Active Learning in Statistics Classes

Andrew Gelman
Dept of Statistics and Dept of Political Science
Columbia University, New York

Workshop at Oxford University, 19 Sept 2013

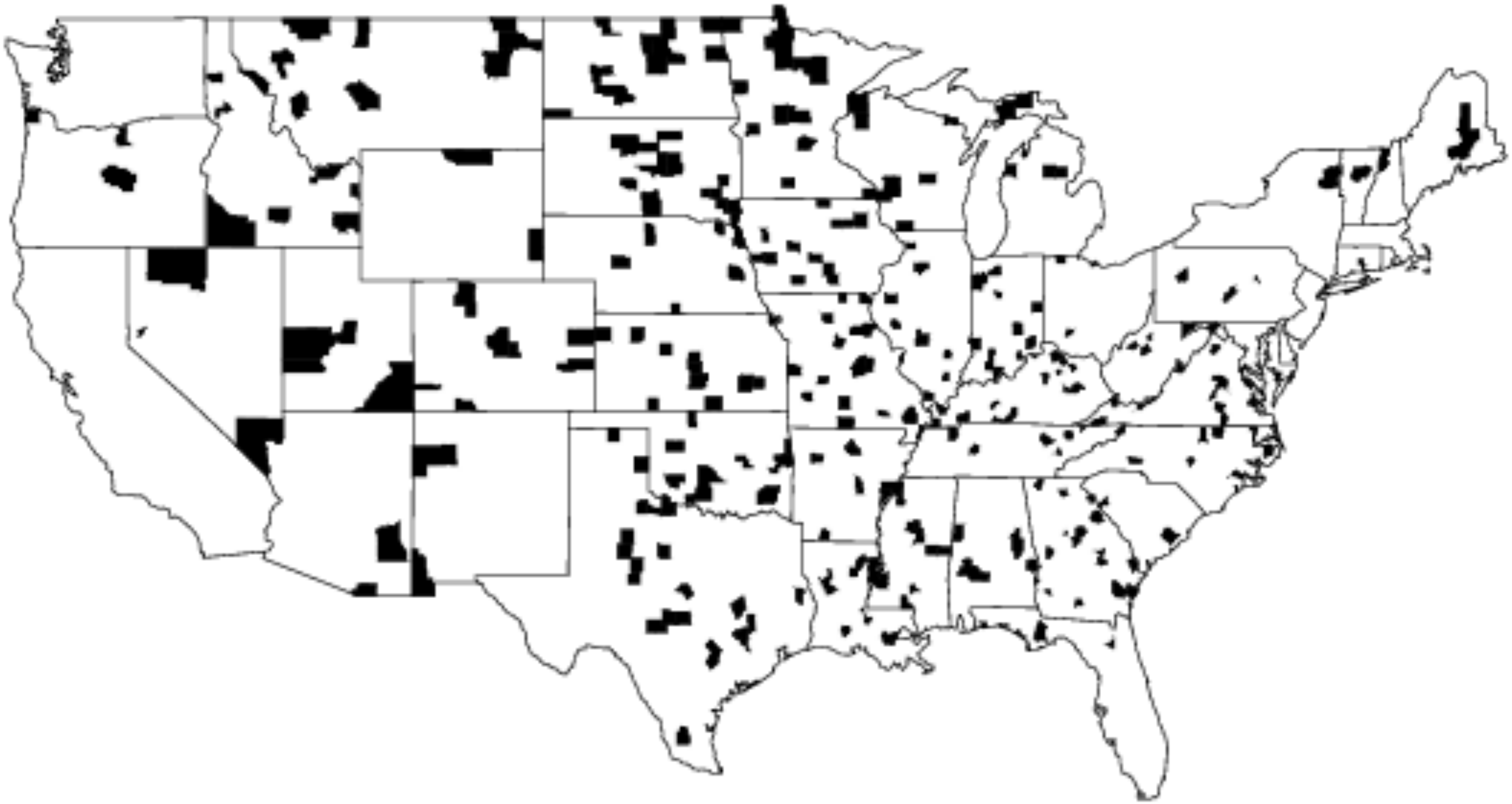
Each pair of you should:

1. Pull 5 candies out of the bag
2. Weigh the candies together
3. Write down the weight
4. Put the candies back in the bag!
5. Pass the scale and bag to your neighbors
6. Silently multiply the weight of the 5 candies by 20

The rules

- Work in pairs
- You can choose your 5 candies using any method—systematic sampling, random sampling, whatever
- Whoever guesses closest to the true weight gets the whole bag

