Teaching Statistics

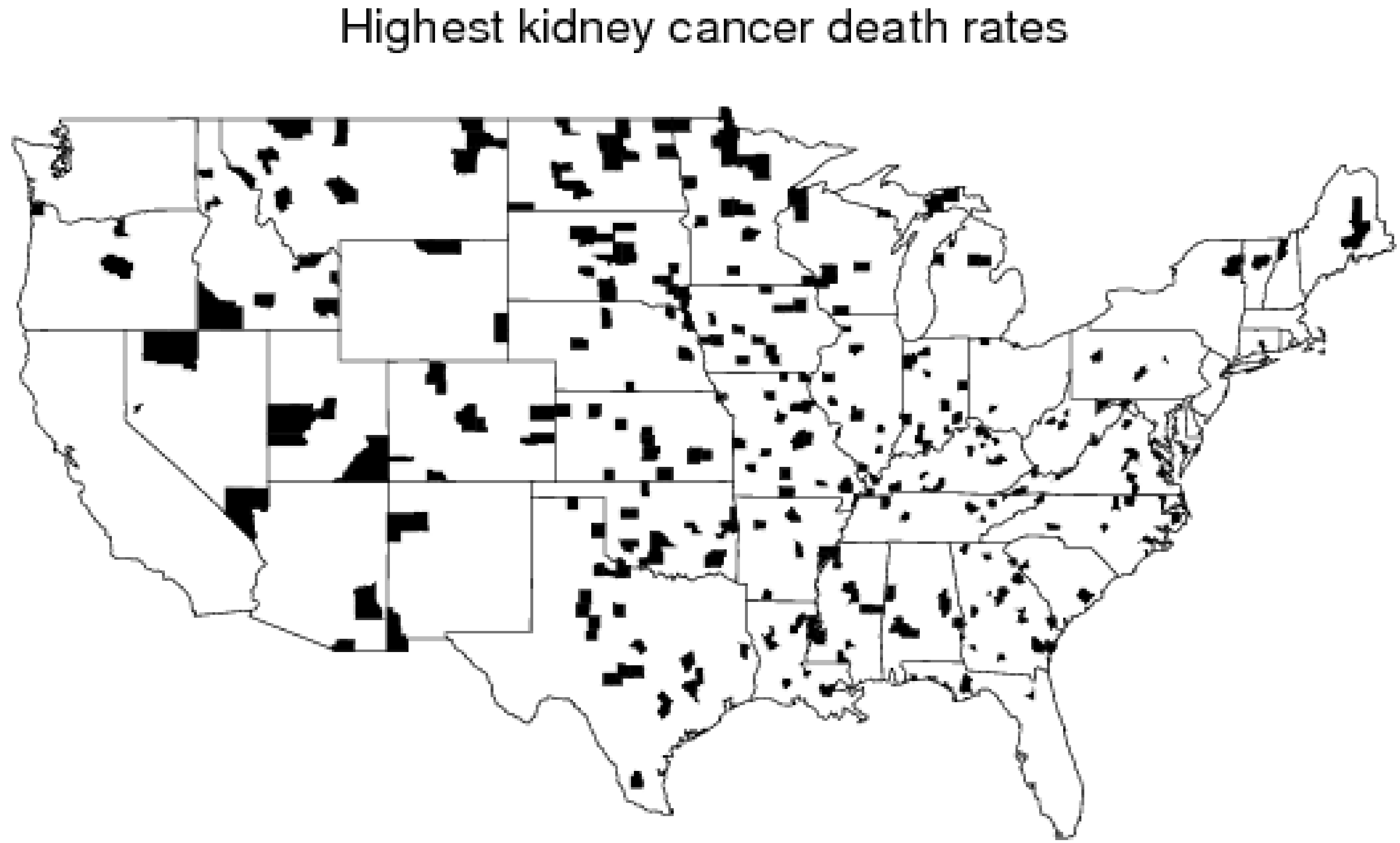
Andrew Gelman
Department of Statistics and
Department of Political Science
Columbia University, New York

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

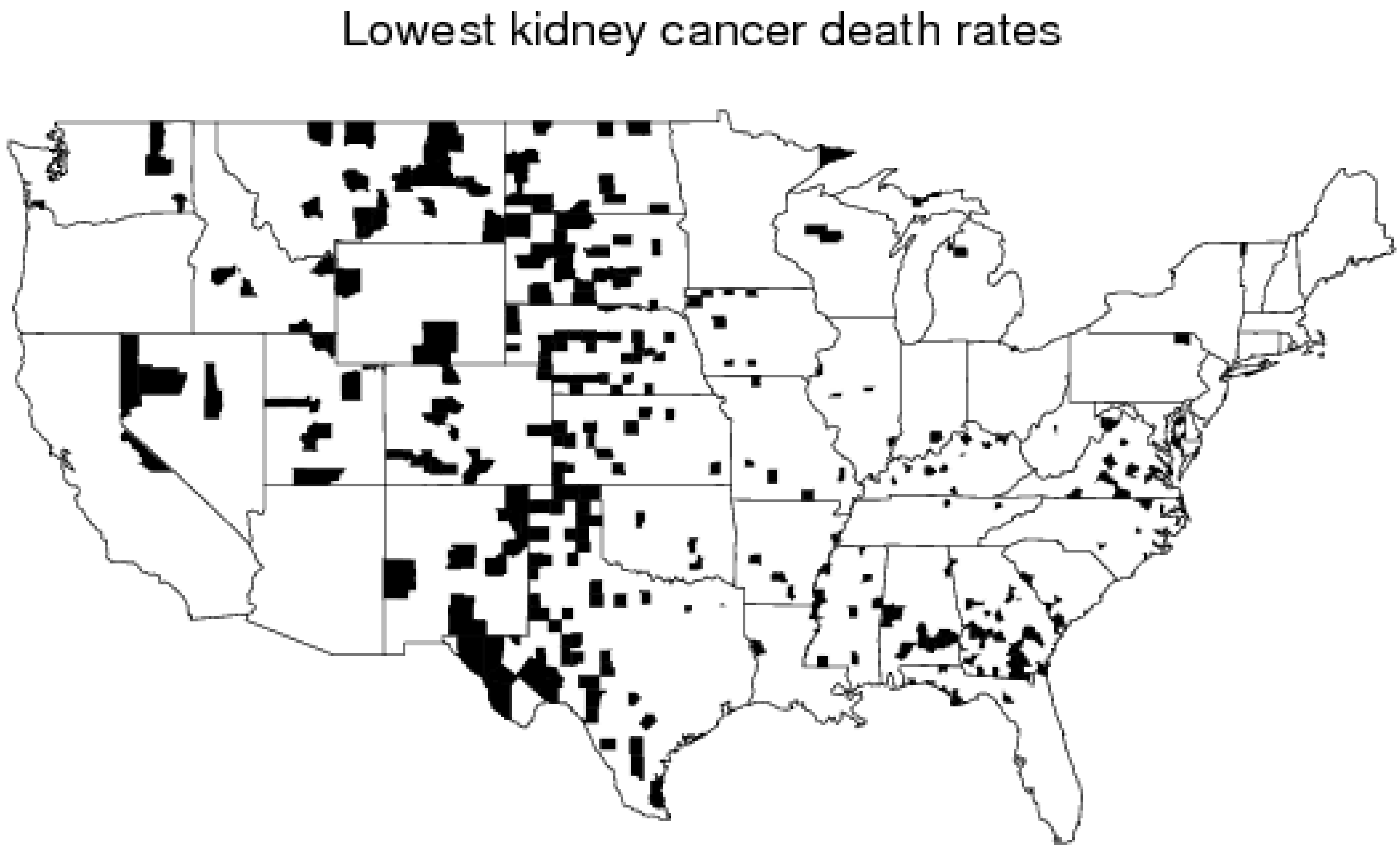
- Work in pairs
- You can choose your 5 candies using any method—systematic sampling, random sampling, whatever
- Whoever guesses closest to the true weight gets the whole bag



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



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.

