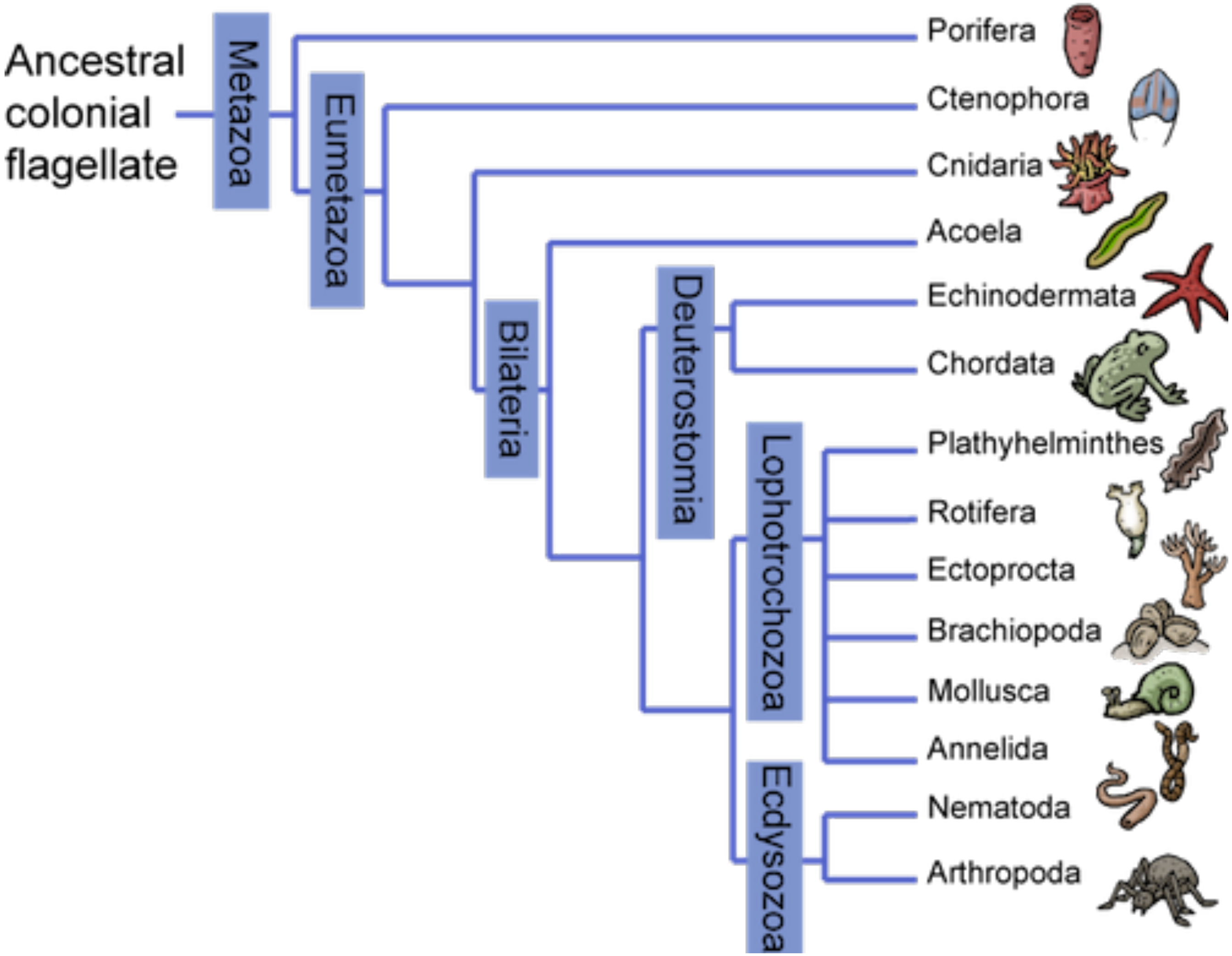
Probability

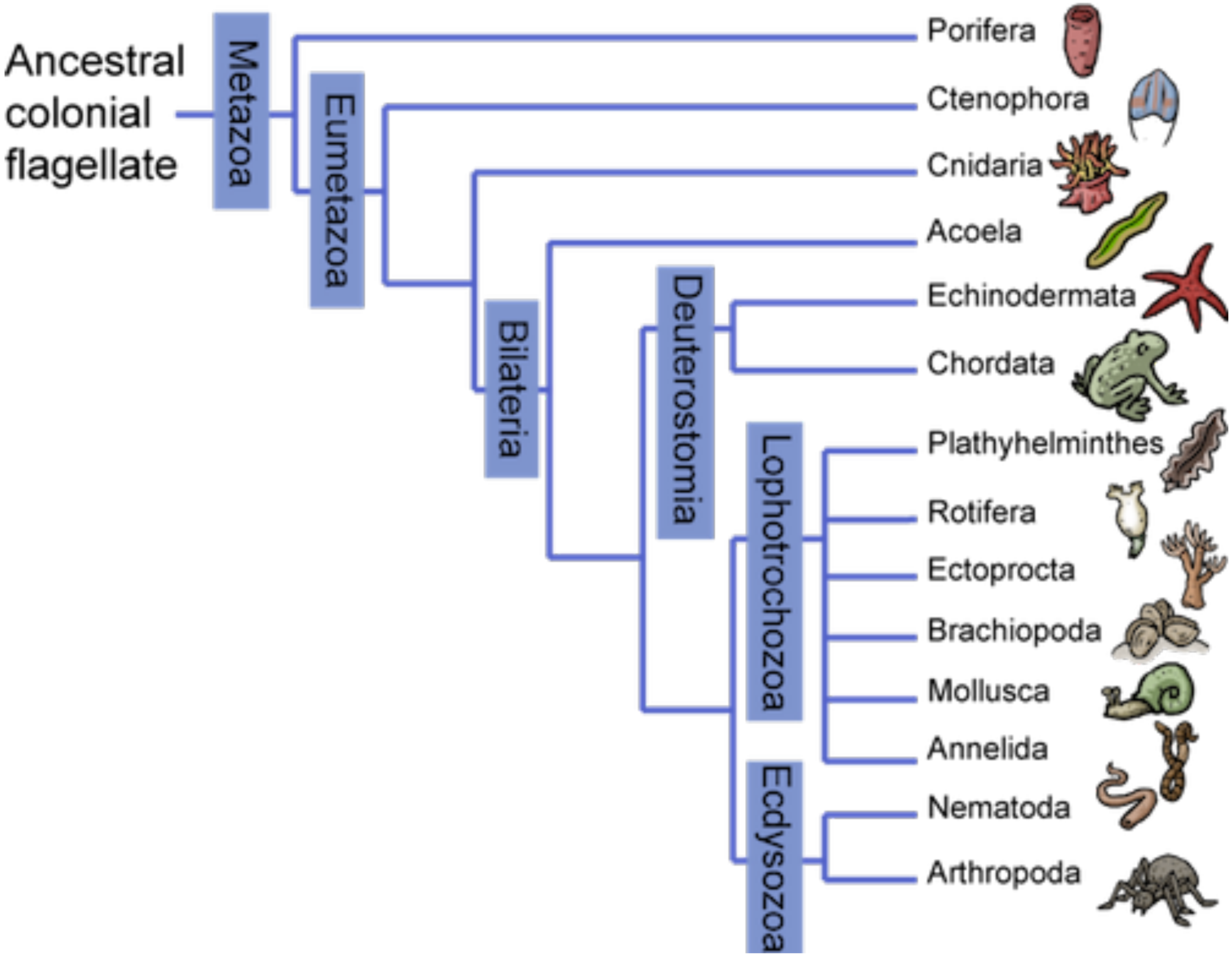
Counting Review

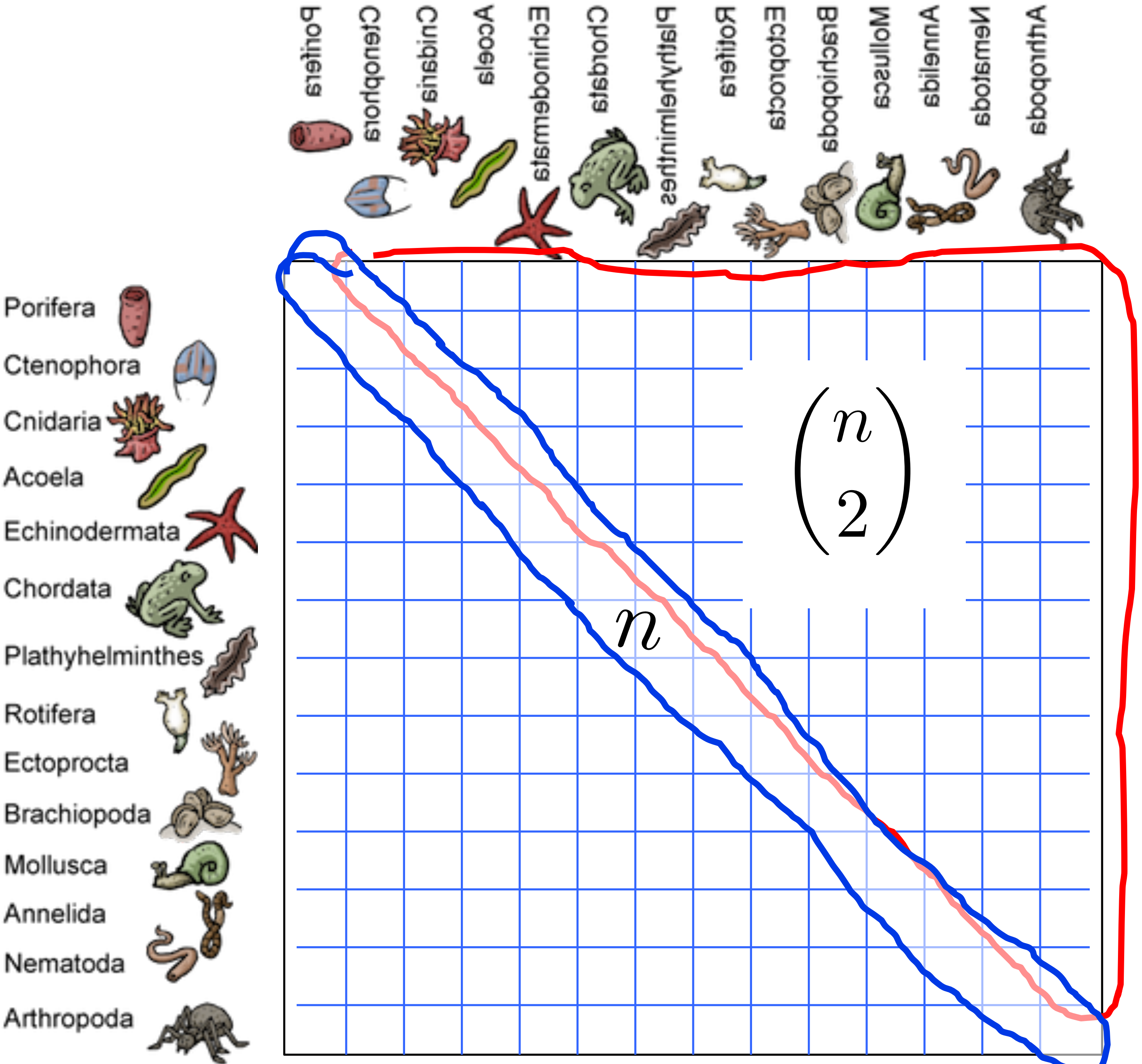
For a DNA tree we need to calculate the DNA distance between each pair of animals. How many calculations are needed?