Highest kidney cancer death rates

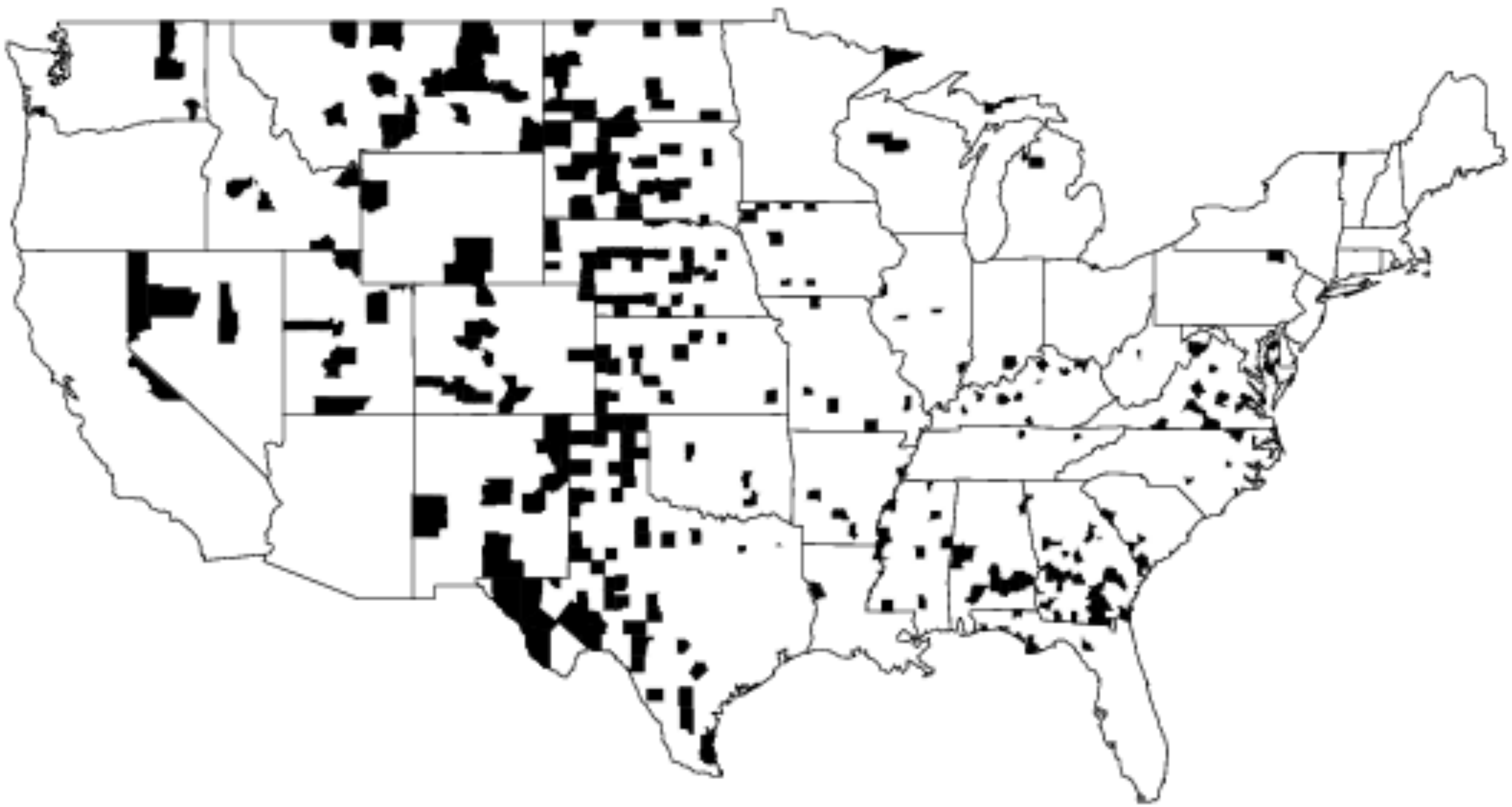


Why are most of the shaded counties in the center-west of the country?

Why are the counties with highest kidney cancer mostly in the center-west?

- Some possible explanations:
  - Pollution in farm areas
  - Poor medical care
  - More old people get cancer

Lowest kidney cancer death rates



Also in the center-west . . .