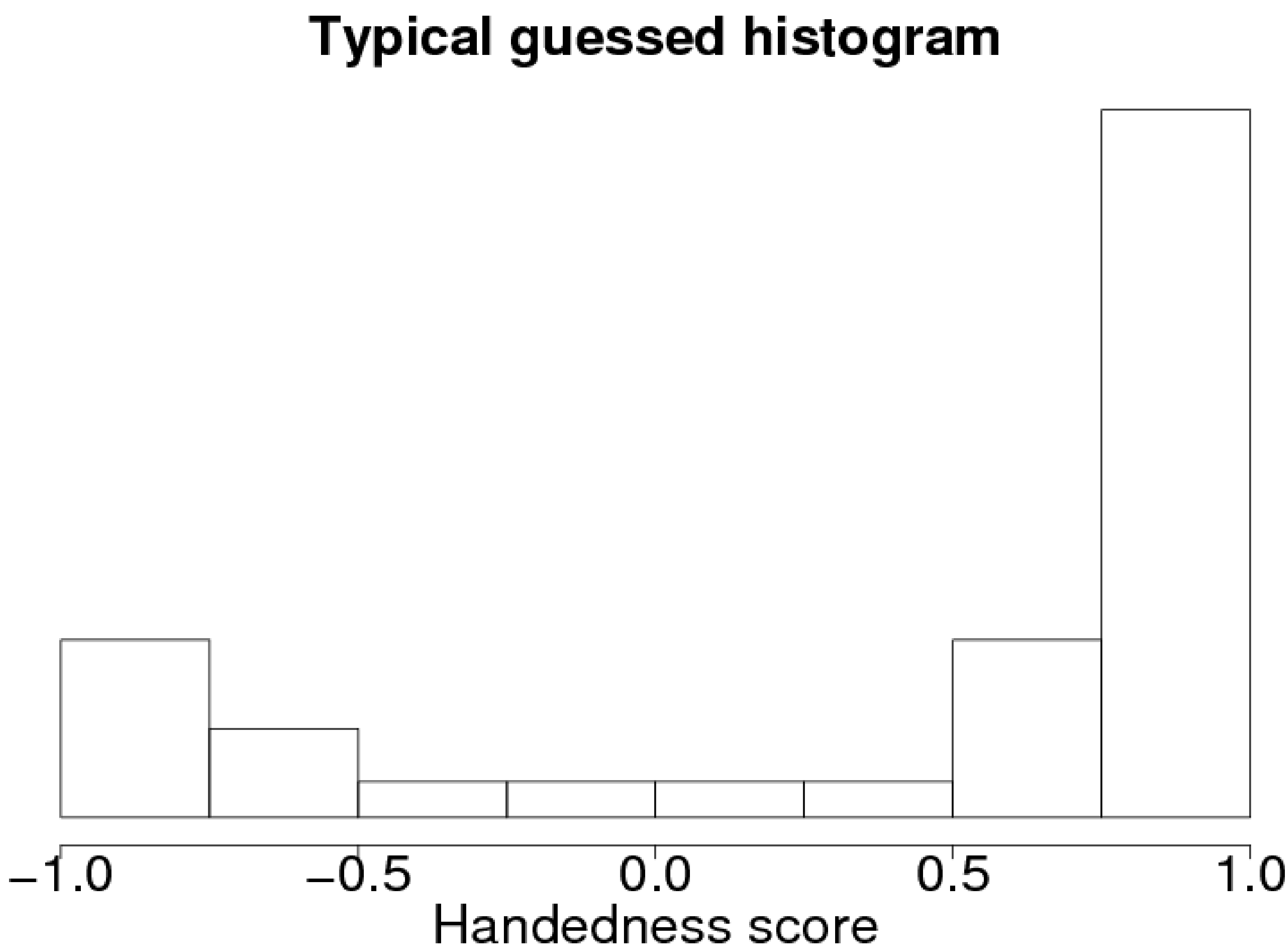
Task	Left	Right
Writing		
Drawing		
Throwing		
Scissors		
Toothbrush		
Knife (without fork)		
Spoon		
Broom (upper hand)		
Striking match (hand that holds the match)		
Opening box (hand that holds the lid)		
Total		

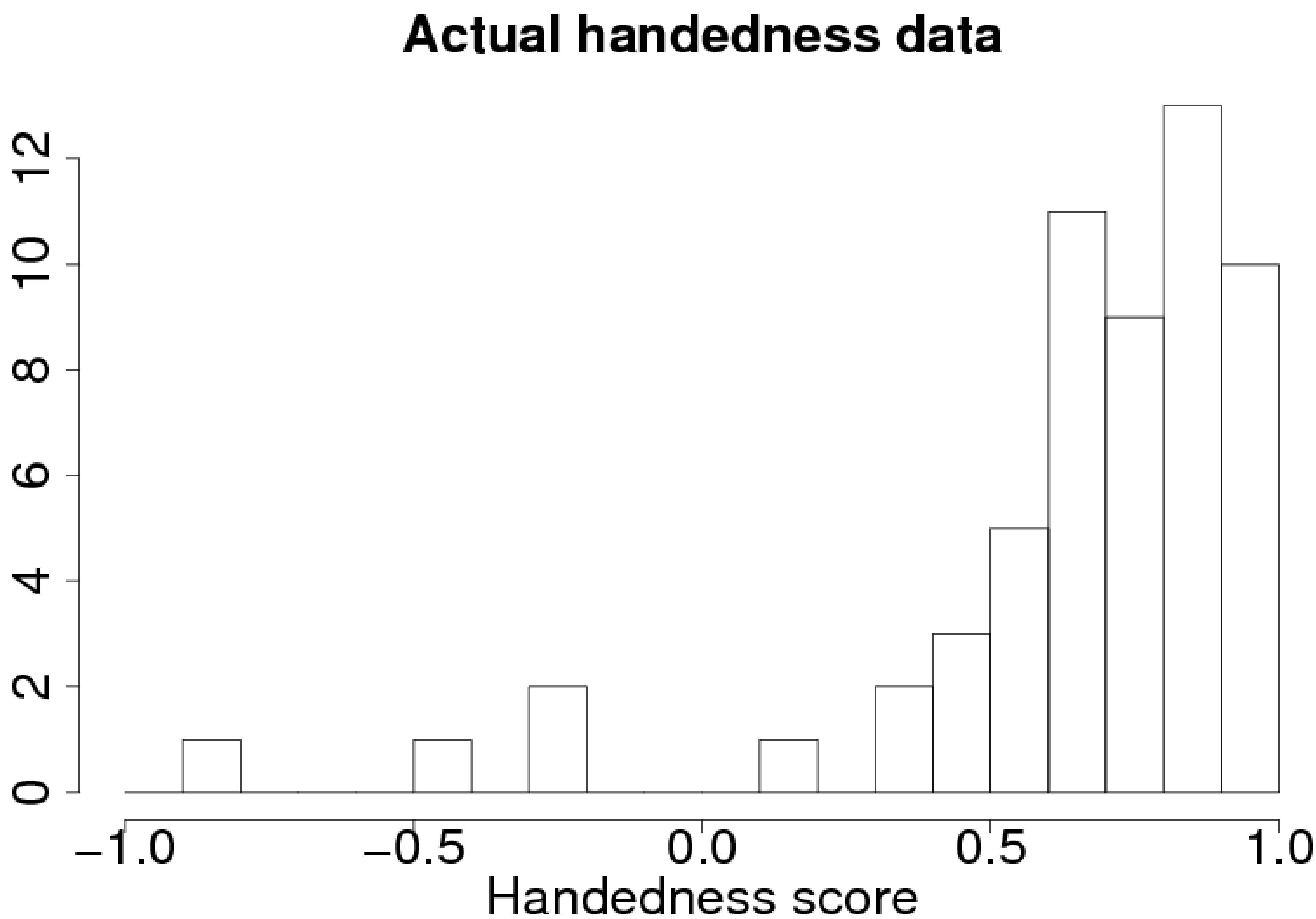
Right − Left: Right + Left: $\frac{\text{Right} - \text{Left}}{\text{Right} + \text{Left}}$:

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?

