Markov's inequality

Ben Langmead

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Department of Computer Science

Please sign guestbook (www.langmead-lab.org/teaching-materials) to tell me briefly how you are using the slides. For original Keynote files, email me (ben.langmead@gmail.com).

Averages & concentration

When more than 90 percent of faculty members rate themselves as above-average teachers, and two-thirds rate themselves among the top quarter, the outlook for much improvement in teaching seems less than promising.

Cross, K. Patricia. "Not can, but will college teaching be improved?."
New Directions for Higher Education 1977.17 (1977): 1-15.

🤔🤔🤔

Definitely not possible; immediate from meaning of "top quarter"

Averages & concentration

Table 1
Distribution of percent of estimates over degree of safe and skillful driving in relation to other drivers. Higher percentiles represent less risky and more skillful driving.

	Estimated position in sample (percentiles)									
	0–10	11–20	21–30	31–40	41–50	51–60	61–70	71–80	81–90	91–100
Safety										
US sample	2.5	0.0	5.0	0.0	5.0	2.5	2.5	22.5	37.5	22.5
Swedish sample	0.0	5.7	0.0	14.3	2.9	11.4	14.3	28.6	17.1	5.7
Skill										
US sample	0.0	2.4	2.4	2.4	0.0	12.2	22.0	12.2	26.8	19.5
Swedish sample	2.2	6.7	2.2	4.4	15.5	17.7	11.1	24.4	13.3	2.2

>75%

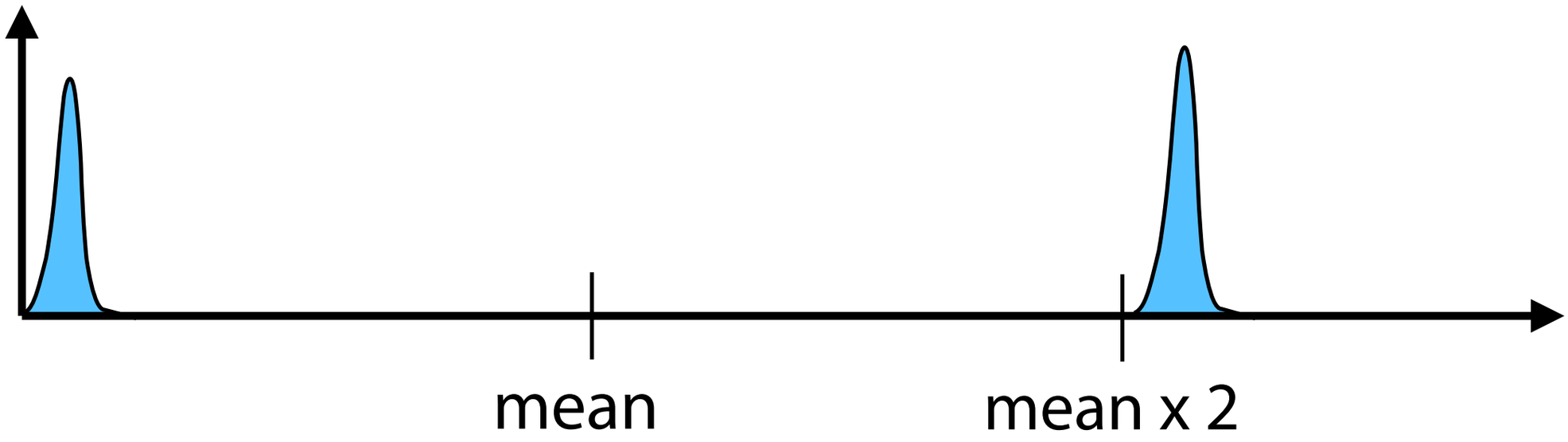
>68%

🤔🤔🤔🤔🤔🤔🤔🤔🤔🤔🤔🤔

Svenson, Ola. "Are we all less risky and more skillful than our fellow drivers?." *Acta psychologica* 47.2 (1981): 143-148.

Markov inequality

Is it possible for >50% of the data to be *more than twice* the mean?



Not possible; violate definition of mean

Not possible for >33.3% to be over 3x mean...

Or for >25% to be > 4x mean...

Markov inequality

$$\text{For } a > 0\text{:}\quad \Pr(X \geq a) \leq \frac{\mathbf{E}[X]}{a}$$

Bounds the probability a ***non-negative*** r.v. is ***much more extreme than its expected value***

"If the mean weight of 10 cars is 2 tons, none can weigh more than ... 20 tons."

Why doesn't it hold when r.v. is allowed to be negative?

Extreme positive values are are easily balanced / nullified by extreme negative ones

Markov inequality

$$\text{For } a > 0\text{:}\quad \Pr(X \geq a) \leq \frac{\mathbf{E}[X]}{a}$$

Strengths:

Only assumption is that r.v. is non-negative

Only have to know expected value (not, say, variance)

Weaknesses:

Not always a tight, useful bound

"If the mean age of 10 kindergartners is 5.5 years,
none can be more than 55 years old"

(Well ... yeah)