Story of the two tests

- Test A: 100 questions, bell-shaped distribution of scores
- Test B: 1 question, your score is 0 or 100
- You're trying to get into grad school: should you take Test A or Test B?
- Connection to cancer maps

Please indicate which hand you use for each of the following activities by putting a + in the appropriate column, or ++ if you use would never use the other hand for that activity. If in any case you are really indifferent, put + in both columns. Some of the activities require both hands. In these cases the part of the task, or object, for which hand preference is wanted is indicated in parentheses.

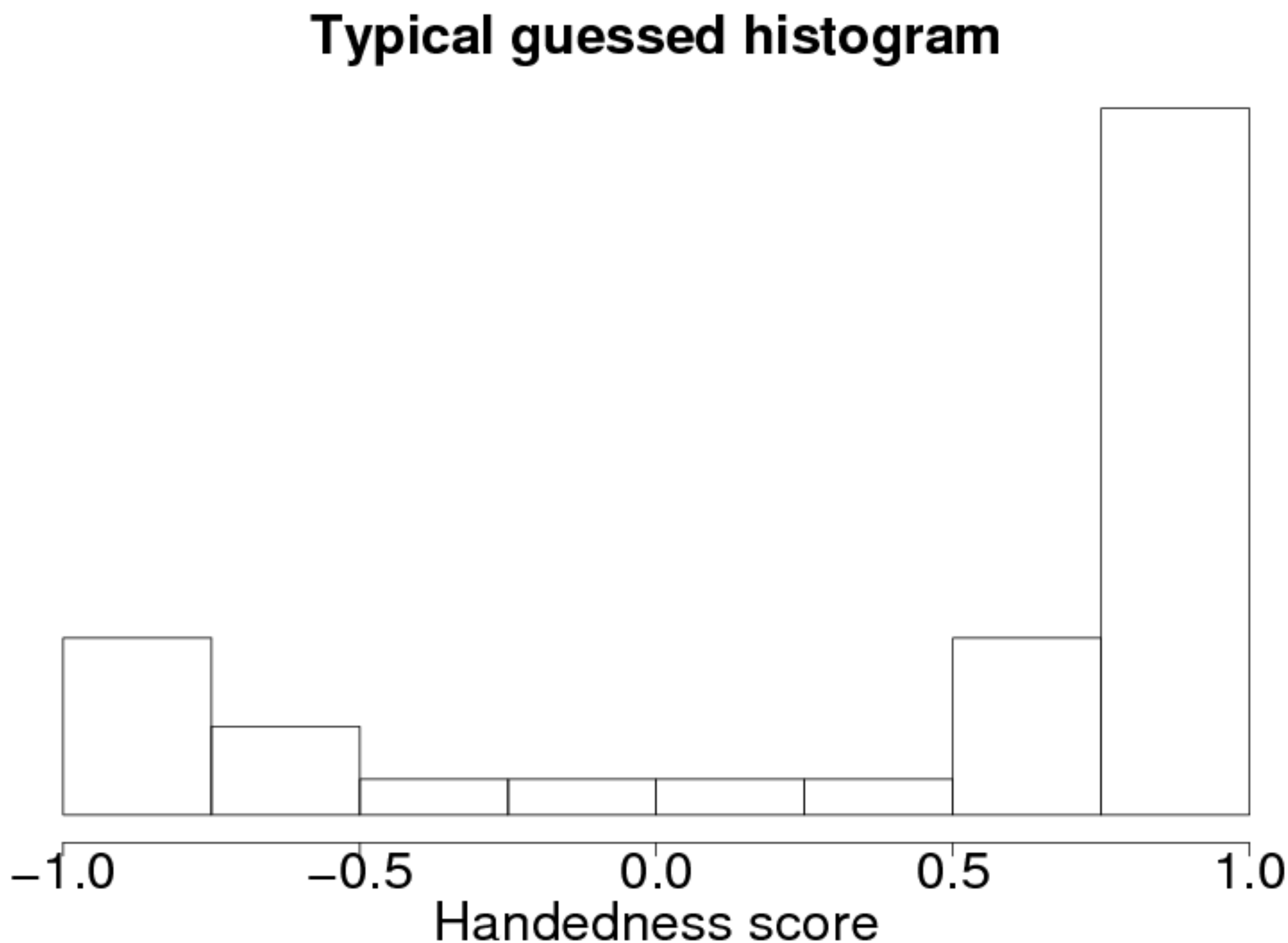
Task	Left	Right
Writing		
Drawing		
Throwing		
Scissors		
Toothbrush		
Knife (without fork)		
Spoon		
Broom (upper hand)		
Striking match (hand that holds the match)		
Opening box (hand that holds the lid)		
Total		

