\theta_2}$$
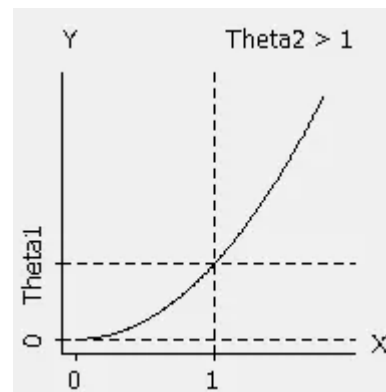
Weibull growth:

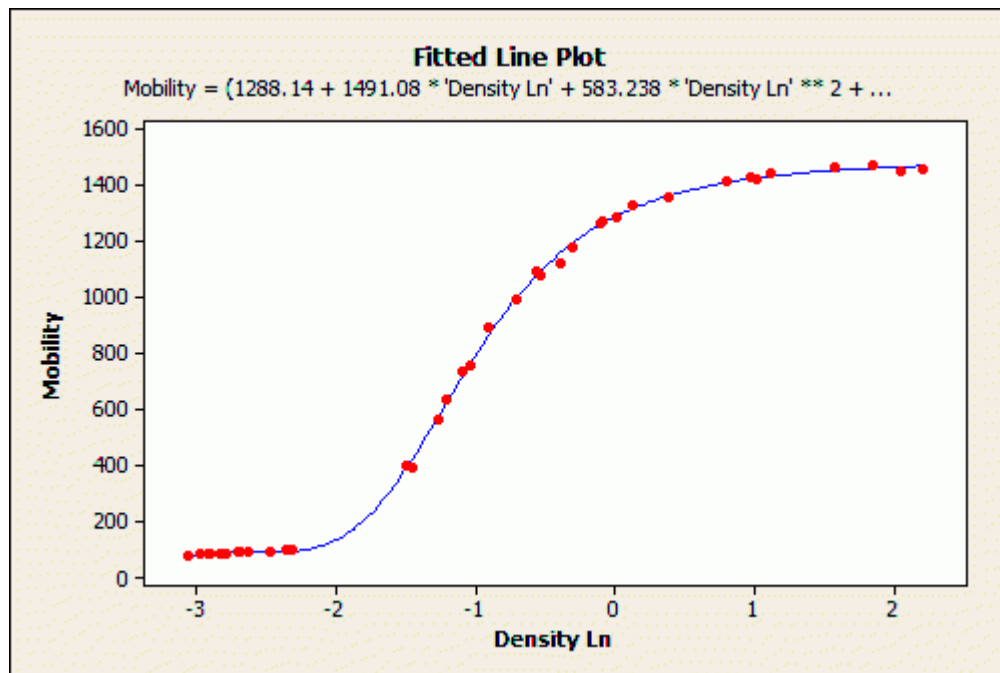$$\theta_1 + (\theta_2 - \theta_1) * exp(-\theta_3 * X^\theta 4)$$

Fourier:

$$\theta_1 * cos(X + \theta_4) + \theta_2 * cos(2 * X + \theta_4) + \theta_3$$

The nonlinear regression example below models the relationship between density and electron mobility.

The equation for the nonlinear regression analysis is too long for the fitted line plot:

Electron Mobility = (1288.14 + 1491.08 * Density Ln + 583.238 * Density Ln^2 + 75.4167 * Density Ln^3) / (1 + 0.966295 * Density Ln + 0.397973 * Density Ln^2 + 0.0497273 * Density Ln^3)

It's important to note that R-squared is invalid for nonlinear models and statistical software can't calculate p-values for the terms.

The defining characteristic for both types of models are the functional forms. If you can focus on the form that represents a linear model, it's easy enough to remember that anything else must be a nonlinear. Now that you understand the differences between

the two types of regression models, learn more about fitting curves and choosing
between them in the following blog posts!

- How to Choose Between Linear and Nonlinear Regression
- Curve Fitting using Linear and Nonlinear Regression

If you're learning regression, check out my Regression Tutorial!