Counting Review

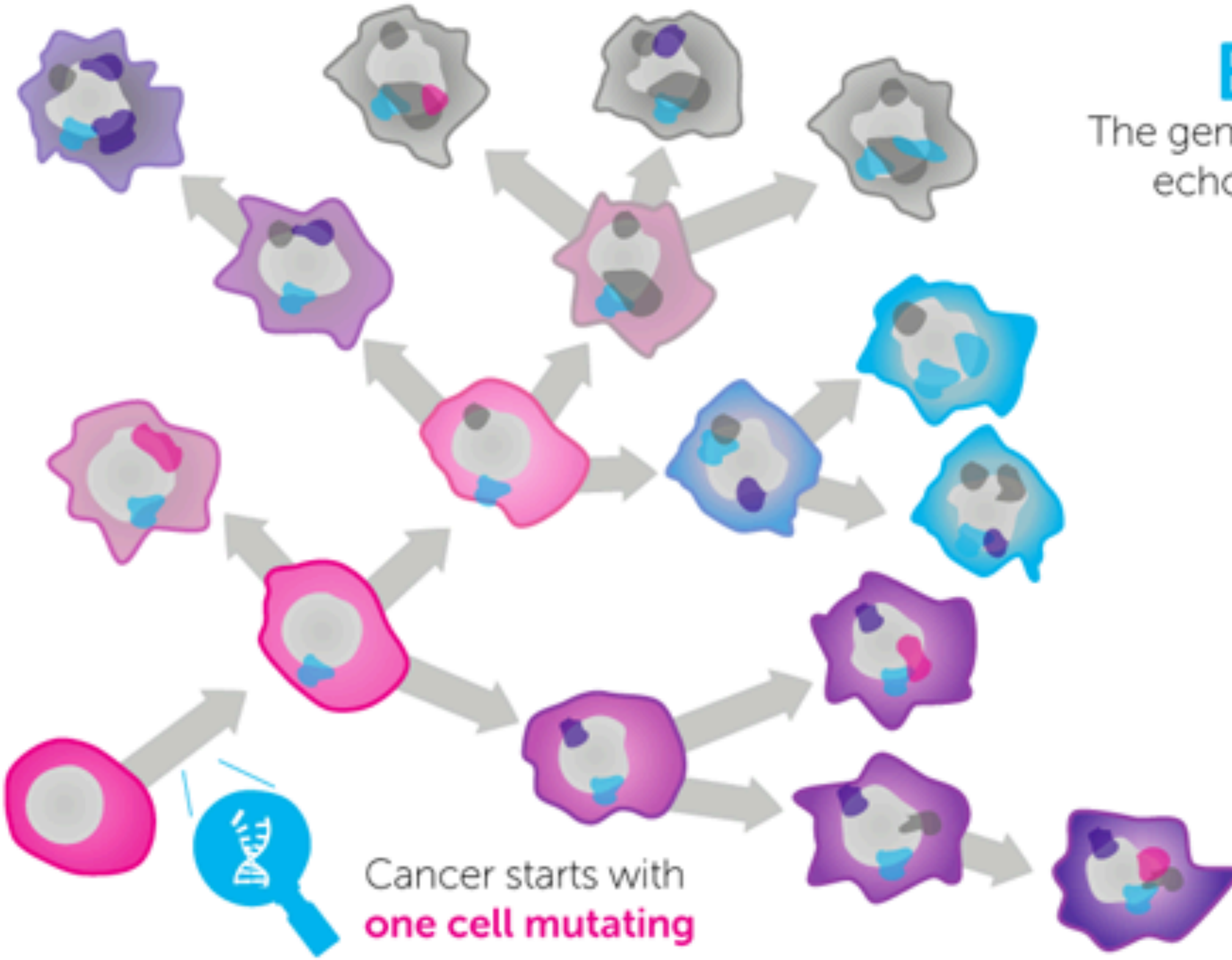
Q: There are *n* animals.
How many distinct pairs of animals are there?




Porifera
Ctenophora
Cnidaria
Acoela
Echinodermata
Chordata
Plathyhelminthes
Rotifera
Ectoprocta
Brachiopoda
Mollusca
Annelida
Nematoda
Arthropoda
$\binom{n}{2}$
n

BRANCHED EVOLUTION

The genetic diversity in a tumour echoes Darwin's **Tree of Life**.