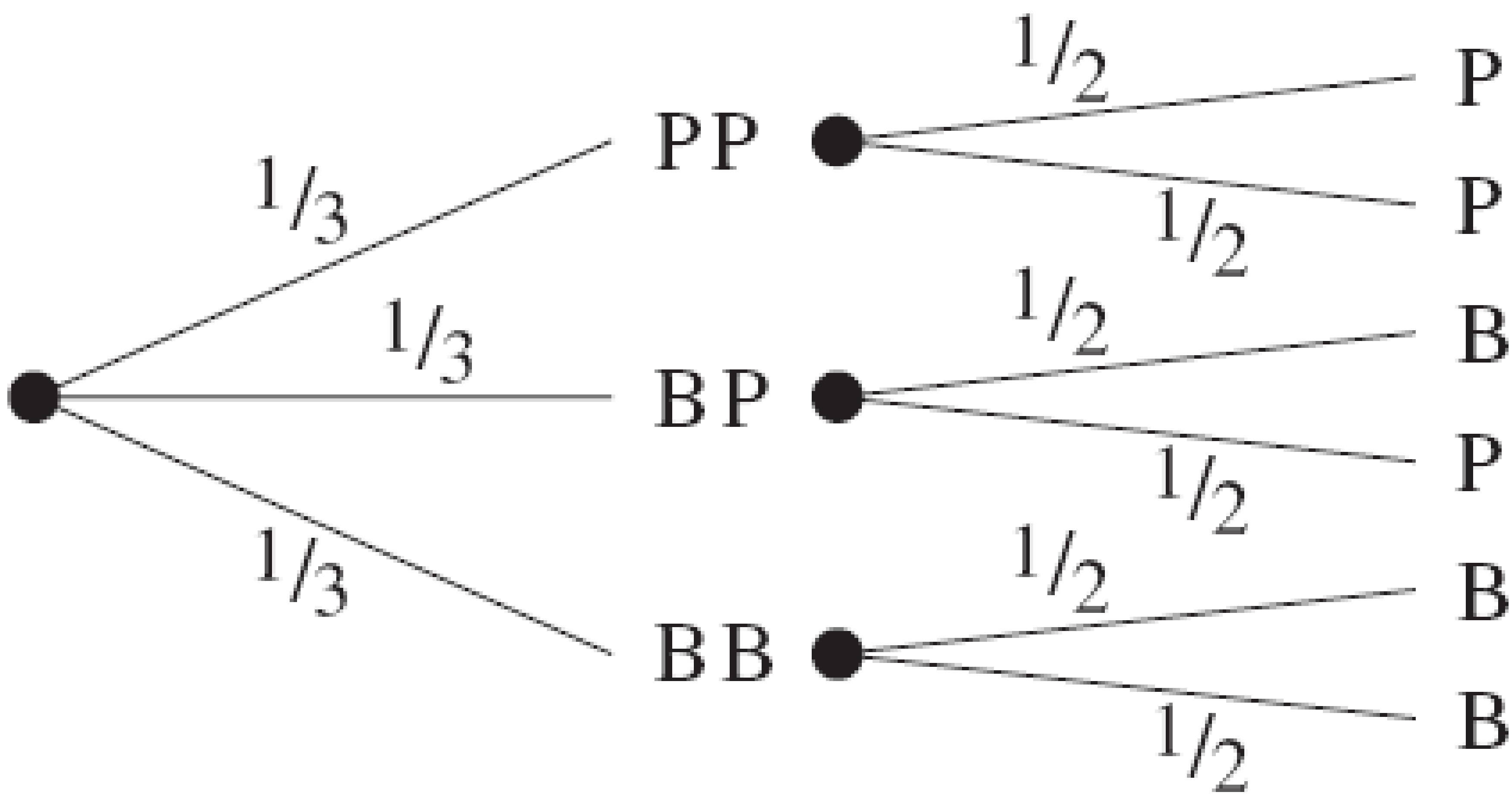
Experimenting with exam questions

- 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

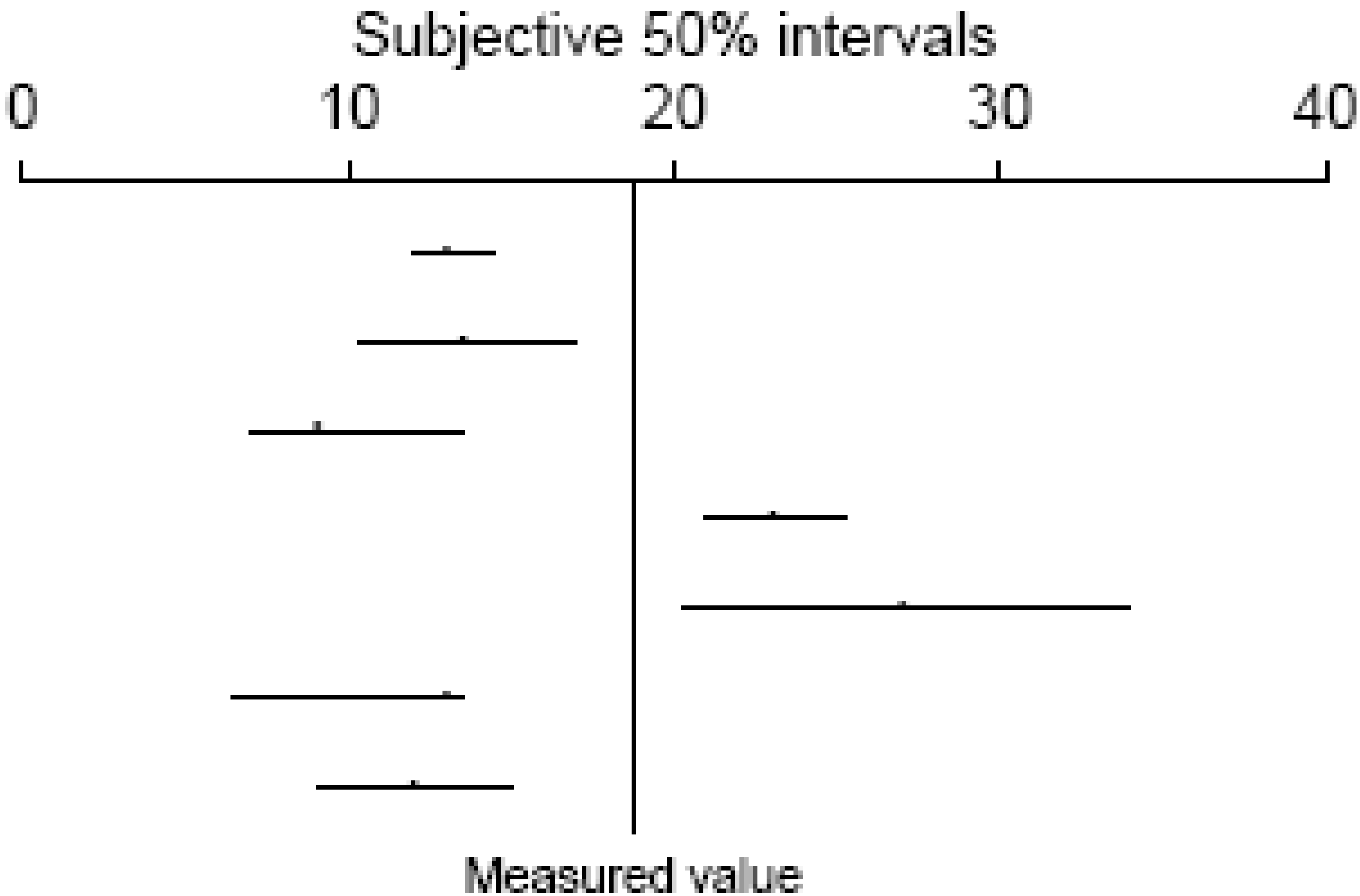
Uncertain quantity	25% lower bound	75% upper bound
% black		
# eggs		
# airline deaths		
% girl births		
% freshmen in phys sciences		
# French speakers		
# Super Bowl watchers		
# babies born		
# abortions		
$ median income		

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 . . .