***Note: I wrote a different version of this post that appeared elsewhere. I've
completely rewritten and updated it for my blog site.***

Share this:

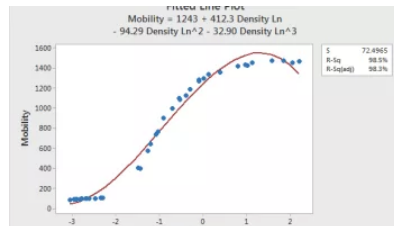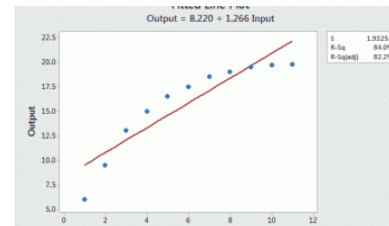| Share 20 | Share | Tweet | Save |

Related Posts on Statistics by Jim

How to Choose Between
Linear and Nonlinear
Regression
In "Regression"

Why Are There No P Values in
Nonlinear Regression?
In "Regression"

Curve Fitting using Linear and
Nonlinear Regression
In "Regression"

Filed Under: Regression   Tagged With: conceptual

Comments

Sotiris says
May 8, 2019 at 7:51 am

Hello there,really nice work! My questions is related to non linear regression.I am trying to create a predictive model using nls(in R). My formula goes like that: $Y \sim a*X*exp(b/Z)$, where "Y" is my dependent variable,"X" and "Z" are my independent variables and "a","b" are my coefficients. My question is what happens when the P value of "b" is more than 0.05 meaning it is not statisticaly significant. Does this mean that i should exclude b from the equation like we normally do in linear regression? Thank you in advance for your time!

★ Loading...

Reply

Jim Frost says
May 10, 2019 at 2:35 pm

Hi Sotiris,

Typically, in nonlinear regression, you don't see p-values for predictors like you do in linear regression. Linear regression can use a consistent test for each term/parameter estimate in the model because there is only a single general form

of a linear model (as I show in this post). In that form, zero for a term always indicates no effect. Consequently, the test for each model term tests whether the difference between the coefficient and zero is statistically significant. All of these tests use the same null and alternative hypotheses, as shown below:

H_0: b_i equals 0
H_A: b_i does not equal 0

This flexibility gives nonlinear regression great flexibility for fitting many types of curves. However, because of the many different forms, you can't assume that zero is the correct null hypothesis value for all parameter estimates. That depends on the function, the parameter's location in it, and the study area. Consequently, statistical software does not show p-value for parameter estimates in nonlinear regression.

Instead, have your software calculate confidence intervals and use your subject area expertise to identify meaningful values and determine whether the CIs include or exclude them. As for whether to include or exclude a term, that should never be determine solely based on statistical significance. You can use the general process I describe in my post about model specification for guidance.

I hope this helps!

★ Loading…

Reply

Steph says
May 1, 2019 at 10:27 pm

Hi,
I'm considering using non-linear regression for a thesis research project as my model will not be linear.
How do you calculate the number of participants needed? I use Gpower for linear regressions, what would be the process for non-linear?
Thanks
S

★ Loading...

Reply

Leonel Nava says
April 26, 2019 at 4:39 pm

FOlks you can convert a Non- Linear Equation(model) into a Linear One. Piece of cake, that is studied in any book of advance calculus.

★ Loading...

Reply

Jim Frost says
April 26, 2019 at 9:19 pm

Hi Leonel,

Yes, this is a known process. I talk about the ability to include nonlinear relationships in linear regression in my post about fitting curvature. However, it's not possible to transform all possible nonlinear relationships into a linear form. And, indeed there are benefits to fitting the raw untransformed data using nonlinear regression.