CANCER RESEARCH UK

Counting Rules

Counting operations on *n* objects

- **Sort (order matters) (perms)**
  - Distinct: $n!$
  - Some Distinct: $\frac{n!}{n_1!n_2!\dots}$
- **Choose *k* (combinations)**
  - Distinct: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$
- **Put in *r* buckets**
  - Distinct: r^n
  - None Distinct: $\frac{(n+r-1)!}{n!(r-1)!}$

End Review

Sample Space

- **Sample space**, S, is set of all possible outcomes of an experiment
  - Coin flip: S = {Head, Tails}
  - Flipping two coins: S = {(H, H), (H, T), (T, H), (T, T)}
  - Roll of 6-sided die: S = {1, 2, 3, 4, 5, 6}
  - # emails in a day: $S = \{x \mid x \in \mathbf{Z}, x \geq 0\}$ (non-neg. ints)
  - YouTube hrs. in day: $S = \{x \mid x \in \mathbf{R}, 0 \leq x \leq 24\}$

Events

- **Event**, E, is some subset of S $\quad (E \subseteq S)$
  - Coin flip is heads: $E = \{\text{Head}\}$
  - ≥ 1 head on 2 coin flips: $E = \{(H, H), (H, T), (T, H)\}$
  - Roll of die is 3 or less: $E = \{1, 2, 3\}$
  - # emails in a day ≤ 20: $E = \{x \mid x \in \mathbf{Z}, 0 \leq x \leq 20\}$
  - Wasted day (≥ 5 YT hrs.): $E = \{x \mid x \in \mathbf{R}, 5 \leq x \leq 24\}$

Note: When Ross uses: $\subset$, he really means: $\subseteq$

What is a probability?

