3. Discrete Probability

CSE 312

Winter 2017

W.L. Ruzzo

Probability theory:

“an aberration of the intellect”

and

“ignorance coined into science”

– John Stuart Mill

Sample space: *S* is a set of all potential outcomes of an experiment (often Ω in text books–Greek uppercase omega)

Coin flip: S = {Heads, Tails}

Flipping two coins: S = {(H,H), (H,T), (T,H), (T,T)}

Roll of one 6-sided die: S = {1, 2, 3, 4, 5, 6}

emails in a day: $S = \{ x : x \in Z, \ x \geq 0 \}$

YouTube hrs. in a day: $S = \{ x : x \in R, 0 \leq x \leq 24 \}$

Some fine print: "sample space" for an experiment isn't uniquely defined, & "potential" outcomes may include literally impossible ones, e.g., S={1,2,3,4,5,6,7} for a 6-sided die; it's all OK if you're sensible and consistent, e.g., if you make probability(7)=0. Rare to see things quite this wacky, but bottom line: a sample space is just a set, any set.

Events: $E \subseteq S$ is an arbitrary subset of the sample space

Coin flip is heads: $E = \{Head\}$

At least one head in 2 flips: $E = \{(H,H), (H,T), (T,H)\}$

Roll of die is odd: $E = \{1, 3, 5\}$

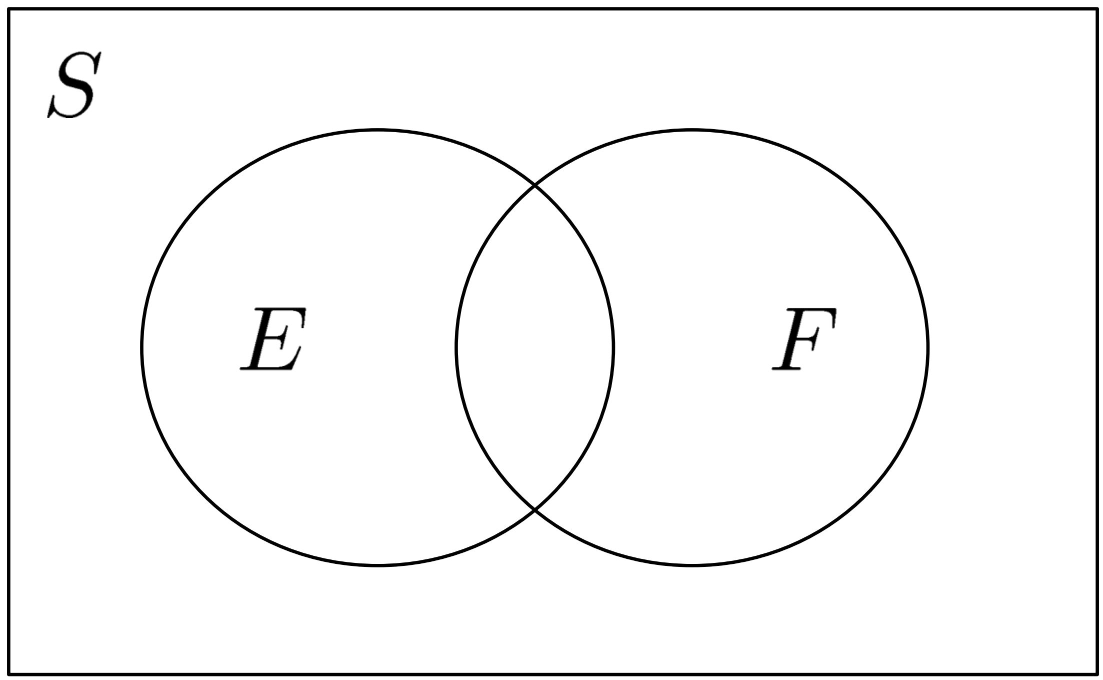
emails in a day < 20: $E = \{ x : x \in Z,\ 0 \leq x < 20 \}$