★ Loading…

Reply

statscurious says
February 18, 2019 at 10:39 am

If I'm reading correctly, I can turn any nonlinear model into a linear one. Say my independent variable is X. I do nonlinear regression which gets me some function G(X). Now if I make G(X) a variable in a linear regression, my model is technically a linear model, yes? At least this doesn't seem any different in principle from raising X to a power or transforming it with any other function which seems to be OK. Assuming this is OK, after I do my linear regression on G(X) is it ok for me to compare the R2 of that

"linear" regression against linear regression of just X? If that's ok I'm not sure I understand why we can't compare R2 of linear and nonlinear models.

★ Loading...

Reply

> ### Jim Frost says
> February 18, 2019 at 4:52 pm
>
> Hi, no, you can't turn all nonlinear models into linear models. Yes, you can use transformations to include *some* nonlinear functions into a linear model. But, you have to be able to express those functions in a linear form. See the example of using log functions in my post about modelling curvature. The log functions fit the linear model specification. I also show this in my discussion about log-log plots.
>
> As for the R-squared, if you can use a transformation in a linear model to fit an underlying nonlinear function, your software will give you an R-squared value. Be aware that the R-squared applies to the transformed data rather than the original data. The variance structure of the transformed data is completely different than the raw data. In this case, R-squared describes something fundamentally different. For transformations, use R-squared to understand how well the model fits the transformed data but do not think that it describes how well the model fits the original data.
>
> ★ Loading...
>
> Reply

aron bereket says

February 2, 2019 at 7:21 am

thank you for your effort .but i could not differentiate the real exact difference between linear and non linear .could you please assist /

★ Loading...

Reply

Jim Frost says

February 4, 2019 at 2:48 pm

Hi Aron, the key point to remember is the linear models follow the one form that I show in this post. If it doesn't follow that form, it is nonlinear. So, make sure you understand that form!

★ Loading...

Reply

Priscilla Branch says

December 24, 2018 at 1:45 pm

Also…could you sort the posts in descending order?

★ Loading...

Reply

Priscilla Branch says
December 24, 2018 at 1:43 pm

Thank you SO much for this refresher on statistics. I haven't touched this stuff in over 20 years, but your contributions are helping me dust off the cobwebs—AND have a better appreciation for the importance of outcomes that are studied and analyzed thoroughly. You present this material simply and elegantly. I pray that my analyses are also thoughtful, simple–and make sense!!

★ Loading...

Reply

Jim Frost says
December 25, 2018 at 12:21 am

You're very welcome! And, thanks so much for the kind comments. They really made my day!

I'm not sure what you mean by sorting the posts in descending order–most recent first to older? That should be how they're sorted currently.

I'm also working on a book about regression analysis that should be out in the first quarter of 2019. All of the content will be in a logical order in that book!

★ Loading…

Reply

Joy says
December 23, 2018 at 5:31 pm

Thank you so much for prompt reply. Just to clear another small confusion, I read somewhere that linear regressions are linear in parameters do $Y \sim \beta_2^{2} x_2 \ldots$ won't be linear, and that
is because we have quadratic on β_2, would not the same logic applied the above equation.

Thanks for the link, going over it now

★ Loading…

Reply

Jim Frost says

December 23, 2018 at 6:05 pm

Hi Joy, quadratics are commonly used to model curvature in linear models. They're ok! In fact, I use a quadratic in the BMI example in this very post. Also in this post, I talk about how linear models are linear in the parameters. I'd reread that section. Additionally, read the post that I mention in my previous comment. I think that'll help!

★ Loading...

Reply

Joy says

December 23, 2018 at 4:57 pm

Hi Jim,

I have a confusion about the core definition of non-linearity. I absolutely love your definition about writing the model as IV * parameter. I am wondering what happens if someone takes a non linear function on the parameters. That is Y ~ exp(\beta_1) * x_1 …. Is it still a linear model. In my definition yes, but it would be great if you shed some light.

★ Loading...

Reply

Jim Frost says

December 23, 2018 at 5:25 pm

Yes it is! And, to see a natural log transformation, you can read my post about modeling curvature, which is related way to use a linear model to fit an otherwise nonlinear function.

★ Loading…

Reply

Akis says

October 22, 2018 at 5:20 pm

Thank you for the answer. I really appreciate it. I hope you keep spreading comprehensive knowledge.