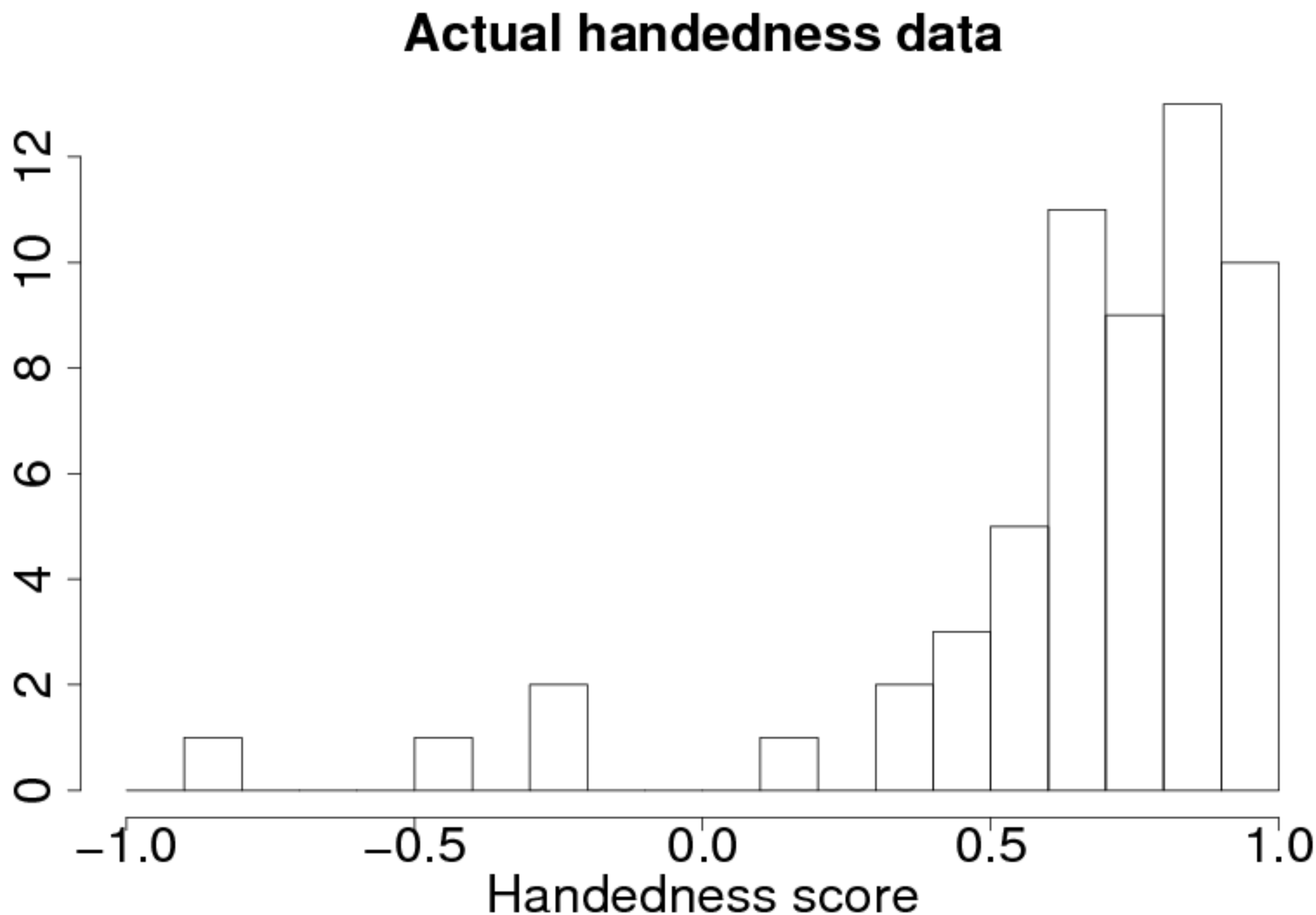
Right − Left: Right + Left: $\frac{\text{Right} - \text{Left}}{\text{Right} + \text{Left}}$:

Create a Left and a Right score by counting the total number of + signs in each column. Your handedness score is (Right − Left)/(Right + Left): thus, a pure right-hander will have a score of score $(20 - 0)/(20 + 0) = 1$, and a pure left-hander will score $(0 - 20)/(0 + 20) = -1$.

