

advanced measurement (logs, quadratics, etc)

Adam Okulicz-Kozaryn
`adam.okulicz.kozaryn@gmail.com`

this version: Monday 20th February, 2023 11:12

outline

interpretation: transforming variables

- ◇ Lin: One unit change in X leads to a β_2 unit change in Y .
- ◇ Log-Lin: One unit change in X leads to a $100 * \beta_2$ % change in Y . (guj ed4:p180 fig6.4; ed5:p163 ex6.4)
- ◇ Lin-Log: One percent change in X leads to a $\beta_2/100$ unit change in Y . (guj: ed4:p182 fig6.5; ed5:p165-6 ex6.5)
- ◇ Log-Log (aka log-linear or “linear in logs”): One percent change in X leads to a β_2 % change in Y (elasticity).

links for logs practice

◇ https:

`//stats.idre.ucla.edu/other/mult-pkg/faq/general/
faqhow-do-i-interpret-a-regression-model-when-some-var`

◇ [*]http:

`//www-stat.wharton.upenn.edu/~stine/stat621/handouts/LogsInRegression.pdf`

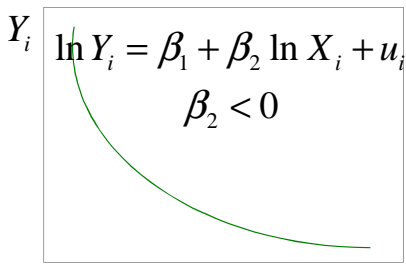
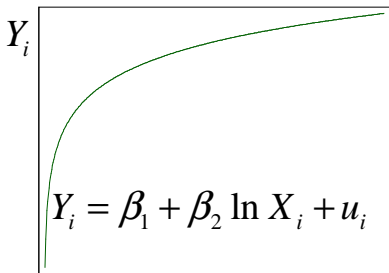
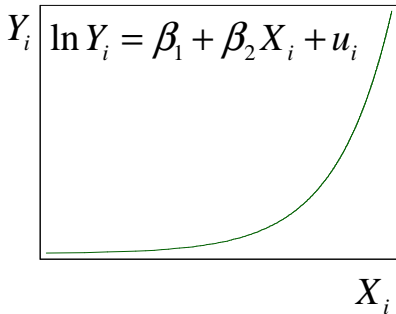
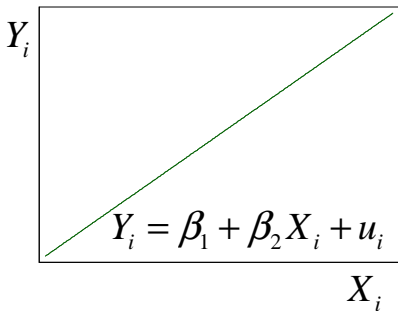
natural logarithm; for simplicity just log or ln

- ◇ `log()` or `ln()`, to reverse it `exp()`
- ◇ it tells us how many times 2.72 was multiplied, eg:
 - $\log(7.4) = 2$, because $2.72^2 = 7.4$
 - $\log(22166) = 10$, because $2.72^{10} = 22166$
- ◇ it compresses distribution!
 - it makes very big numbers relatively small
[*]https://en.wikipedia.org/wiki/Wheat_and_chessboard_problem
 - it is often used to remove outliers
 - say wage, prices, any \$ amounts are very skewed
 - and log will make them more normally distributed

log makes cool interpretations in regressions

- ◇ again it makes very big numbers small
 - actually, the bigger the number, the much more it is compressed
 - $\log(10) \approx 2$; $\log(100) \approx 4$; $\log(1000) \approx 7$; $\log(10,000) \approx 9$
- ◇ what you think will happen if you log transform your x?
- ◇ as the values of x increase, the effect on Y would be weaker
 - originally huge differences at high levels of X are now tiny
- ◇ what if we log transform Y variable? the other way round!
 - so the effect of X would be increasing

it makes a difference



lin-lin [dofile: measurement]

- ◇ eg people with more education earn higher wages...
- ◇ $Y_i = \beta_1 + \beta_2 X_i + u_i$
- ◇ This model specifies that the change is constant regardless of the level of X (because β is constant)
- ◇ $wage_i = \beta_1 + \beta_2 educ_i + u_i$
 - $\widehat{wage}_i = -0.75 + 0.75educ_i$
 - $\widehat{wage}_{10} = \$6.75 \quad \widehat{wage}_{11} = \$7.50 \quad \Delta \widehat{wage} = \0.75
the change is the same for any 1 year change in educ

relative change: log-lin

- ◇ take log of Y first, then regress $\ln Y$ on X
- ◇ regression treats $\ln Y$ the same as any other var
 - percent change in Y per unit change in X is
 - $100 * \beta_2$ times the unit change in X (for small changes)
- ◇ still a linear regression, but with a new DV: $\ln Y$.
 - not linear in terms of Y

eg log-lin [dofile: measurement]

