Interpreting regression coefficients

How to interpret regression coefficients

When interpreting OLS coefficients, the key elements that you need to include depend on whether you're dealing with

1. a simple or multivariate regression, and
2. whether your variables are in levels or logs

Therefore, before you start, before you start interpreting, make sure that you

- know what y and x mean and their units
- check whether $x = 0$ or $x = 1$ (for logs) is plausible in your data
- don't claim causality ("effect of x"), instead say "associated with"
- in multiple regression, don't forget to mention ceteris paribus ("holding other explanatory variables fixed")

Simple regression

Level-level: $y = \beta_0 + \beta_1 x$

Intercept

- β_0 is the predicted value of [dependent variable] when the [explanatory variable] equals zero
- if asked, comment on whether this is economically meaningful (i.e. could the explanatory variable ever take on a value of zero (e.g., earnings) or is that a purely hypothetical scenario (e.g., height))

Slope

- a one-unit increase in [explanatory variable] is associated with a β_1 increase/decrease in [dependent variable]
- include appropriate units for both variables and don't forget the direction of the change (positive negative)

Example

$$\text{wage} = 5.2 + 0.75 \times \text{education}$$

with wage denoting hourly wage and education measured in years

- Intercept: The predicted hourly wage for a person with zero years of education is $5.20. The meaningfulness of zero education will depend on the context: in some low-income countries it's possible that a respondent will have no formal schooling, whereas if you're looking at data from Australia, the intercept has a less meaningful interpretation.

- Slope: A one-year increase in education is associated with a $0.75 higher hourly wages, on average.

Log-linear: $\ln(y) = \beta_0 + \beta_1 x$

Intercept

- The predicted value of [dependent variable] when the [explanatory variable] equals zero is around e^{β_0} units
- Note that we have to exponentiate the intercept to get back original units

Slope

- A one-unit increase in [explanatory variable] is associated with a $(\beta_1 \times 100)\%$ change in [dependent variable]
- Note that we multiply the coefficient by 100 to get percentage
- Make sure to include correct units for the explanatory variable

Example

$$\ln(\text{wage}) = 1.8 + 0.12 \times \text{education}$$

with wage denoting hourly wage and education measured in years

- Intercept: Our model predicts that a person with zero years of education would have an hourly wage of around $e^{1.8} \approx \$6.05$. Depending on whether zero years of education is realistic in our data, the intercept may serve mostly to anchor our regression line properly
- Slope: each additional year of education is associated with approximately $0.12 \times 100 = 12\%$ higher wages on average

---

Linear-log: $y = \beta_0 + \beta_1 \ln(x)$

Intercept

- When the explanatory variable equals one (remember that $\ln(1) = 0$), the predicted value of [dependent variable] is β_0 units
  - Don't forget the units of the dependent variable
  - Note that the intercept here occurs when your explanatory variable equals one, since it's in logs!

Slope

- A 1-percent increase in [explanatory variable] is associated with a $\frac{\beta_1}{100}$-unit change in [dependent variable]
- Note that we divide the coefficient by 100 to get the correct units
- Don't forget to include units for the dependent variable

Example

$$\text{consumption} = 2000 + 800 \times ln(\text{income})$$

with consumption and income both in dollars

- Intercept: The predicted consumption for a person with income of $1 is $2000
- Slope: A one-percent increase in income is associated with $8-dollar greater consumption on average. Note that we divide the estimated coefficient by 100 to get the correct units

---

Log-log: $\ln(y) = \beta_0 + \beta_1 \ln(x)$

Intercept

- When [explanatory variable] equals one unit, the predicted value of [dependent variable] is e^{β_0} units
- Remember
  - exponentiate to recover units
  - $\ln(1) = 0$

Slope

- A one-percent increase in [independent variable] is associated with a β_1 percent change in [dependent variable]
- Alternatively: β_1 represents the elasticity of [dependent variable] with respect to [explanatory variable]
- The coefficient in a log-log regression gives the elasticity directly

Example

$$\ln(\text{quantity}) = 4.5 - 1.2 \times \ln(\text{price})$$

- Intercept: When the price is $1, the predicted quantity demanded is $e^{4.5} \approx 90$ units
- Slope:
  - A one-percent increase in price is associated with a 1.2% decrease in quantity demanded, or
  - The price elasticity of demand is -1.2

Multivariate regression

When interpreting coefficients in multivariate regression, always remember to include the ceteris paribus (all else equal) qualifier!

Example

$$\text{wage} = 2.1 + 0.65 \times \text{education} + 0.43 \times \text{experience}$$

with wage denoting hourly wages, education and experience measured in years

- Intercept: our model predicts that a person with zero years of education and experience would earn $2.1/hour on average

- Slope on education: one more year of education is associated with 65 cents higher hourly wages, on average, holding experience constant
- Slope on experience: one more year of experience is associated with 43 cents higher hourly wages on average, holding education constant

---

General tips

- Remember that the approximations for log transformations work best for small changes (review the week 2 Essential Concepts recordings if you are feeling fuzzy)
- Be careful about units
- Always state whether the coefficient is positive/negative (i.e. the predicted change in the dependent variable is an increase or a decrease)
- I've said it above, but I'll say it again: for multivariate regression, always remember the ceteris paribus (all else constant) assumption