Your handedness and others

- Fill out the handedness inventory and compute your score
- Sketch a histogram of what you think the distribution of handedness scores will look like
- Scores range from -1 (pure lefty) to +1 (pure righty)


Typical guessed histogram
−1.0
−0.5
0.0
0.5
1.0
Handedness score


Actual handedness data
0
2
4
6
8
10
12
−1.0
−0.5
0.0
0.5
1.0
Handedness score

Regression of earnings on height

```
. regress earn height
  Source |       SS       df       MS              Number of obs = 1379
---------+------------------------------           F(  1,  1377) = 137.21
   Model |  4.8773e+10     1  4.8773e+10           Prob > F      = 0.0000
Residual |  4.8948e+11  1377   355470204           R-squared     = 0.0906
---------+------------------------------           Adj R-squared = 0.0900
   Total |  5.3826e+11  1378   390606004           Root MSE      = 18854
------------------------------------------------------------------------------
    earn |      Coef.   Std. Err.       t     P>|t|    [95 Conf. Interval]
---------+--------------------------------------------------------------------
  height |   1563.138   133.4476     11.713   0.000    1301.355     1824.92
   _cons |  -84078.32   8901.098     -9.446   0.000   -101539.5   -66617.15
------------------------------------------------------------------------------
. graph earn yhat height, connect(.s) symbol(Oi) xlabel ylabel
```

Graph the regression line and the data
(consistent with the Stata output)

Earnings and height example

- Graphs on graph paper and on the blackboard
- How did it feel to make the graphs?
- How did it feel to work in pairs?
- What skills are the students learning?

Teaching multiple regression

- Usual focus is on normal distribution, statistical significance, p-values, etc.
- But . . . the *deterministic* part of the model is most important

Regressions of earnings on height

- Earnings = –84000 + 1560*height + err
  – Height in inches
  – Graph the line
- Earnings = 1600 + 550*height – 11300*sex + err
  – Sex: 1 for men, 2 for women
  – Graph the parallel lines
- Earnings = –41000 + 1200*height + 16000*sex
  – 400*height*sex + err
  – Graph the non-parallel lines

Example: grading on a curve

In pairs, work on these questions:

- How to assign grades?
- What are some possible systems? What is best?
- What are your goals?
- How could you design a study and gather evidence to decide what grading system to use?