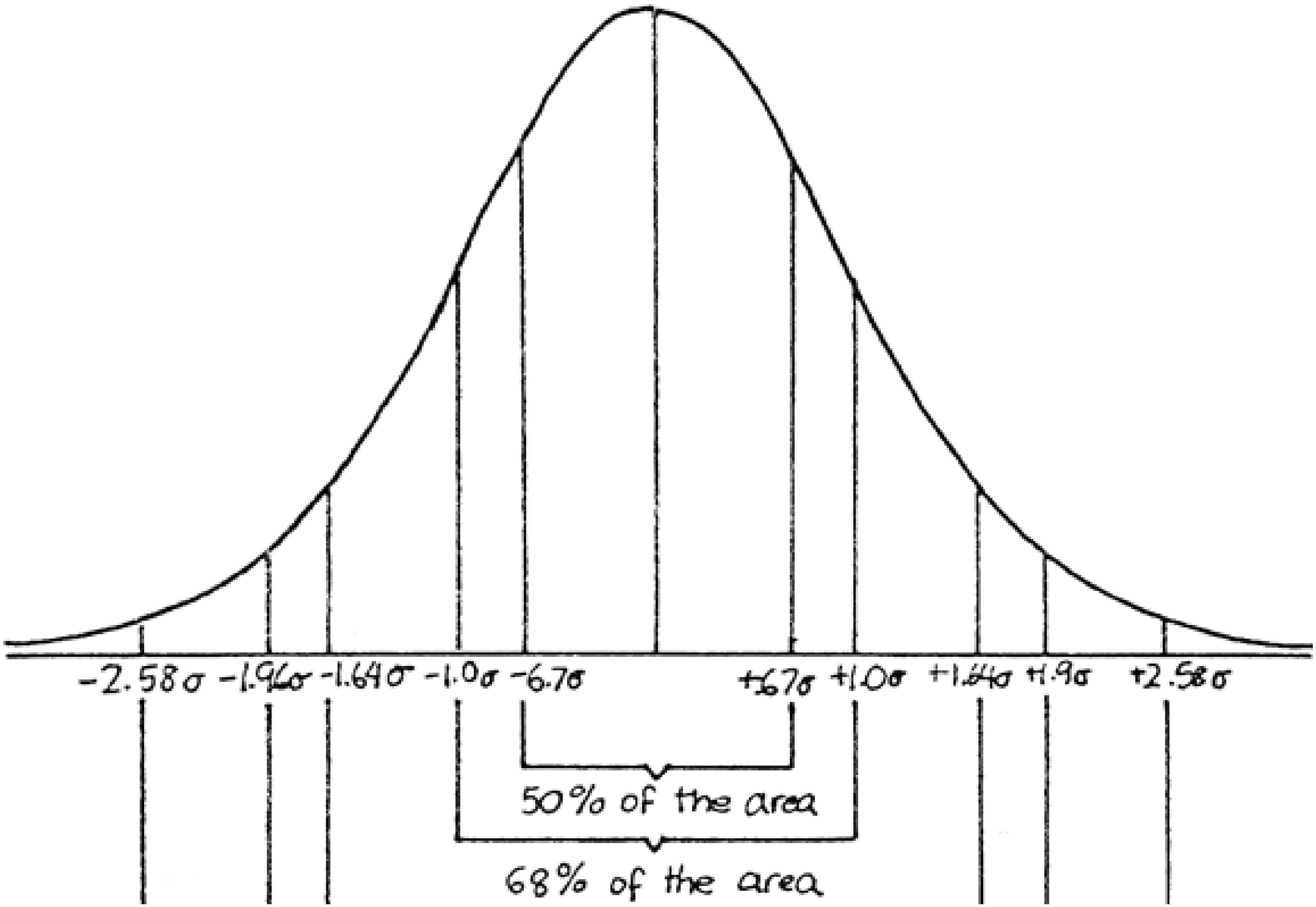
Experts are overconfident too



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

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

Projects: some principles

- Give students a good template
- Data collection or data analysis
- Can they study something interesting to them?

Not doing it

- Teachers love 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

Example: learning logarithms

- 2 topics: log and log-log
- Log transformation
  - Amoebas
  - World population
- Log-log transformation
  - Squares and cubes
  - Metabolic rates

Log transformation

- You have an amoeba that takes 1 hour to divide, then 2 amoebas divide in one more hour, . . .
- What is y(x), the formula of #amoebas as a function of time?
- Now suppose amoebas take 3 hours to divide; what is y(x)?

Log transformation

- You have an amoeba that takes 1 hour to divide, then 2 amoebas divide in one more hour, . . .
- What is y(x), the formula of #amoebas as a function of time?

  $y = 2^x$

  $\log y = (\log 2)\, x = 0.30\, x$

- Now suppose amoebas take 3 hours to divide . . .

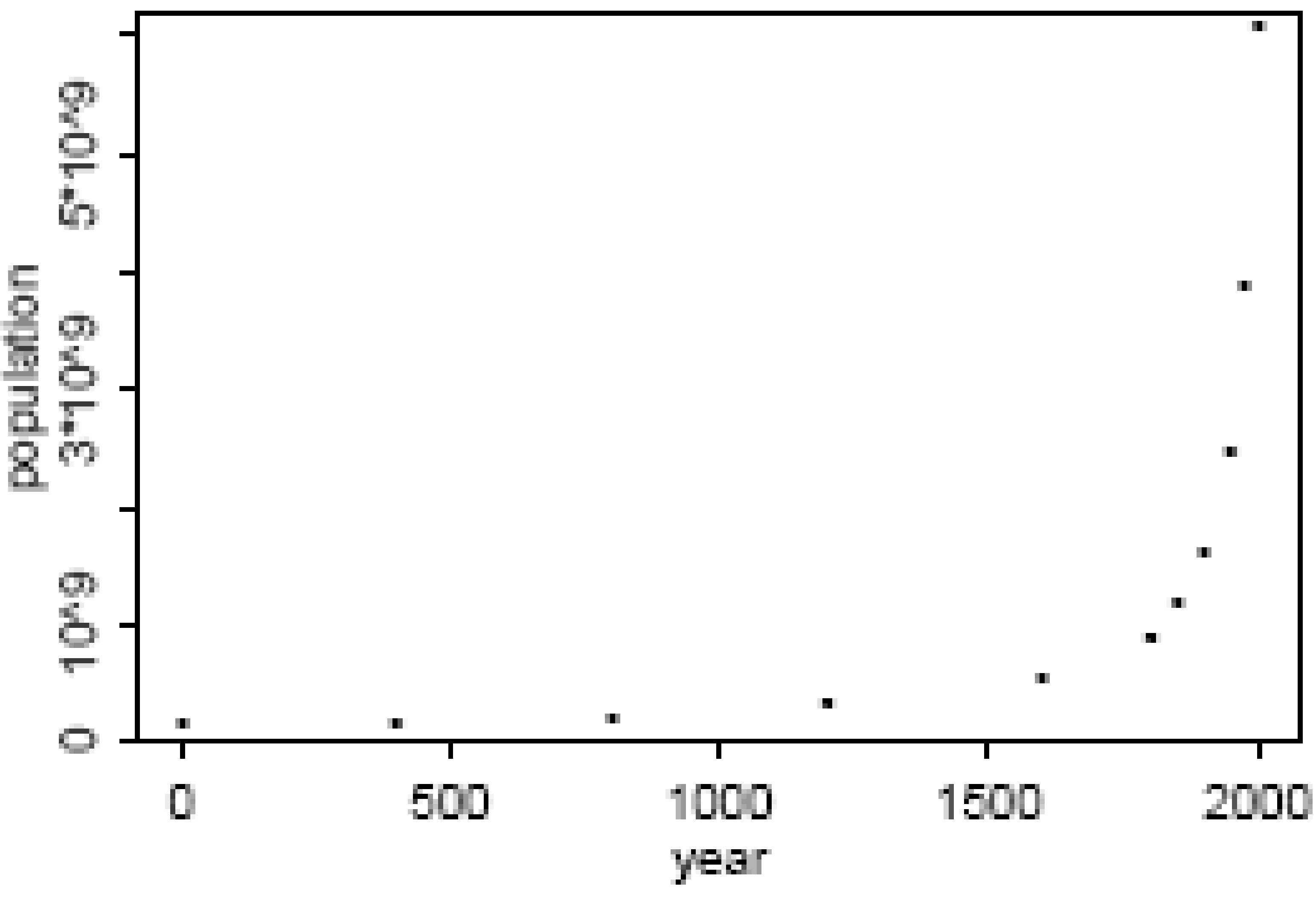
  $y = 2^{(x/3)} = (2^{(1/3)})^x = 1.26^x$