Number between 0 and 1

Ascribe Meaning

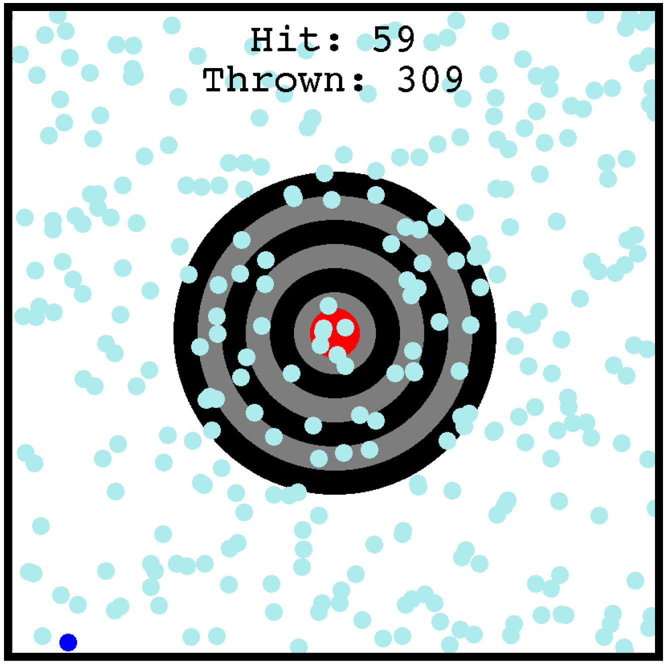
$$P(E)$$

* Our belief that an event E occurs

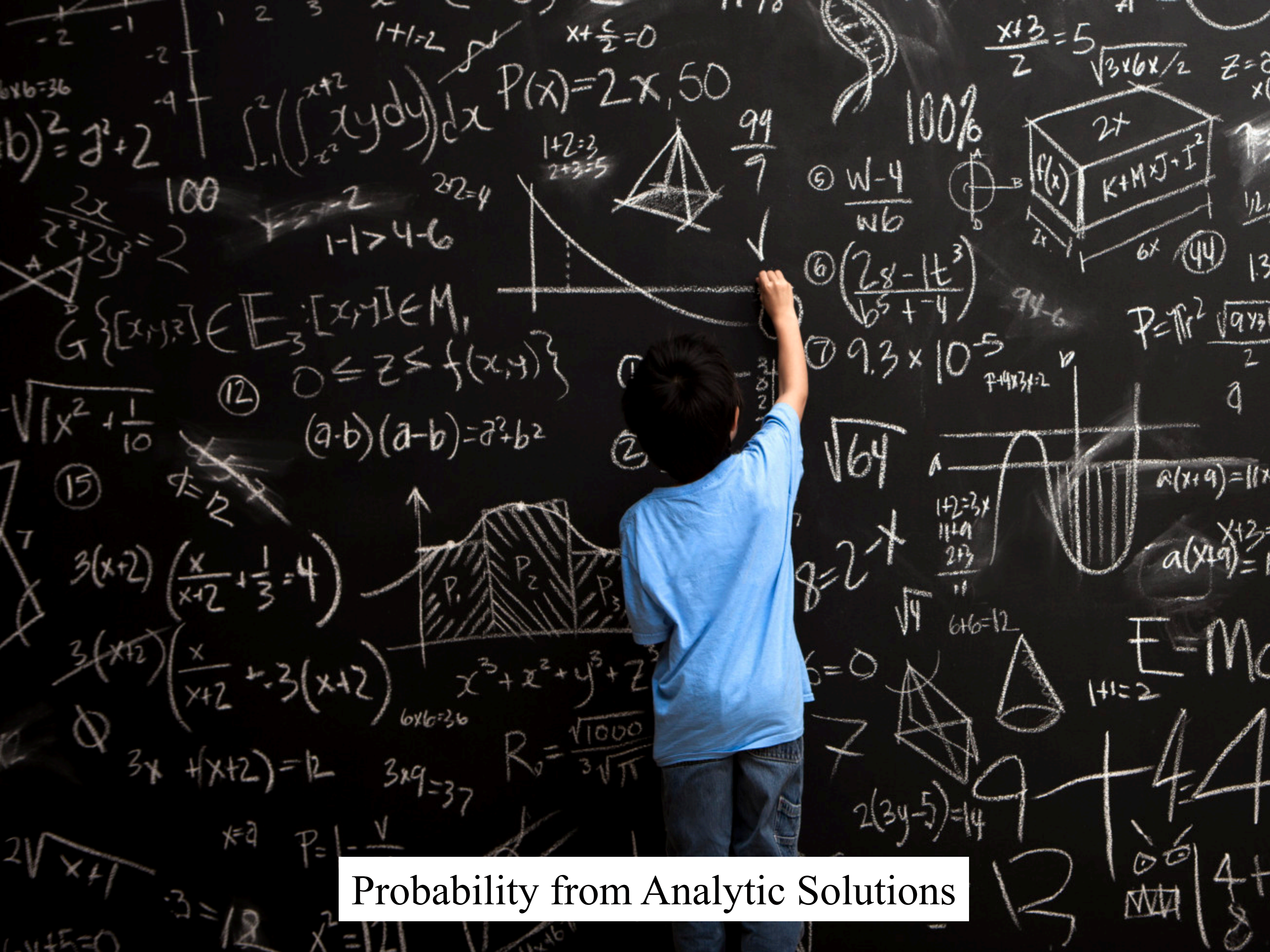
What is a Probability

What is a probability?

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$




90
MLB 040



Probability from Analytic Solutions

Axioms of Probability

Recall: S = all possible outcomes. E = the event.

- Axiom 1: $0 \leq P(E) \leq 1$
- Axiom 2: $P(S) = 1$
- Axiom 3: $P(E^c) = 1 - P(E)$

Equally Likely Outcomes

- Some sample spaces have equally likely outcomes
  - Coin flip: S = {Head, Tails}
  - Flipping two coins: S = {(H, H), (H, T), (T, H), (T, T)}
  - Roll of 6-sided die: S = {1, 2, 3, 4, 5, 6}
- P(Each outcome) = $\frac{1}{|S|}$
- In that case, P(E) = $\frac{\text{number of outcomes in E}}{\text{number of outcomes in S}} = \frac{|E|}{|S|}$

Rolling Two Dice

- Roll two 6-sided dice.
  - What is P(sum = 7)?

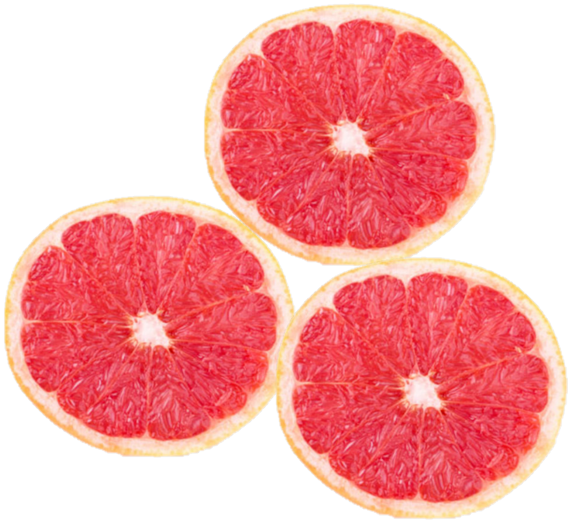
- S = {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6),
  (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6),
  (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6),
  (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6),
  (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6),
  (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)}
- E = {(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)}

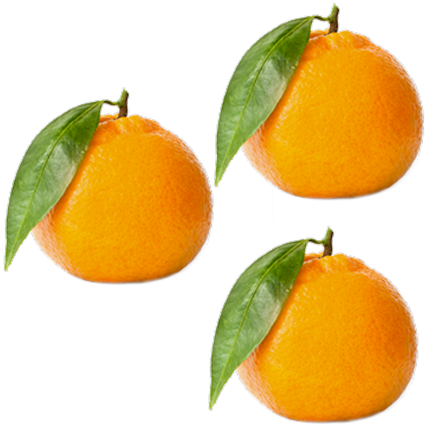
- P(sum = 7) = |E|/|S| = 6/36 = 1/6

Mandarins and Grapefruit

- 4 Mandarins and 3 Grapefruit in a Bag. 3 drawn.
  - What is P(1 Mandarin and 2 Grapefruits drawn)?