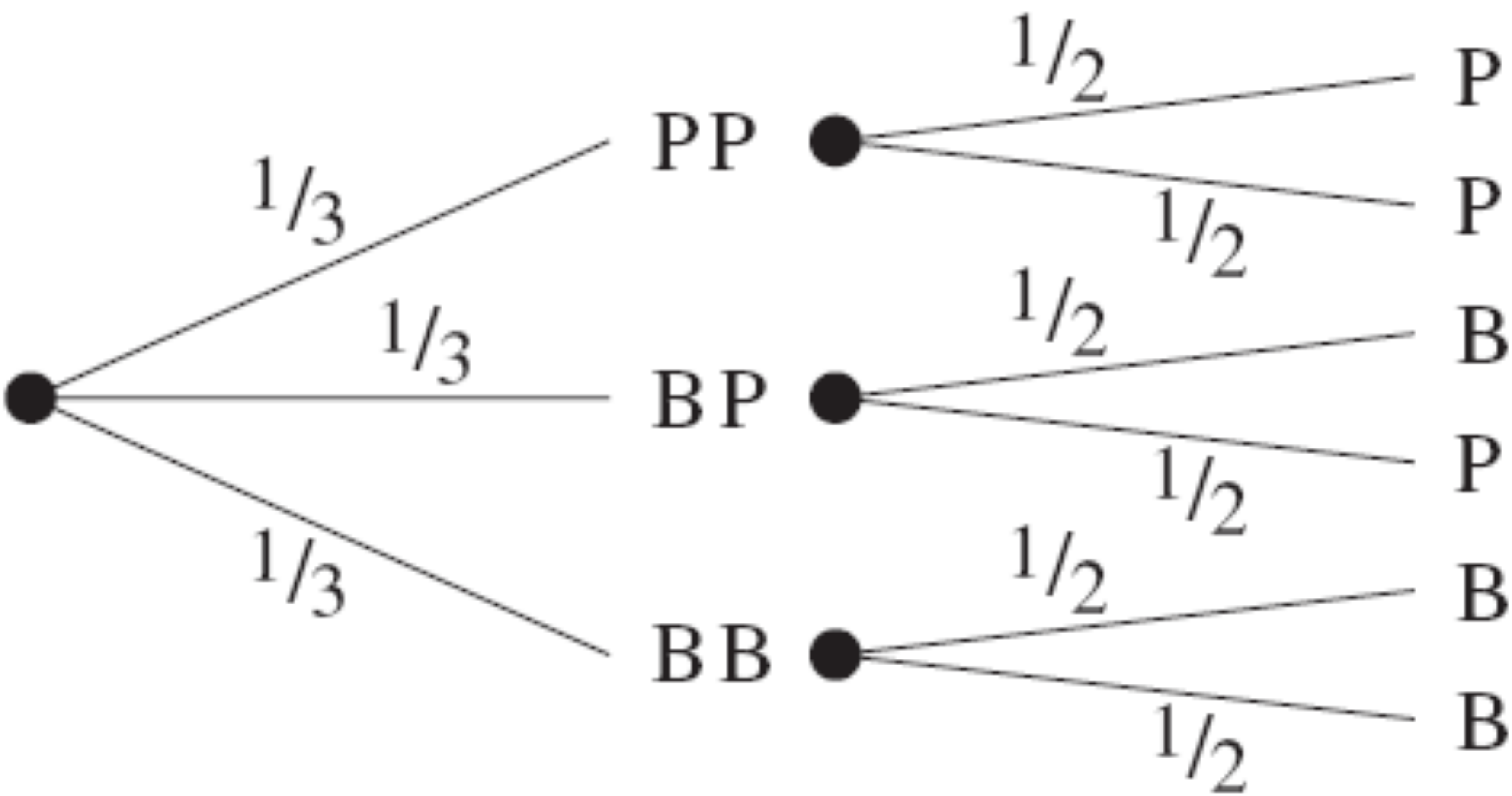
Experimenting with exam questions

- Half the students get one midterm exam form, half get another
- Compute average grades on each
  - Form A: avg grade is 65
  - Form B: avg grade is 70
- Should grades be adjusted?

Probability demonstrations

- Classic examples
  - Birthday problem
  - Monty Hall
  - Three cards
- How to get student involvement
- Avoiding “trickiness”
- Key techniques
  - Probability trees