emails in a day is prime: $E = \{ 2, 3, 5, 7, 11, 13, \ldots \}$

Wasted day (>5 YT hrs): $E = \{ x : x \in R,\ x > 5 \}$

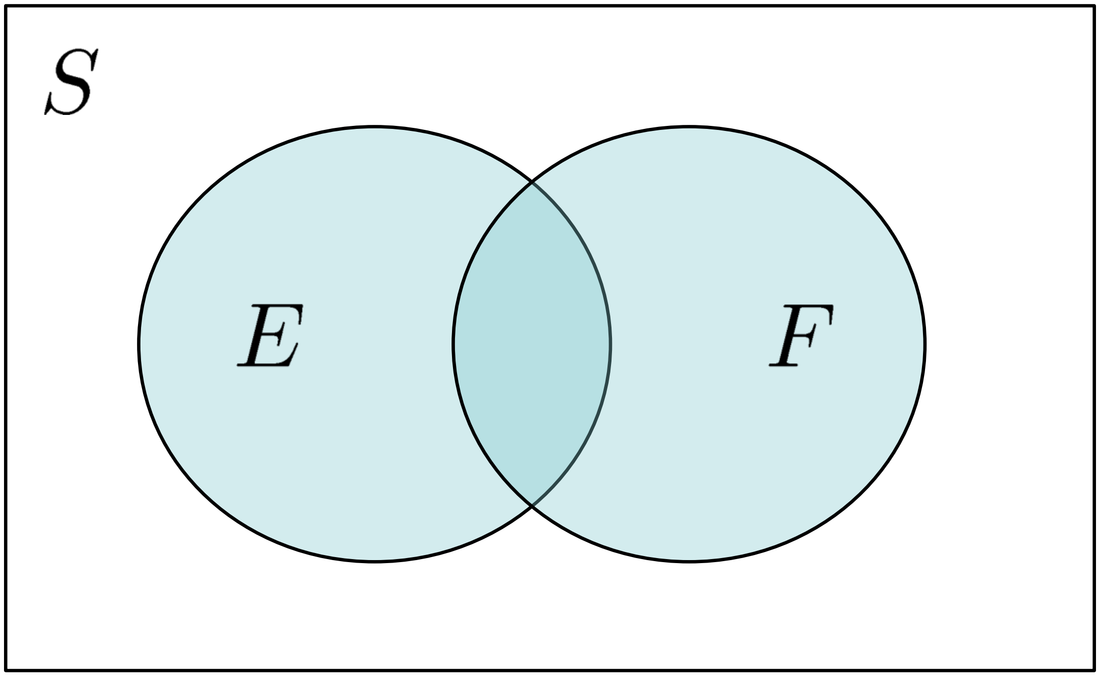
Note: an *event* is not an *outcome*, it is a set of outcomes. E.g., the outcome of rolling a die is always a *single* number in 1..6; "roll is odd" aggregates 3 potential outcomes as one event; "roll is >5" aggregates 1 potential outcome as the event E = {6} (a singleton set; sole element is the outcome 6).