Equally likely sample space? Thought experiment





Mandarins and Grapefruit

- 4 Mandarins and 3 Grapefruit in a Bag. 3 drawn.
  - What is P(1 Mandarin and 2 Grapefruits drawn)?
- Ordered:
  - Pick 3 ordered items: |S| = 7 * 6 * 5 = 210
  - Pick Mandarin as either 1st, 2nd, or 3rd item:
    |E| = (4 * 3 * 2) + (3 * 4 * 2) + (3 * 2 * 4) = 72
  - P(1 Mandarin, 2 Grapefruit) = 72/210 = 12/35
- Unordered:
  - $|S| = \binom{7}{3} = 35$
  - $|E| = \binom{4}{1}\binom{3}{2} = 12$
  - P(1 Mandarin, 2 Grapefruit) = 12/35

Often make indistinct items distinct to get equally likely sample space outcomes

*You will need to use this “trick” with high probability

Chip Defect Detection

- n chips manufactured, 1 of which is defective.
- k chips randomly selected from n for testing.
  - What is P(defective chip is in k selected chips)?
- $|S| = \binom{n}{k}$
- $|E| = \binom{1}{1}\binom{n-1}{k-1}$
- P(defective chip is in k selected chips)

$$= \frac{\binom{1}{1}\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{\frac{(n-1)!}{(k-1)!(n-k)!}}{\frac{n!}{k!(n-k)!}} = \frac{k}{n}$$

Any "Straight" Poker Hand

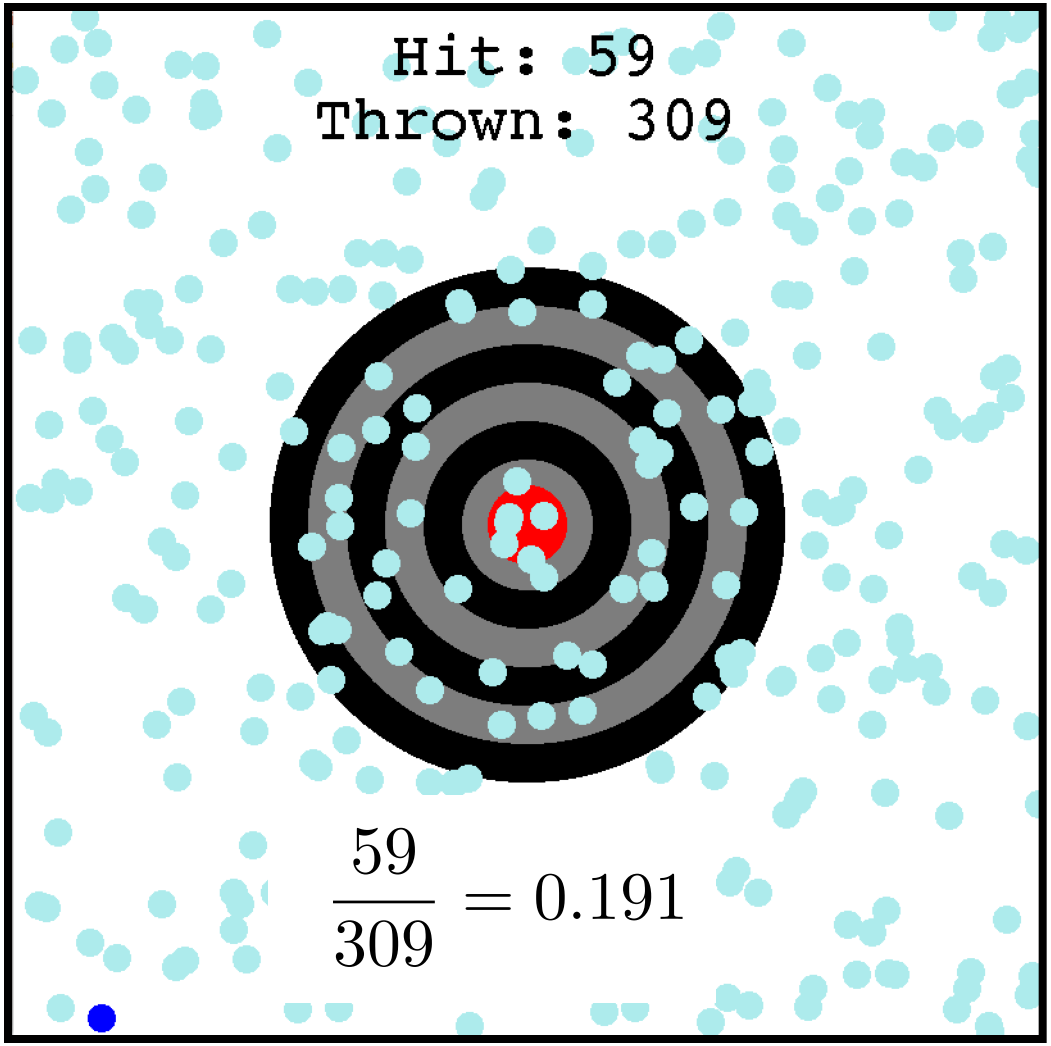
- Consider 5 card poker hands.
  - "straight" is 5 consecutive rank cards of any suit
  - What is P(straight)?
  - Note: this is a little different than the textbook
- $|S| = \binom{52}{5}$
- $|E| = 10\binom{4}{1}^5$
- $P(\text{straight}) = \dfrac{10\binom{4}{1}^5}{\binom{52}{5}} \approx 0.00394$

Official "Straight" Poker Hand

- Consider 5 card poker hands.
  - "straight" is 5 consecutive rank cards of any suit
  - "straight flush" is 5 consecutive rank cards of same suit
  - What is P(straight, but not straight flush)?
- $|S| = \binom{52}{5}$
- $|E| = 10\binom{4}{1}^5 - 10\binom{4}{1}$
- $\text{P(straight)} = \dfrac{10\binom{4}{1}^5 - 10\binom{4}{1}}{\binom{52}{5}} \approx 0.00392$

When approaching a problem, start by defining events.