  $\log y = (\log 1.26)\, x = 0.10\, x$

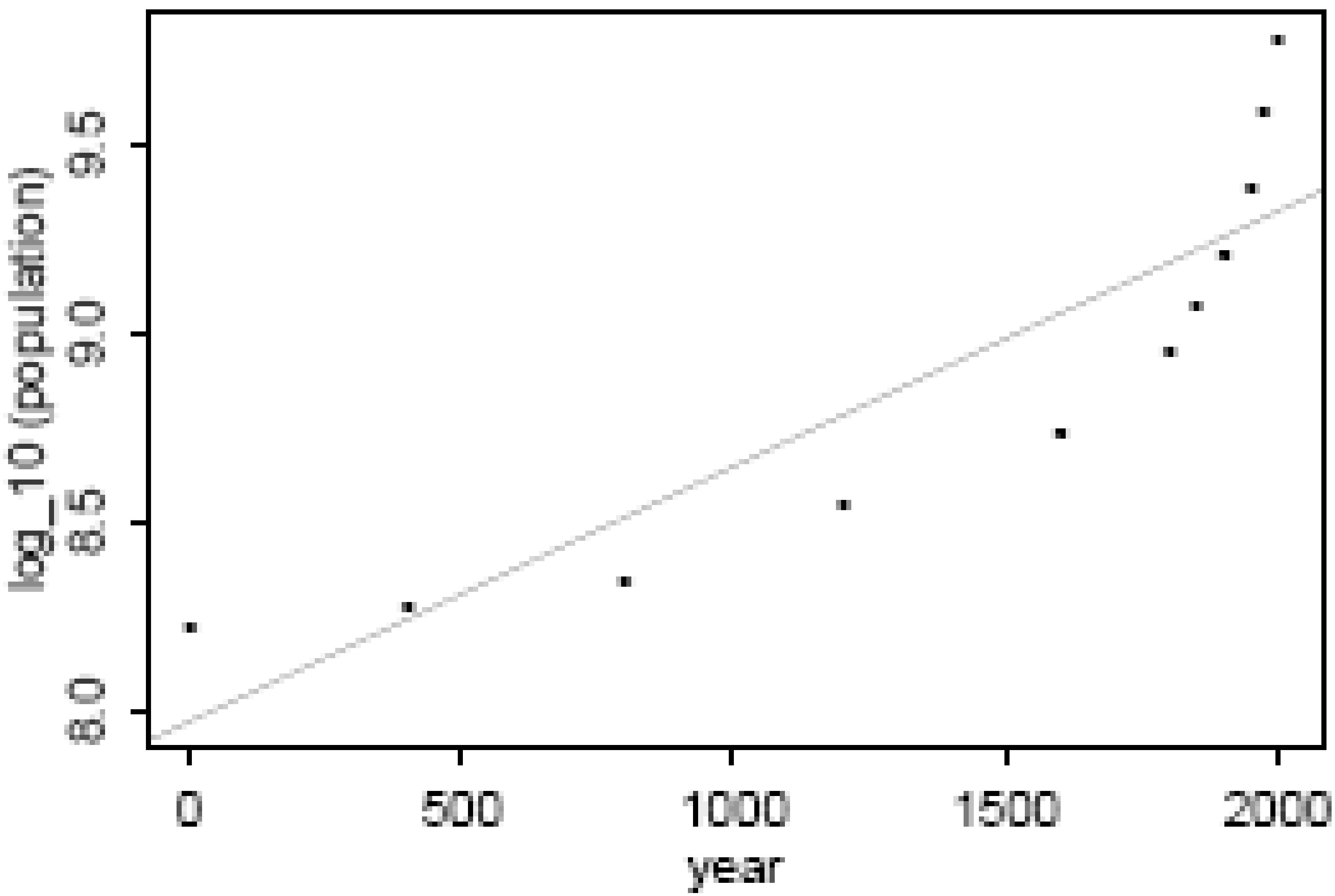
World population

Year	Population	$\log_{10}$ (population)	Residual	10^{residual}
1	170 million	8.230	.258	1.81
400	190	8.279	.037	1.09
800	220	8.342	−.171	.68
1200	360	8.556	−.227	.59
1600	545	8.736	−.318	.48
1800	900	8.954	−.236	.58
1850	1200	9.079	−.145	.72
1900	1625	9.200	−.047	.90
1950	2500	9.398	.107	1.28
1975	3900	9.591	.283	1.92
2000	6080	9.784	.459	2.88

World population, graphed



World population on log scale



Log-log transformation

- How does the area of a square relate to its circumference?
- If the side of a square has length L, then area is L^2 and circumference is 4L . . .

  (express area in terms of circumference)
- Now consider a cube: express the surface area as a function of volume

Log-log transformation

- How does the area of a square relate to its circumference?
- If the side of a square has length L, then area is L^2 and circumference is 4L; thus

  Area = (circumference/4)^2

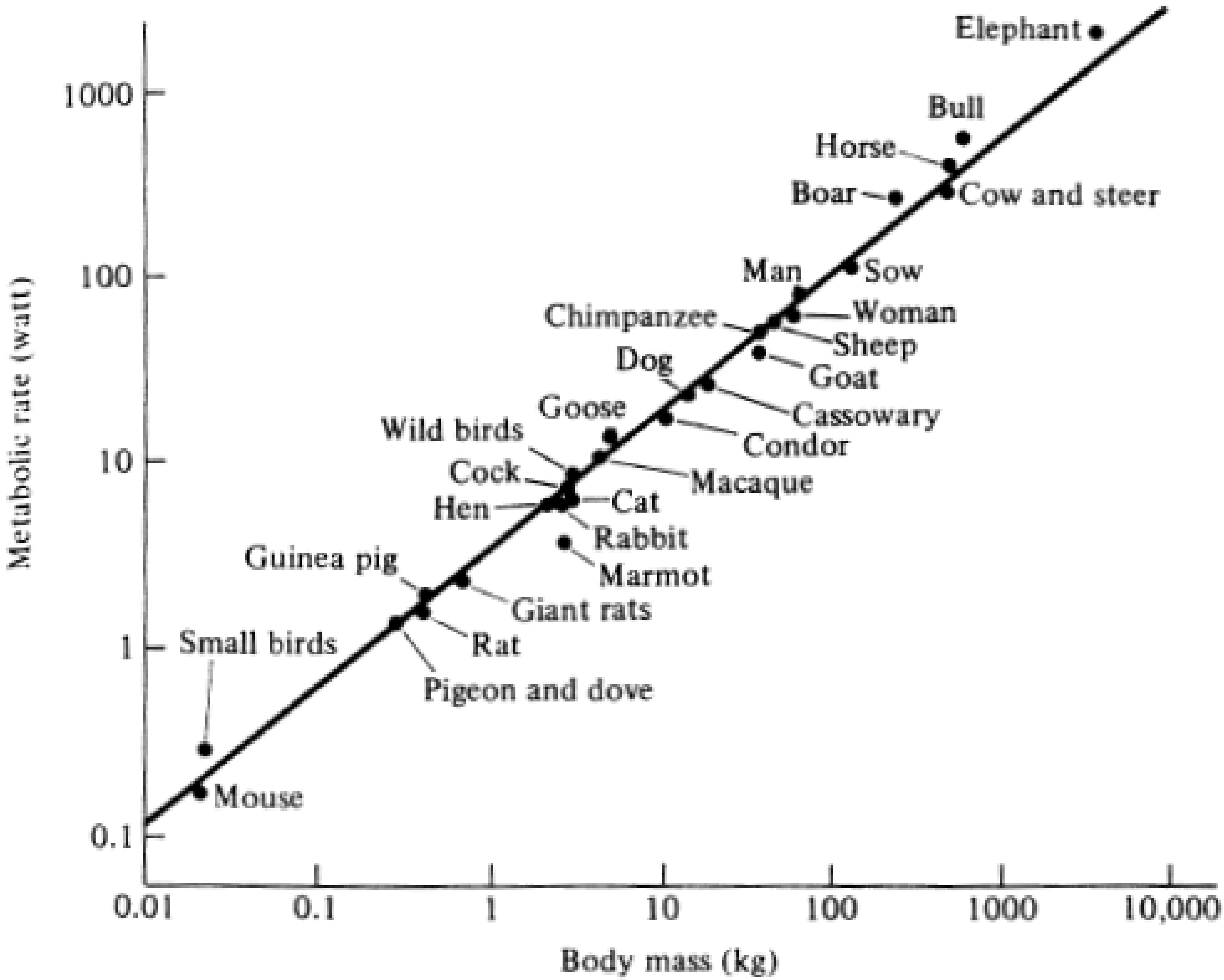
  log (area) = 2(log(circumference) – log(4))