E and F are events in the sample space S



E and F are events in the sample space S

Event "E OR F", written $E \cup F$

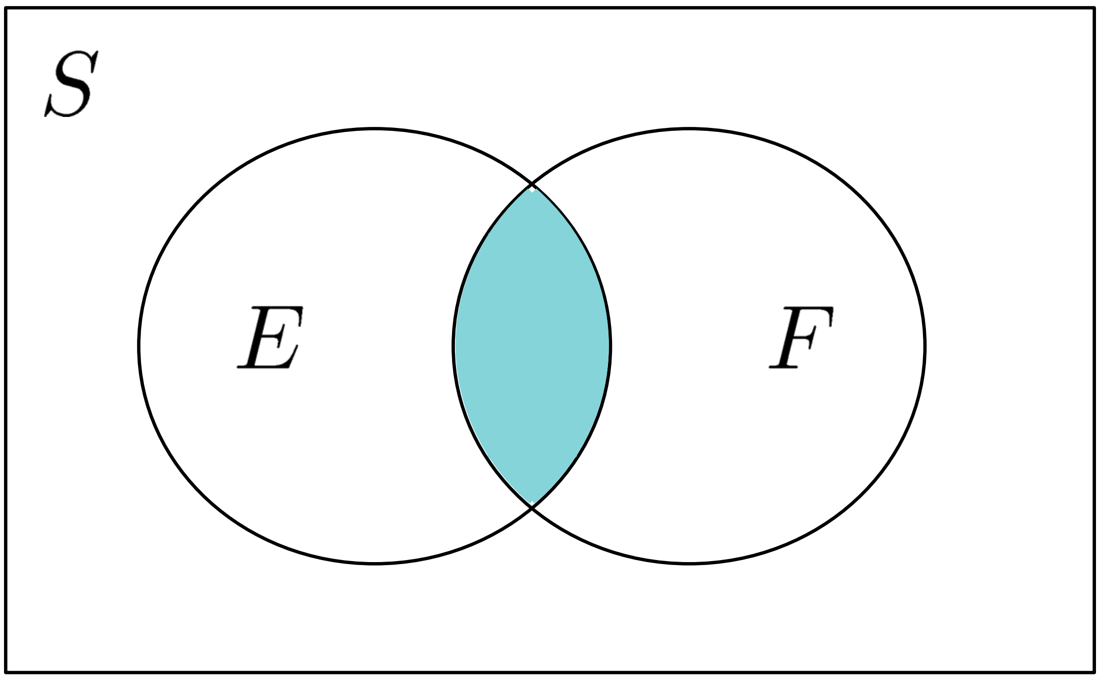


$S = \{1,2,3,4,5,6\}$
outcome of one die roll

$E = \{1,2\}$, $F = \{2,3\}$
$E \cup F = \{1, 2, 3\}$

E and F are events in the sample space S

Event "E AND F", written E ∩ F or EF