- ◇ $\ln(wage_i) = \beta_1 + \beta_2 educ_i + u_i$
- ◇ $\widehat{\ln(wage)}_i = 1.06 + 0.08 educ_i$
- ◇ $\widehat{\ln(wage)}_{10} = 1.06 + 0.08(10) = 1.86$
- ◇ this is the predicted $\ln(wage)$
 - but what about the predicted wage?
- ◇ $\widehat{wage}_{10} = e^{1.86} = \6.42 **exp()**
- ◇ $\widehat{wage}_{11} = e^{1.94} = \6.96
- ◇ $\% \Delta \widehat{wage}_{10 \rightarrow 11} = \frac{\$6.96 - \$6.42}{\$6.42} = 0.08 = 8\%$

the change varies in dollar terms

- ◇ but let's examine the change in wage for an additional year of graduate school, eg master's degree years.
- ◇ $\widehat{\ln(wage)}_i = 1.06 + 0.08educ_i$
- ◇ $\widehat{\ln(wage)}_{17} = 1.06 + 0.08(17) = 2.42$ $\widehat{wage}_{17} = \$11.25$
- ◇ $\widehat{\ln(wage)}_{18} = 1.06 + 0.08(18) = 2.50$ $\widehat{wage}_{18} = \$12.18$
- ◇ $\% \Delta \widehat{wage}_{17 \rightarrow 18} = 0.08 = 8\%$
- ◇ the change in relative (percentage) terms is constant at 0.08 (8 percent), but the dollar change is larger.

lin-log [dofile: measurement]

- ◇ $Y_i = \beta_1 + \beta_2 \ln X_i + u_i$
- ◇ we generate the natural log of education and regress dollar wage on the log of education

eg: lin-log

- ◇ wage as a function of relative change in education
- ◇ $wage_i = \beta_1 + \beta_2 \ln(educ_i) + u_i$
- ◇ $\widehat{wage}_i = -10.15 + 7.54 \ln(educ_i)$
- ◇ $\widehat{wage}_{10} = -10.15 + 7.54 \ln(10) = 7.21$
- ◇ $\widehat{wage}_{11} = -10.15 + 7.54 \ln(11) = 7.93$
- ◇ $\% \Delta \widehat{wage}_{10 \rightarrow 11} = \0.72
- ◇ $\widehat{wage}_{17} = -10.15 + 7.54 \ln(17) = 11.21$
- ◇ $\widehat{wage}_{18} = -10.15 + 7.54 \ln(18) = 11.64$
- ◇ $\% \Delta \widehat{wage}_{17 \rightarrow 18} = \0.43
- ◇ for a 1% (0.01) change in X, the change in Y is $\beta_2/100$,
in this case 0.0754

relative change in education

educ	%change	educ	%change
1		11	10%
2	100%	12	9%
3	50%	13	8%
4	33%	14	8%
5	25%	15	7%
6	20%	16	7%
7	17%	17	6%
8	14%	18	6%
9	13%	19	6%
10	11%	20	5%

The relative

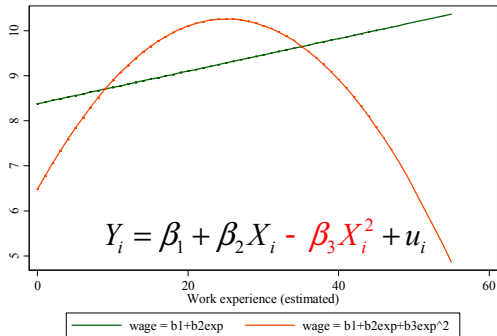
change in education per year is declining because the base is getting larger. So the lin-log model will predict a smaller impact on wage each year (see graph few slides back)

and now quadratic regression

- ◇ not bivariate regression anymore, but trivariate
- ◇ the third var is just a sq of 2nd var
- ◇ and it does what logs do—fits a curve as opposed to a line
- ◇ and i think it is more intuitive than logs or reciprocals !
- ◇ the idea is that quadratic coef is smaller than linear, and opposite sign
- ◇ but as X gets bigger, its square get huge, and so quadratic coef with opposite sign overpowers first term and curve flips to positive or negative

quadratic model

If a *non-linear relationship* between X and Y is suspected, a *polynomial function of X* can be used to model it.

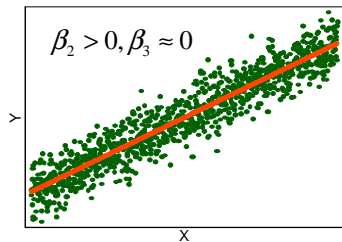
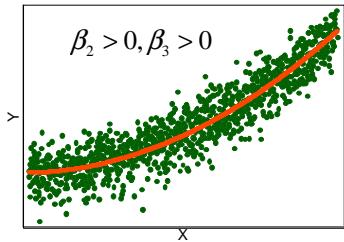
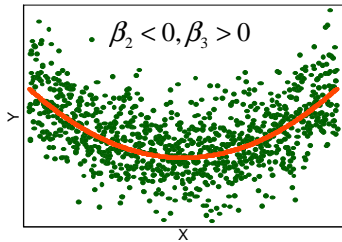
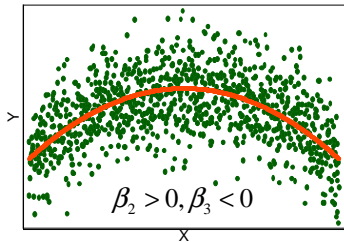


when it flips:

$$X_i^* = -\frac{\beta_2}{2\beta_3}$$

This curve reaches a maximum wage at the point where the marginal effect of experience is zero.

quadratic model



quadratic: interpretation

- ◇ the slope changes with X , it is not constant
- ◇ the best way to show the quadratic relationship is to graph it
- ◇ there is always a tipping point, but it may be outside the range of the data; in fact, the estimated line may be approximately linear for the observed data range even if the quadratic term is significant !
- ◇ the t test on squared term has a null hypothesis of linearity
 - if it is not significant, only linear term is left
- ◇ dofile: quadratic,bonus