  = –1.20 + 2 log(circumference)
- Surface area and volume of a cube:

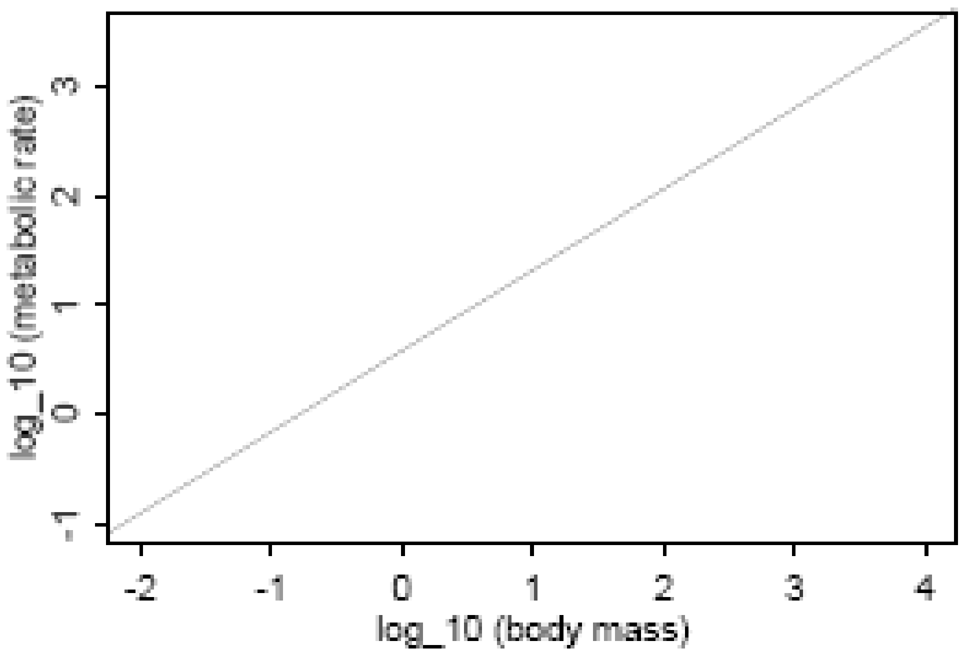
  Surface area = 6 (volume)^(2/3)

  log (surface area) = log(6) + (2/3) log(volume)

Metabolic rate and body mass



Reading the graph of metabolic rate and body mass



Line has slope 0.74 and goes through the point (3, 2.8):

$y - 2.8 = 0.74\ (x - 3)$

$y = 0.58 + 0.74\ x$

log(metabolic rate) = 0.58 + 0.74 * log (body mass)

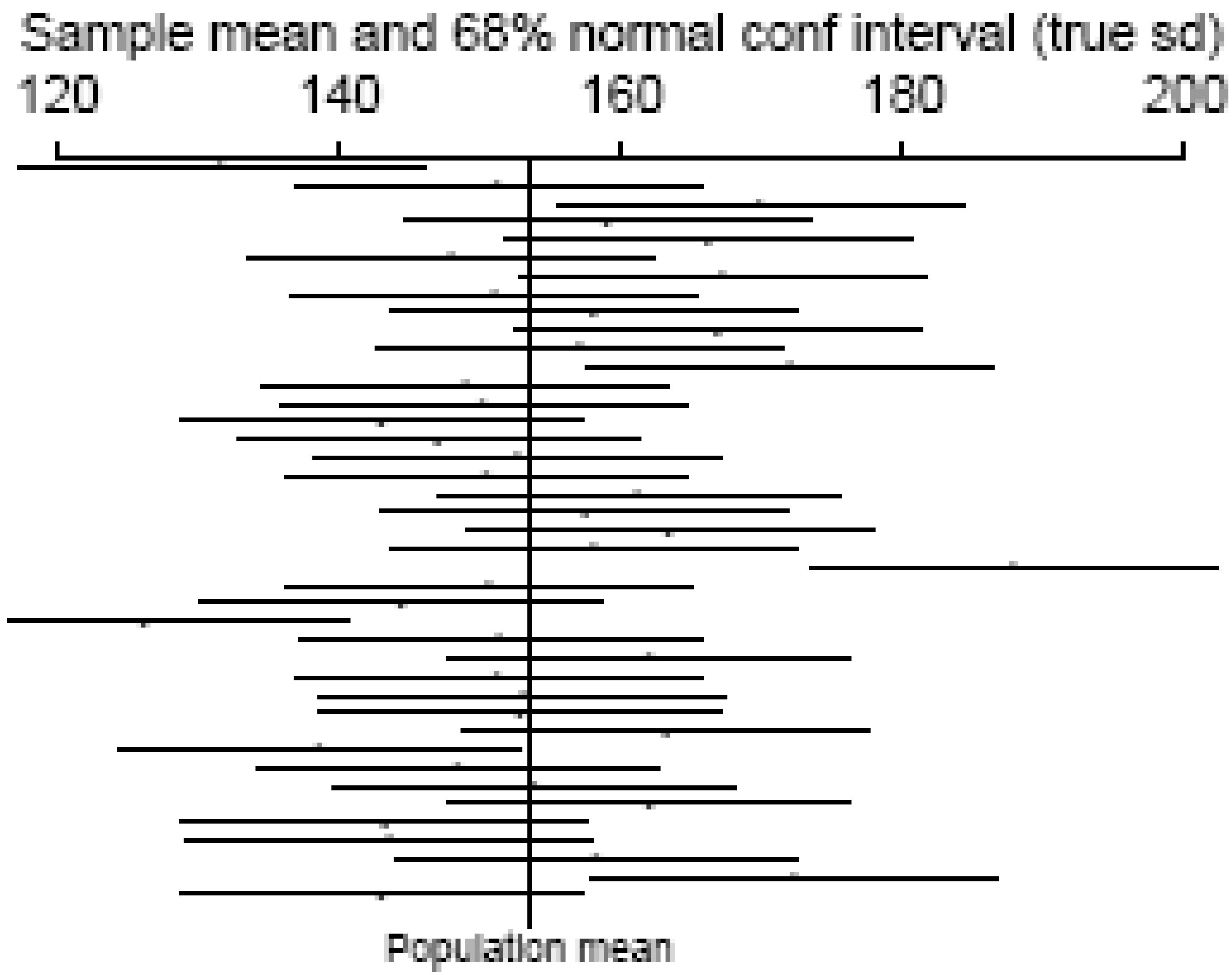
met.rate = 3.8 (mass)^0.74

Consider an elephant, 100 men, or 10000 rats . . .