★ Loading…

Reply

Akis says
October 22, 2018 at 8:57 am

So, this is the first time reading your articles and I find them really interesting and comprehensive, good job with this. I have a quick question though. As you mentioned:

"Linear models can also contain log terms and inverse terms to follow different kinds of curves and yet continue to be linear in the parameters."

Then, what helps us understand which model is linear or not? I mean, if linear models can have inverse terms too, then why is the model above (density and electron mobility) a nonlinear one?

I am also trying to find some other sources that point out this difference and it would be greatly appreciated if u could link some references.

Thanks in advance.

★ Loading...

Reply

Jim Frost says
October 22, 2018 at 11:12 am

Hi Akis,

As I describe in this post, to be considered a linear model, the form of the model must fit a very specific format. However, you can transform the variables that fit

within this format. If it doesn't fit this very specific format, it's a nonlinear model. Again, it is based on the form of the model that I describe in this post.

The density and electron mobility is nonlinear because it doesn't fit that specific linear form that I describe.

Keep in mind that the difference between linear and nonlinear is the form and not whether the data have curvature. Nonlinear regression is more flexible in the types of curvature it can fit because its form is not so restricted. In fact, both types of model can sometimes fit the same type of curvature. To determine which type of model, assess the form.

I don't have a reference handy for you. However, these are the basic properties of these types of models and any textbook about linear models and nonlinear models will talk about these forms.

★ Loading...

Reply

---

shashi says
October 10, 2018 at 9:14 am

How you decide to use linear regression or non linear regression ?

★ Loading...

Reply

Jim Frost says
October 10, 2018 at 10:29 am

Hi Shashi,

I wrote a blog post about this topic specifically! How to Choose Between Linear and Nonlinear Regression.

I also talk about it in a post about curve fitting.

Read those two posts and you'll have your answer!

★ Loading...

Reply

Ashutosh Kumar says
September 17, 2018 at 8:17 pm

Dear Jim,

I have read a couple of your articles now and sharing it around as well and they are really very helpful and easy to interpret. Thank you for this !!

With regards to this post, my question is that – In cases when we use a higher order polynomial term in the linear regression model, to mode the curvature, does this not

fail the multicollinearity assumption because now we cannot change the value of x keeping x^2 constant?

★ Loading...

Reply

> Jim Frost says
> September 18, 2018 at 12:07 am
>
> Hi Ashutosh,
>
> Thank you so much for you kind words! And, thanks for sharing my articles. I really appreciate that!
>
> Yes! You're pretty much correct with that. However, the assumption only excludes perfect correlation. Some degree of correlation is OK but if it increases too much it becomes problematic. For information about the assumptions (including this one), read my post about the classical OLS assumptions.
>
> In terms of polynomials specifically, yes, these terms often increase multicollinearity to problematic levels. To correct for this type of multicollinearity, you can center the continuous variables. Read about this approach in my post about multicollinearity. This method typically reduces multicollinearity caused by polynomials to acceptable levels.
>
> I hope this helps!
>
> ★ Loading...

Reply

Natalie says
September 12, 2018 at 10:24 pm

You have helped me understand statistics!!!! Thank you

★ Loading...

Reply

Jim Frost says
September 12, 2018 at 11:27 pm

You're very welcome! I'm happy to hear that my website has helped you with statistics!

★ Loading...

Reply

jw says
September 9, 2018 at 4:12 am

Hey Jim,
I was just researching this topic and found a similar article elsewhere.

The time stamp on your article is older so I hope that they were copying from you and not the other way around. I just wanted to let you know about this, maybe you already did □

Cheers

★ Loading...

Reply

Jim Frost says
September 9, 2018 at 12:35 pm

Hi JW,

Thanks so much for pointing this out to me. Everything is OK because I am the author of both. I actually completely rewrote these articles so the text is different. They're basically different articles on the same topic. In fact, if you use the Internet Archive Wayback Machine and look at older versions, you'll see that I am listed as the author. For unknown reasons, the organization removed most authors' names from their blog posts.