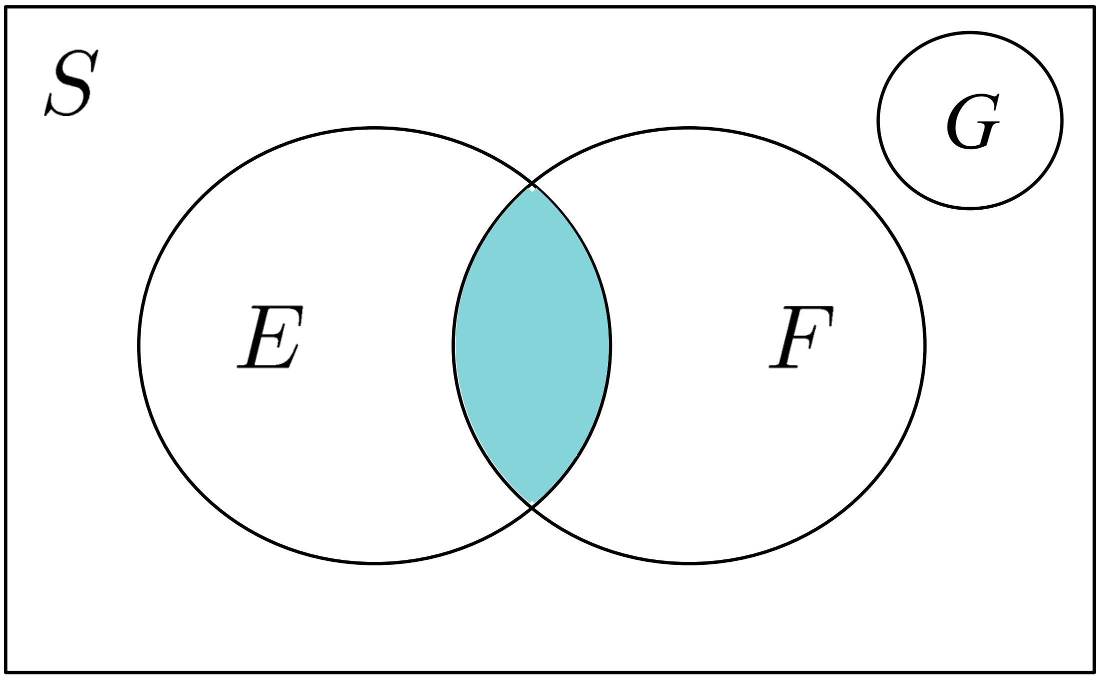


S = {1,2,3,4,5,6}
outcome of one die roll

E = {1,2}, F = {2,3}
E ∩ F = {2}

E and F are events in the sample space S

EF = ∅ ⇔ E,F are "mutually exclusive"

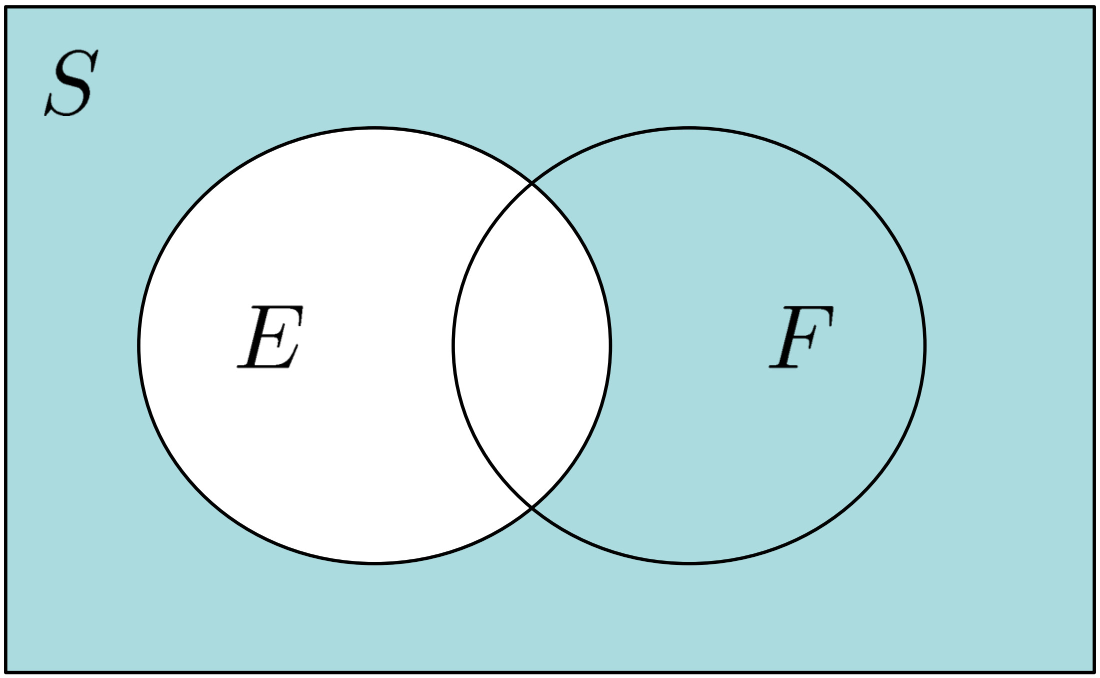


S = {1,2,3,4,5,6}
outcome of one die roll

E = {1,2}, F = {2,3}, G={5,6}
EF = {2}, *not* mutually exclusive, but E,G and F,G are