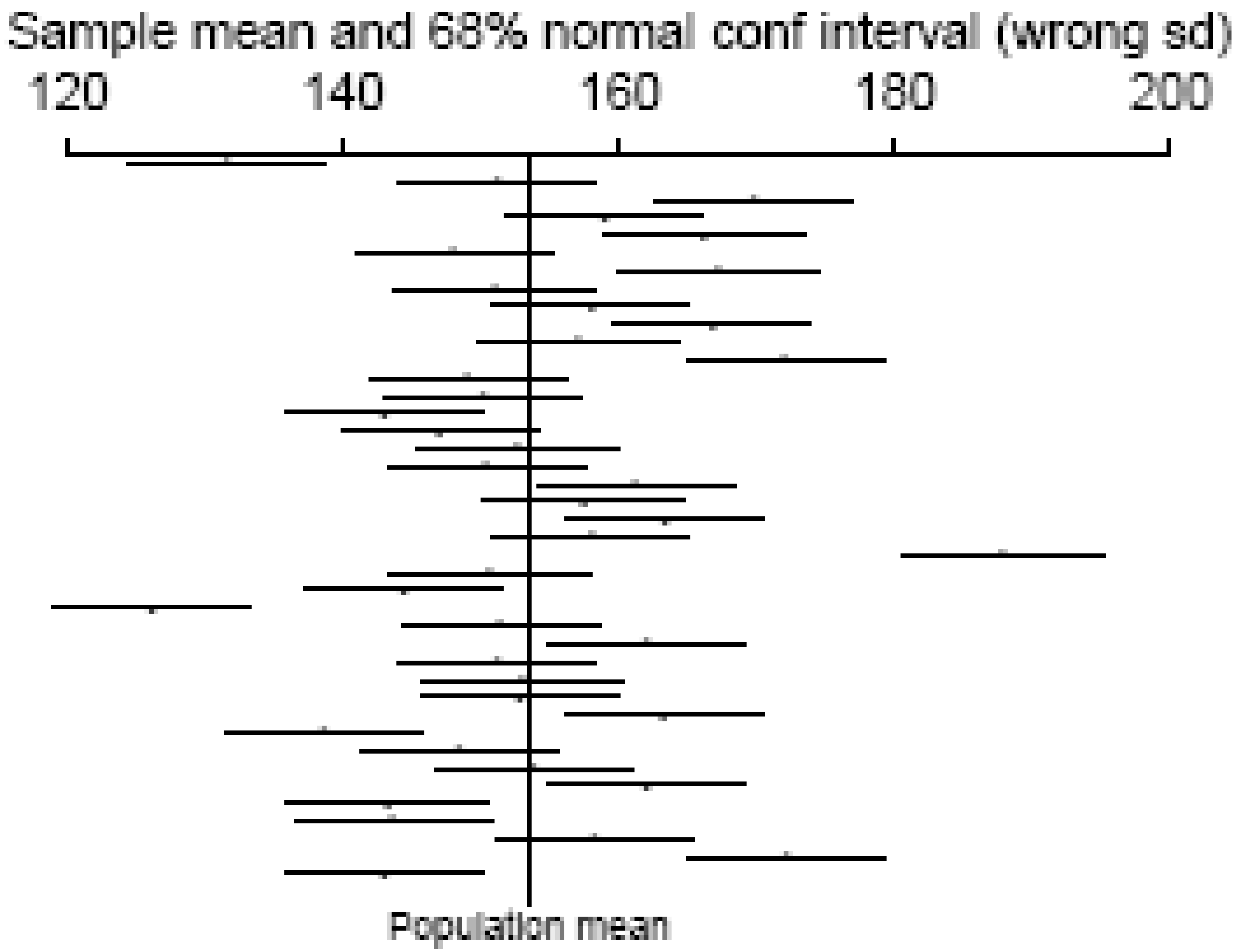
Example: coverage of conf intervals

- Each student writes his or her weight on a slip of paper
- Put the slips in a bag
- Assistant quickly enters the numbers on computer
- Pass bag around the room, each pair of students takes a sample of size 5
  - 68% conf interval for mean: estimate $\pm$ sd/$\sqrt{5}$
- Display all the intervals, compare to true mean

Example: coverage of conf intervals



Same demo but we lie about the sd



Struggles: demos and games

- Scatterplot charades
  - My example
  - Students examples?
- Phone book sampling
  - Random numbers via dice
  - Sampling: missing data and duplicate listings
  - Phone numbers and addresses

Phone book sampling

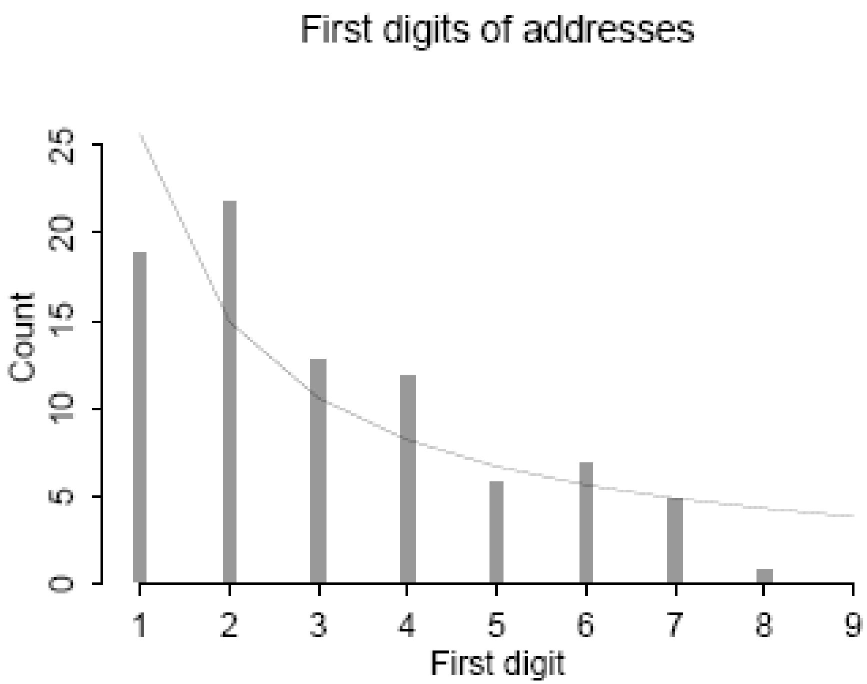
KATOPIS Theodore	120 E 82	212 249-3047
KATOVITZ Michael	299 W 12	212 929-9511
KATOWSKY Marc	215 E 95	212 706-2855
KATRAGADDA Sireesha	31 E 31	212 532-6457
KATRANCI Elif	155 E 99	212 722-1951
KATRI Edmond	160 E 48	212 588-0118
KATRITSIS A		212 741-0174
KATROV Marat P	747 10 Av	212 757-4845
KATS Amir	531 W 48	212 333-5811
Ester	15 Willett	212 477-2490
Guyora	230 W 82	212 362-5351
I		212 588-1244
Inna	1277 3 Av	212 288-7739
Michael	345 E 93	212 987-2902
Victor	75 West St	212 385-1686
KATSAMAKIS Basil	315 E 69	212 628-9512
Basil	530 E 72	212 628-0312
KATSANOS Andrew	321 E 71	212 717-9393
Christina	417 W 47	212 459-2304
KATTULA Jennafer	409 E 69	212 327-2845
KATUN Mosammat	316 W 95	212 666-4817
KATUS B	210 W 89	212 362-9715
KATUSAK F J	176 E 77	212 737-8955
KATVAN Moshe	40 W 17	212 627-2169
Moshe	40 W 17	212 627-4362
Moshe	40 W 17	212 627-5035
Moshe & Rivka	117 W 17	212 627-5034
KATWAROO Dianna	434 W 163	212 568-0636
Errol	434 W 163	212 568-3629
KATYAL Monica	617 W 115	212 222-3669
KATYANG Keo	104 W 96	212 749-8386
KATZ A		212 721-3504
A		212 725-6758
A	268 E Bway	212 982-8619
A	737 Park Av	212 517-8897
A	2 5 Av	212 533-9692
A	148 10 Av	212 366-6487
A	315 E 86	212 831-7554
A D	433 W 21	212 255-1769