Target Revisited



$$\frac{59}{309} = 0.191$$

Screen size = 800x800
Radius of target = 200

The dart is equally likely to land anywhere on the screen.

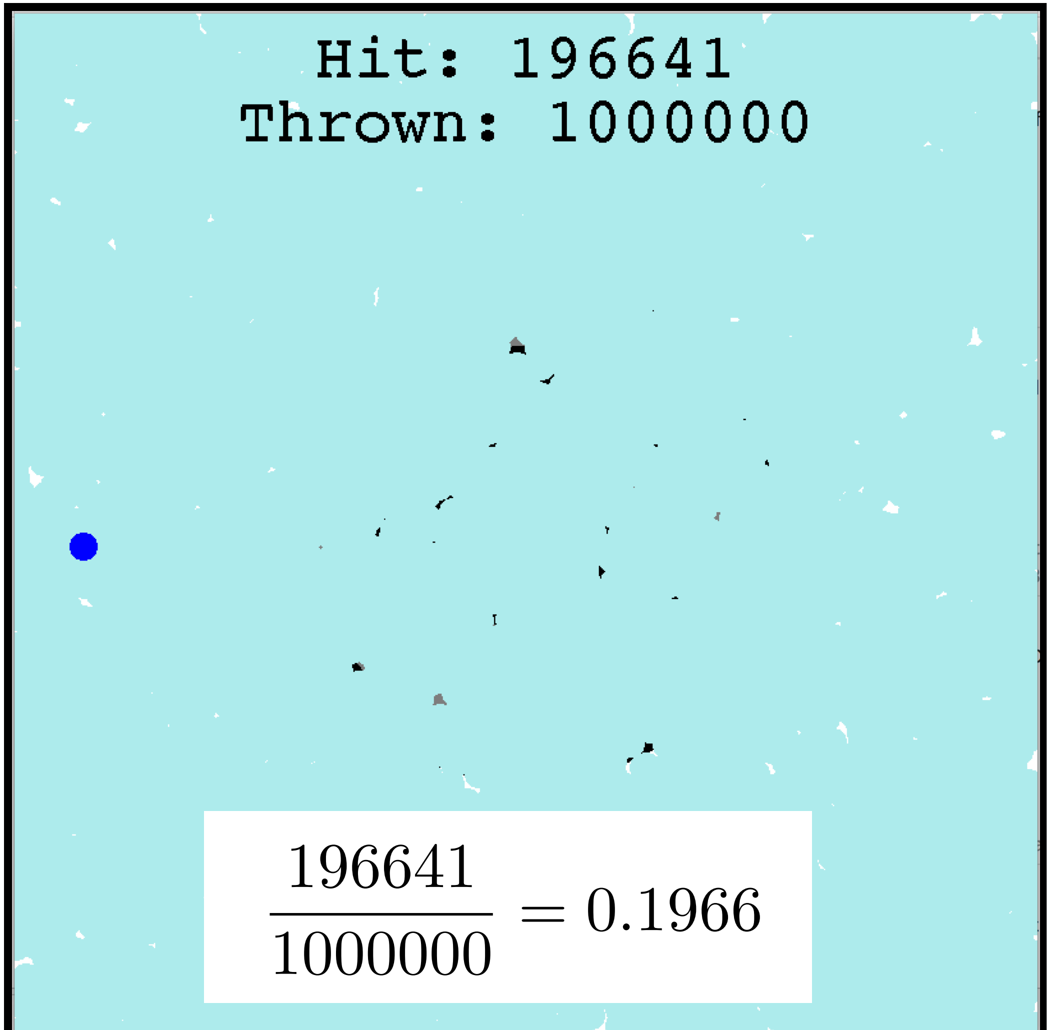
What is the probability of hitting the target?

$$|S| = 800^2$$

$$|E| = \pi 200^2$$

$$p(E) = \frac{\pi \cdot 200^2}{800^2} \approx 0.1963$$

Target Revisited



Screen size = 800x800
Radius of target = 200

The dart is equally likely to land anywhere on the screen.

What is the probability of hitting the target?

$$|S| = 800^2$$

$$|E| = \pi 200^2$$

$$p(E) = \frac{\pi \cdot 200^2}{800^2} \approx 0.1963$$

Let it find you.

SERENDIPITY

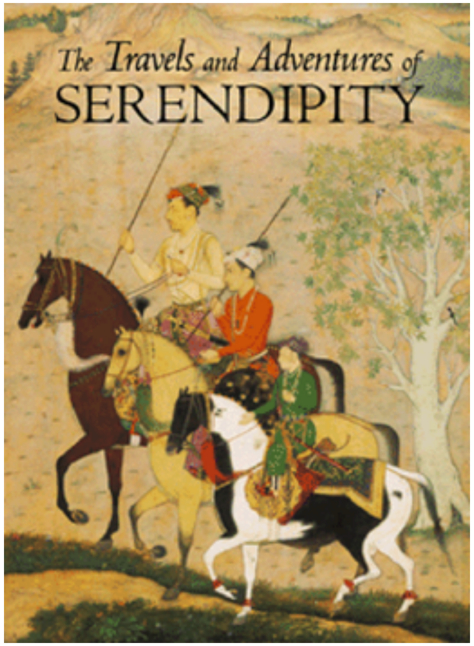
the effect by which one accidentally stumbles upon something truely wonderful, especially while looking for something entirely unrelated.