Thanks again!

★ Loading...

Reply

Alisha Bansal says
August 14, 2018 at 7:14 am

Hi Jim, found this article very helpful. Many many thanks for posting this!

★ Loading...

Reply

Alee says
July 24, 2018 at 1:22 pm

Hello Sir!
if the property values is dependent variable and (house characteristics including, no of beds rooms, larger area, green spaces near to home and other amenities) are independent variables, how can we apply liner and non linear regression on them?

★ Loading...

Reply

Jim Frost says
July 24, 2018 at 2:39 pm

High Alee!

Long, long ago I had a professor who published a study about this exact subject! By the way, he found that the single thing a homeowner could do to increase the value of their home is to use add a bathroom! He used linear regression for his analysis.

As for how to perform this analysis, you simply use the sale prices as the dependent variable. And, you include all of the house characteristics as independent variables. In this manner, you can see how changes in the independent variables relate to changes in the average sale price of a home. For example, you can see how the average sales price changes when you add a square foot or add a bathroom!

I hope this helps!

★ Loading...

Reply

Amir says

November 21, 2017 at 10:35 am

Dear Jim,

The way you teach statistics is exclusive! It comes from deep experience, and this makes people to bypass all the fear and fuss around mathematical expressions – which were developed to explain the world around us. As Einstein said, "if you can not explain it simple you have not understand well". Thanks for explaining simple and thanks for understanding statistics well.

Amir

★ Loading...

Reply

> Jim Frost says
> November 21, 2017 at 2:39 pm
>
> Hi Amir, thank you so much for your kind words–that means a lot to me! I strongly believe that statistics doesn't have to be scary! I'm happy that you have found my posts helpful. By the way, that's a great quote from Einstein! –Jim
>
> ★ Loading...
>
> Reply

Leave a Reply
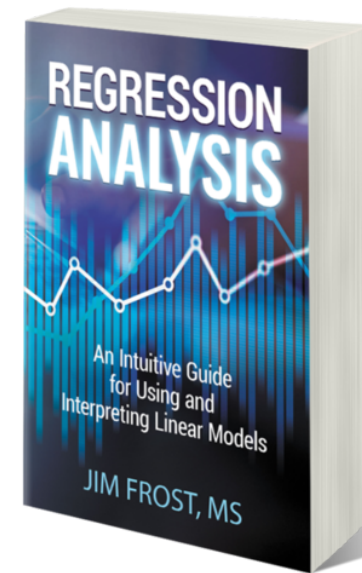
Enter your comment here...

Meet Jim

I'll help you
intuitively
understand statistics
by focusing on
concepts and using
plain English so you

can concentrate on understanding your
results.

Read More…

Buy My eBook!

Regression Analysis: An Intuitive Guide [ebook]

Over the course of this full-
length ebook, you'll progress
from a beginner to a skilled

practitioner. I'll help you intuitively understand regression analysis by focusing on…

$14.00 USD

Buy it now

Search this website

Subscribe via Email!

Enter your email address to receive notifications of new posts by email.

Email Address

Subscribe

Follow Me

Facebook

RSS Feed

Twitter

Popular   **Latest**

How To Interpret R-squared in Regression Analysis

How to Interpret P-values and Coefficients in Regression Analysis

Understanding Interaction Effects in Statistics

How to Interpret the F-test of Overall Significance in Regression Analysis

Measures of Central Tendency: Mean, Median, and Mode

Multicollinearity in Regression Analysis: Problems, Detection, and Solutions

The Importance of Statistics

Recent Comments

Jim Frost on Boxplots vs. Individual Value Plots: Graphing Continuous Data by Groups

Jim Frost on Check Your Residual Plots to Ensure Trustworthy Regression Results!

Jim Frost on Choosing the Correct Type of Regression Analysis

Tanvir Ahammed on Choosing the Correct Type of Regression Analysis

Sanne Nelissen on Check Your Residual Plots to Ensure Trustworthy Regression Results!