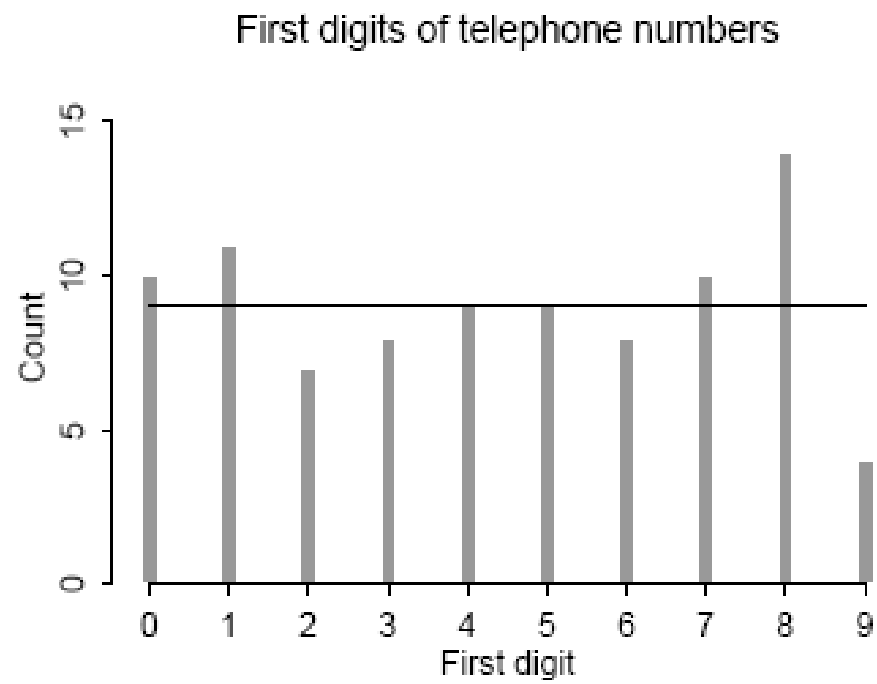
Doing the sampling

A sample from the phone book

	Page	Column	Entry	Address #	Telephone #
1	520	5	100	15 W 53 St	586-7149
2	519	2	116	240 W 116 St	663-1076
3	519	4	087	710 West End Ave	749-2245
4	520	2	081	511 E 20 St	533-0614
5	519	4	115	2 Horatio St	206-7914
6	519	3	124	256 ...	304-2769
7	519	2	110	350 ...	308-4620
8	520	1	107	129 ...	xxx-2xxx
9	520	5	126	315 ...	xxx-2xxx
10	520	2	040	104 ...	xxx-1xxx

Sampling distributions





Struggles: demos and games

- Real and fake coin flips

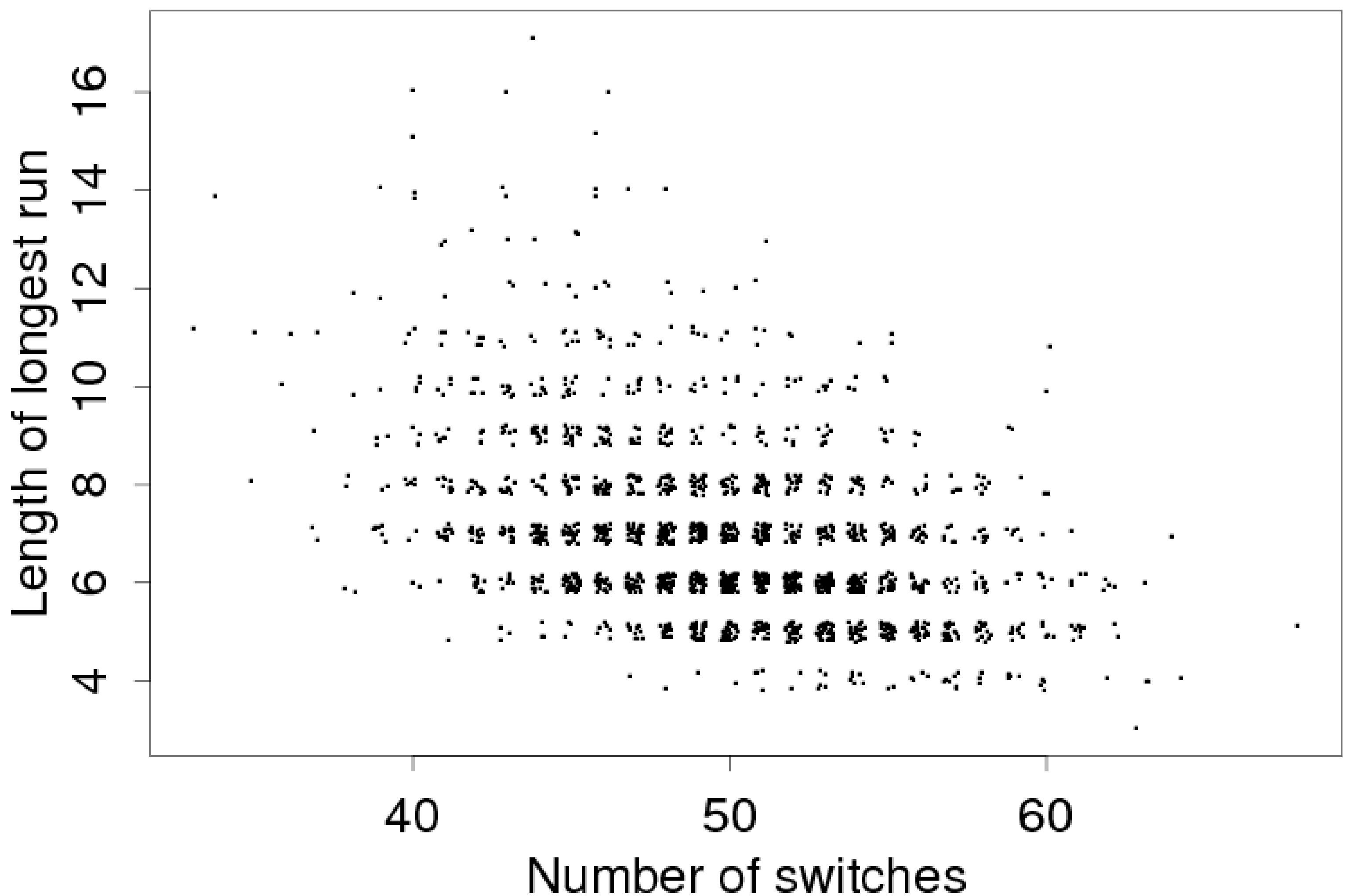
00111000110010000100	01000101001100010100
00100010001000000001	11101001100011110100
00110010101100001111	01110100011000110111
11001100010101100100	10001001011011011100
10001000000011111001	01100100010010000100

- Lie detection
- Role playing

Bridging from demo to course material

- Coin flipping demo
- Look at #switches
  - 99 chances to switch
  - For random flips, Pr(switch)=0.5 each time
  - Binomial (99, 0.5)
  - Graph distribution, show p-value
- This doesn’t work well
  - The math is tricky, distracts from main point
  - Also, it’s a low-power test

This didn't work either!



Struggles: group projects

- Lots of ideas . . .
- . . . But student projects are usually disappointing
- Struggles
  - The topic
  - Data collection
  - Data analysis
  - Working together

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

From your perspective

- Sharing teaching tips
- Where to put your teaching effort?
- Connections to empirical research?

Larger questions about stat classes

- Intro undergraduate courses
- Basic graduate courses
- Advanced courses

The sampling distribution of the sample mean

- Traditional central point of intro course
  - Culmination of expectations and variances
  - Applied to one-sample estimates, two-sample comparisons, regression, Anova, etc.
- Problem: students don't understand it
  - They can't derive it
  - They can't apply it
- Solution: I don't teach it!

Challenges

- Integrating homeworks, drills, exams
- Monitoring progress of individual students
- Statistics is a sloppy subject
- Motivating non-quantitative students