WHEN YOU MEET YOUR BEST FRIEND

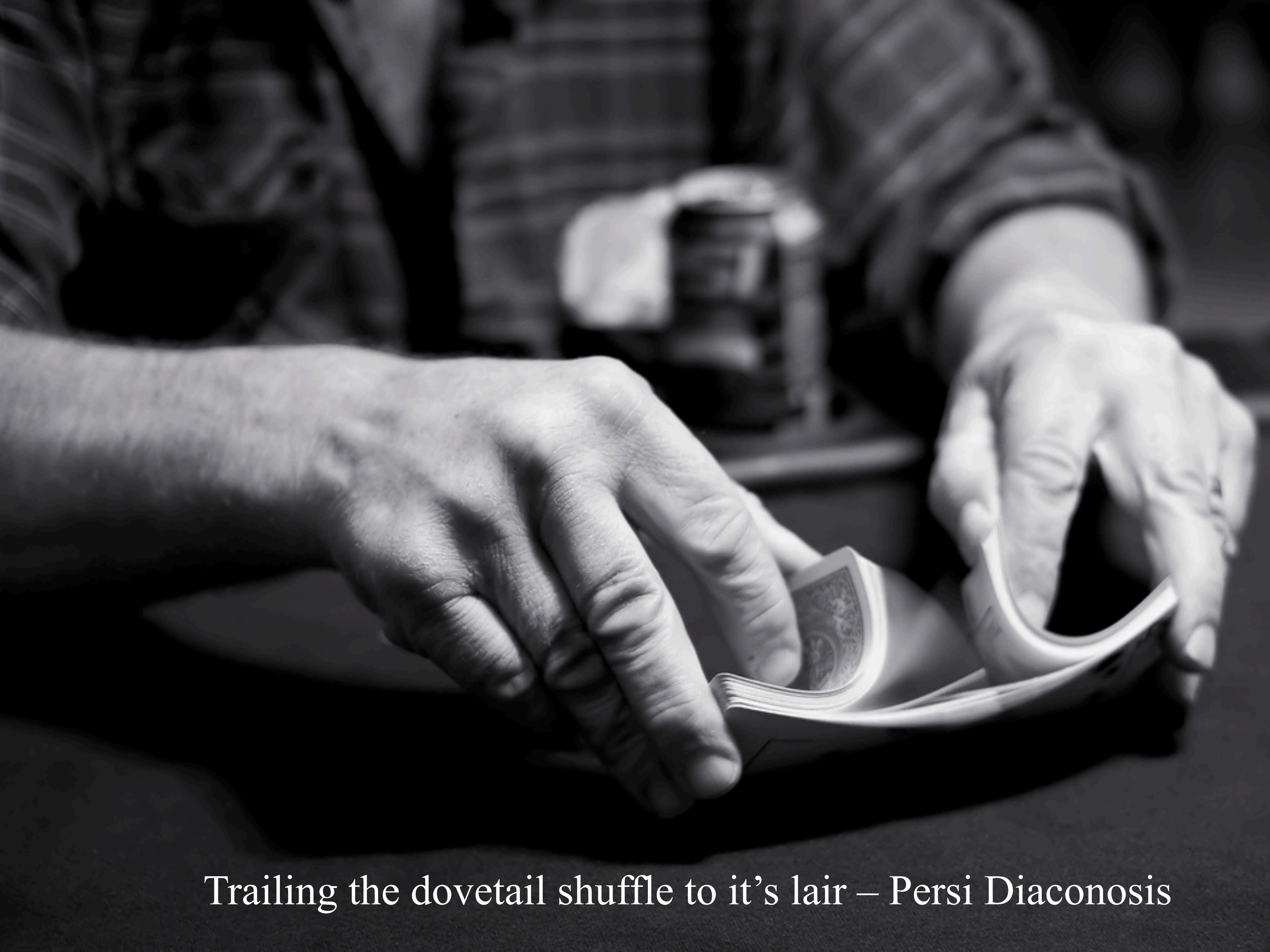
Somewhere you didn't expect to.

Serendipity

- Say the population of Stanford is 21,000 people
  - You are friends with ?
  - Walk into a room, see 240 random people.
  - What is the probability that you see someone you know?
  - Assume you are equally likely to see each person at Stanford



Many times it is easier to calculate $P(E^C)$.


Trailing the dovetail shuffle to it’s lair – Persi Diaconosis

Making History

- What is the probability that in the *n* shuffles seen since the start of time, yours is unique?
  - $|S| = (52!)^n$
  - $|E| = (52! - 1)^n$
  - P(no deck matching yours) = $(52!-1)^n/(52!)^n$

- For $n = 10^{20}$,
  - P(deck matching yours) < 0.000000001

* Assumes 7 billion people have been shuffling cards once a second since cards were invented