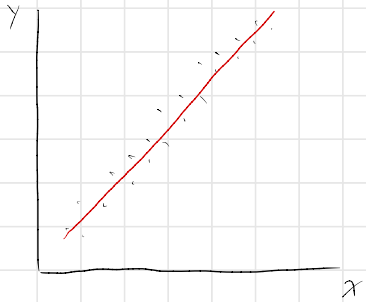
Feb 13th



OLS: minimize residuals

$y = \alpha + \beta_x$

Case whether statistically significant

t-statistic: $\left| \frac{\beta}{\text{Std. error}} \right| > 2 \rightarrow$ statistically significant

dummy variable: use for categories

omitted variable bias: might find spurious correlation
- include control variables as R&D variable

Group Exercise: Regression Table

Determinants of State High School Graduation Rates (standard errors in parentheses)		
	1910	1928
log per capita wealth	0.236	0.852
	(0.090)	(0.370)
% labor force in manufacturing	-0.067	-0.144
	(0.034)	(0.097)
South (0/1 dummy variable)	-0.0449	-0.0935
	(0.010)	(0.030)
New England (0/1 dummy variable)	0.0444	0.1
	(0.010)	(0.030)
Constant	-0.136	-0.468
	(0.07)	(0.27)

Source: Goldin and Katz, "Human Capital and Social Capital," Table 1

Questions

1. What does each column represent?
   HS grad rates in 1910 & 1928, regressed against both factors
2. What is the dependent variable?
   HS grad rate
3. What are the independent variable(s)?
   pc wealth, % in manufacturing, South, NE
4. For each variable, what are?

   Top #'s: coefficient

   Bottom #'s in parentheses: standard error

For the questions below, only answer for the 1910 regression.

5. Suppose we have a state in New England with the following made-up values in 1910 (chosen so that the math is easy!): log per capita wealth = 2, % LF in manufacturing = 20% (or 0.20). What is that state's predicted graduation rate?

~36 %

6. Suppose two states are alike in all regards, except that one state is in the South and one is in New England. What is the predicted difference in high school graduation rates between these two states?

~9 %

7. Suppose two states are alike in all regards, except one state has 100% of its labor force in manufacturing (i.e. 1.0) and the other 0% (i.e. 0.0). What is the predicted difference in high school graduation rates between these two states?

~7 %

8. Two states are alike in all regards, except one state has 50% of its labor force in manufacturing (i.e. 0.5) and the other 30% (i.e. 0.30). What is the predicted difference in high school graduation rates between these two states?

~1 %

9. What variables are statistically significant?

pc wealth

Another Exercise: "Do students go to class? Should they?"

UC Berkeley professor David Romer took attendance at six meetings of his Intermediate Macro course. He regressed final class grade (on a 4.0 scale) on the fraction of class meetings attended:

	(1)
Constant	1.25 (0.27)
Fraction of meetings attended	2.19 (0.35)
Sample size	195
R^2	0.31

Source: Romer, David. "Do students go to class? Should they?" *Journal of economic perspectives* (1993): 167-174.

Questions

1. Please identify the following:
   - Dependent variable: GPA
   - Independent variable(s) proportion of class attended
   - # students in sample: 195
   - R^2 (measures the fraction of the variation of the dependent variable associated with variation in the independent variable(s)): 0.31

2. Is the variable "fraction of meetings attended" statistically significant? Why or why not?
   Yes, $t > 2$

3. What is the predicted course grade for a student who attends
   - No class meetings: 1.25
   - Half of class meetings: 2.3
   - All class meetings: 3.34

4. What omitted variable might bias the results?
   time spent studying

Motivation may be an omitted variable.

- Students who attend more classes may study more, so the variable on attendance may be picking up some of the effects of being more academically motivated.

To reduce omitted variable bias, David Romer estimates another regression that includes a variable designed to control for motivation: Fraction of problem sets completed.

- Students who complete more problem sets are likely more motivated and hardworking.

	(1)	(2)
Constant	1.25 (0.27)	1.07 (0.23)
Fraction of meetings attended	2.19 (0.35)	1.74 (0.46)
Fraction of problem sets completed		0.60 (0.32)
Sample size	195	195
R^2	0.31	0.33

5. Is the variable "Fraction of meetings attended" still statistically significant?

Yes