  - Probabilities as frequencies

Probability tree for the 3-card example



More advanced material

- We still do demonstrations, work in pairs, etc.

Subjective uncertainty bounds

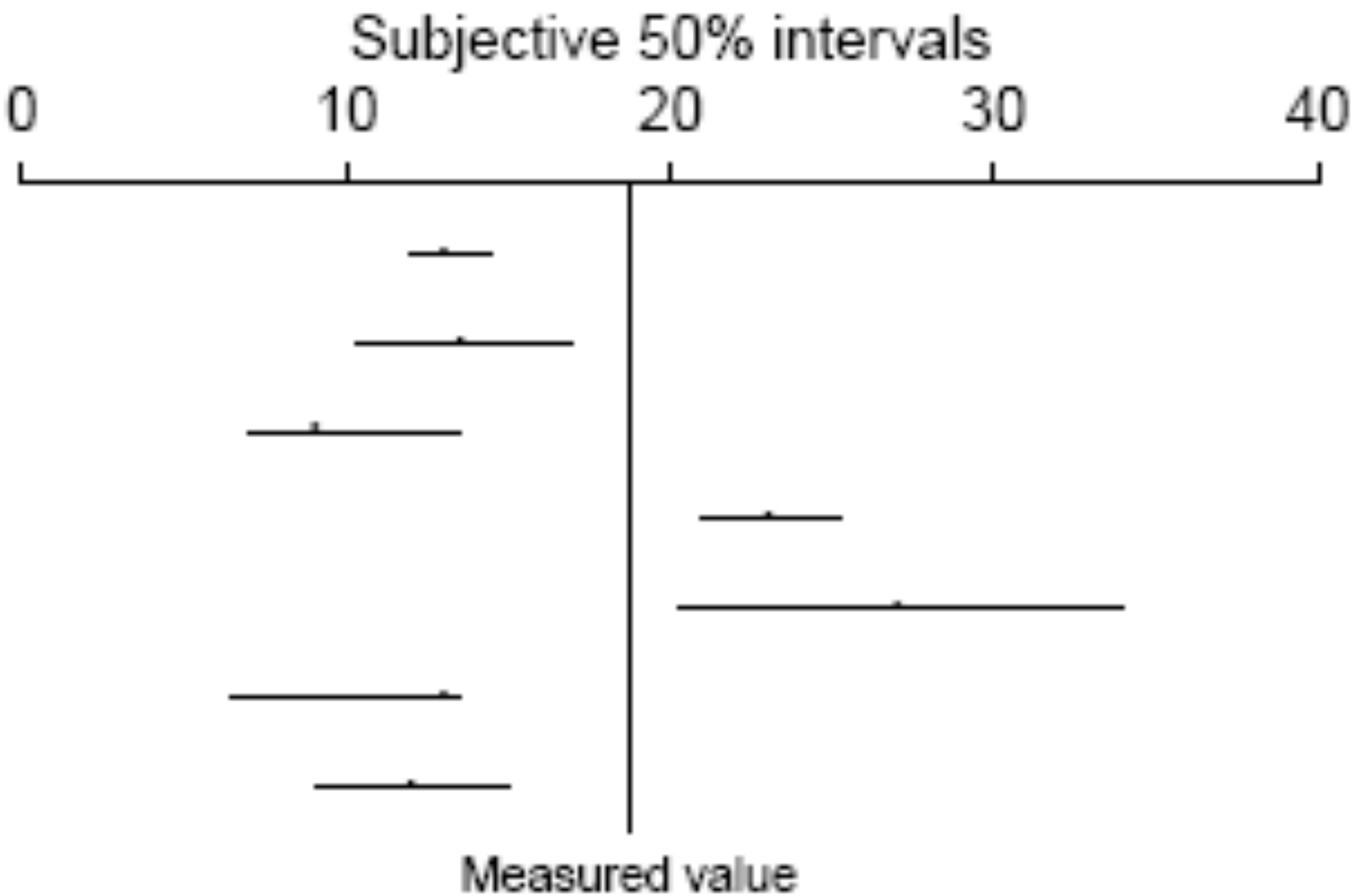
Uncertain quantity	25% lower bound	75% upper bound
% black		
# eggs		
# airline deaths		
% girl births		
% freshmen in phys sciences		
# French speakers		
# Super Bowl watchers		
# babies born		
# abortions		
$ median income		

