# Review
# 18.05 Spring 2018

Here are some board problems to finish the semester...

# Board question: mileage

Each time it is turned off, your car reports how far you have travelled and how much gasoline you used. Here are the reports (distance,gas) from last week:

$$(0.8, 0.06),\ (1.1, 0.08),\ (0.8, 0.05), (36.2, 0.74),\ (1.1, 0.07)$$

Do a linear regression to estimate mileage.
**Hint:** R says that the line best fitting these points is

$$\text{gallons} = 0.019 * \text{distance} + 0.047$$

**Better Hint:** R says that the line with intercept zero best fitting is

$$\text{gallons} = 0.021 * \text{distance}$$

# Board question: make it fit

Bivariate data:

$$(1,3),\ (2,1),\ (4,4)$$

**1.** Do linear regression to find the best fitting parabola.

**2.** Do linear regression to find the best fitting cubic.

# Board Question

(a) Count the number of ways to get exactly 2 heads in 10 flips of a coin.

(b) For a fair coin, what is the probability of exactly 2 heads in 10 flips?

(c) If you flip a coin 10 times and get 2 heads, should you reject the null hypothesis that the coin is fair with 95% confidence?

## Board question: gourmet chocolate

The Atlas Gourmet Chocolate Company (gcc) manufactures 10 million chocolate bars each year. Before a bar is sold as a gcc bar, it is subjected to eight independent quality control tests. Three-fourths of the bars pass any one test, but passing all eight is difficult.

**1.** How many gcc bars should Atlas *expect* each year?

**2.** As production manager for the factory, would you advise Atlas to *count on* producing a million gcc bars?

**3.** In a recent year Atlas produced just 998,000 gcc bars. Is this evidence of possible sabotage in the factory?

# Board Question: Find the pmf

$X = \#$ of successes before the *second* failure of a sequence of independent Bernoulli($p$) trials.

Describe the pmf of $X$.

*Hint: this requires some counting.*

# Board question

I've noticed that taxis drive past 77 Mass. Ave. on the average of once every 10 minutes.

Suppose time spent waiting for a taxi is modeled by an exponential random variable

$$X \sim \text{Exponential}(1/10); \qquad f(x) = \frac{1}{10}\mathrm{e}^{-x/10}$$

**(a)** Sketch the pdf of this distribution

**(b)** Shade the region which represents the probability of waiting between 3 and 7 minutes

**(c)** Compute the probability of waiting between between 3 and 7 minutes for a taxi

**(d)** Compute and sketch the cdf.