Give 25% and 75% probability bounds for each of these quantities. You should specify the bounds so that, for an unknown quantity x, there should be a 50% chance that x is between your upper and lower bounds. Fill in all the blanks on the table. You will then be told the true values of these quantities.

Are you calibrated?

- People's 50% intervals typically are correct only 1/3 of the time
  - 90% intervals are correct only ½ of the time
  - 100% intervals are only correct about ½ of the time, too!
- But there is a foolproof way of being calibrated . . .

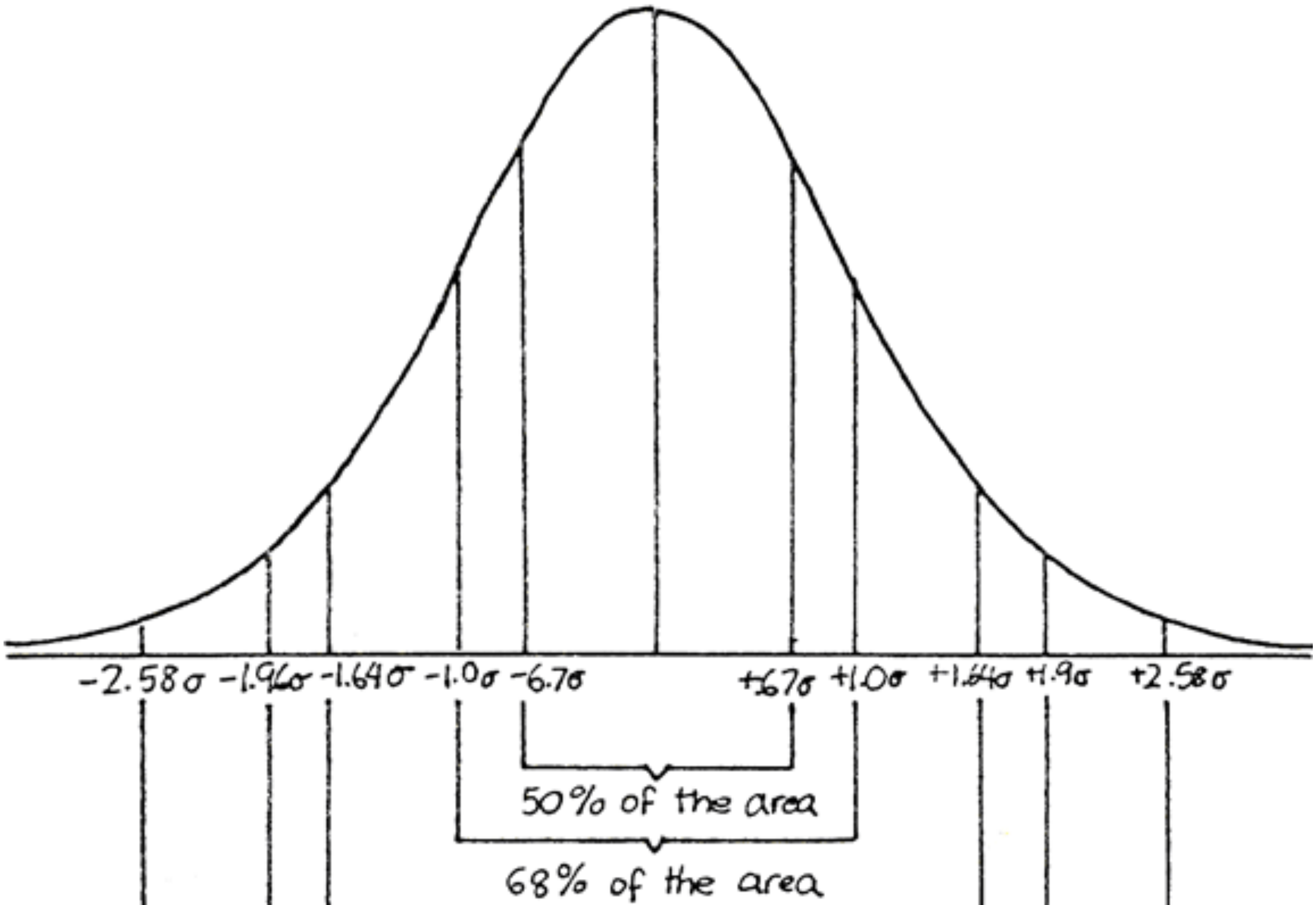
Experts are overconfident too



How many quarters are in the jar?

- We'll negotiate on a point estimate, then a 50% interval, then we'll use this to get a subjective probability distribution
- Then I want a single guess from the class
- If you guess correctly, you keep all the quarters!

Representing your uncertainty using a normal distribution



Maximizing your expected gain

- Let x be your guess
- Expected gain is approximately:

  $x * \exp[-\frac{1}{2}((x-\mu)/\sigma)^2] * \sqrt{(1/(2\pi\sigma^2))}$
- Differentiate with respect to x, set d/dx to 0
- Solve for x:
  - Optimal $x = \frac{1}{2}[\mu + \sqrt{(\mu^2+4\sigma^2)}]$
  - This is a little more than μ
- For your distribution, optimal x is . . .

What happened in this demo?

- Students learn about empirical calibration of probabilities
- Expected value = (Possible value) * probability
- Students actually get to use the normal distribution formula
- Optimization by setting derivative equal to 0
- Solution uses the quadratic formula!

- Just complicated enough . . .

Examples, demos, drills, projects

- In class: keep students awake and learning
- Identify problem areas
- Motivate students to practice

- Why did I need a “bag of tricks”?

  The Deborah Nolan story

Examples: some principles

- Relevance
  - Surveys and experiments on topics of interest (e.g., beauty and student evaluations; drinking and academic performance)
  - For probability examples: boy and girl births, not tricky dice games, poker hands, etc.
  - Straight math problems are OK (and needed) too
- Active participation of students
- Work in pairs

Demonstrations: some principles

- Clear instructions
- Working in pairs
- Debriefing afterward: connect to statistical topics

Drills: some principles

- Easy questions
- Involve all the students
- Don't make it a lecture

Not doing it

- Teachers love reading about these activities but don't actually use them!
- Why?
  - Limited class time
  - Awkwardness of trying something new, losing control

- It's not "what's covered in class" that matters, it's "what's learned in class"

How we do it

What we do

- "Covering the material"
  - Students learn by doing homeworks
  - Rely on the textbook. Students will rely on it anyway!
  - Give students tips on how to do well on exams
- Active learning in class
  - Time sharing (candy demo)
  - 1 demo and 1 drill per lecture
- Just-in-time teaching assignments

Some ups and downs

- Real and fake coin flips

00111000110010000100	01000101001100010100
00100010001000000001	11101001100011110100
00110010101100001111	01110100011000110111
11001100010101100100	10001001011011011100
10001000000011111001	01100100010010000100

- Lie detection
- Role playing

Putting it all together

- Integrating drills, hwks, exams, and lectures
- Goal: a more teacher-friendly (and student-friendly) package
- Just the good stuff—no “filler”
- Motivate students to do the hard work to learn