E and F are events in the sample space S

Event "not E," written $\overline{E}$ or $\neg E$



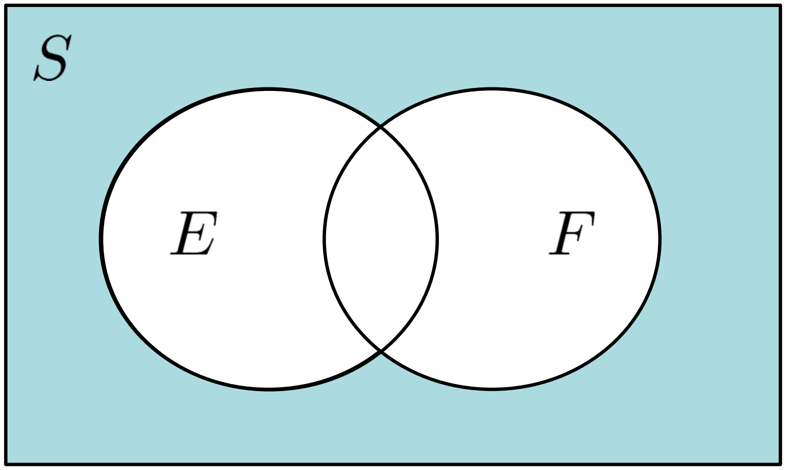
S = {1,2,3,4,5,6}
outcome of one die roll

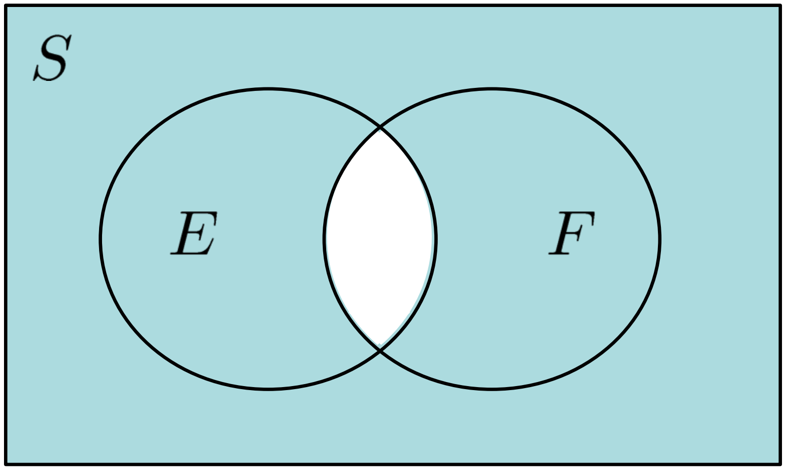
$E = \{1, 2\}$ $\neg E = \{3, 4, 5, 6\}$

DeMorgan's Laws

$$\overline{E \cup F} = \bar{E} \cap \bar{F}$$

$$\overline{E \cap F} = \bar{E} \cup \bar{F}$$





Intuition: Probability as the relative frequency of an event

$$Pr(E) = \lim_{n \to \infty} (\text{\# of occurrences of } E \text{ in n trials})/n$$

Mathematically, this proves messy to deal with.

Instead, we define "Probability" via a function from subsets of S ("events") to real numbers

$$Pr: 2^S \to \mathbb{R}$$

satisfying the properties (axioms) below.

axioms of probability

Intuition: Probability as the relative frequency of an event

$\Pr(E) = \lim_{n\to\infty}$ (# of occurrences of E in n trials)/n

Axiom 1 (Non-negativity): $0 \leq \Pr(E)$

Axiom 2 (Normalization): $\Pr(S) = 1$

Axiom 3 (Additivity):

If E and F are mutually exclusive ($EF = \emptyset$), then

$$\Pr(E \cup F) = \Pr(E) + \Pr(F)$$

For any sequence $E_1, E_2, \ldots, E_n$ of mutually exclusive events,

$$\Pr\left(\bigcup_{i=1}^{n} E_i\right) = \Pr(E_1) + \cdots + \Pr(E_n)$$

$\Pr(\overline{E}) = 1 - \Pr(E)$

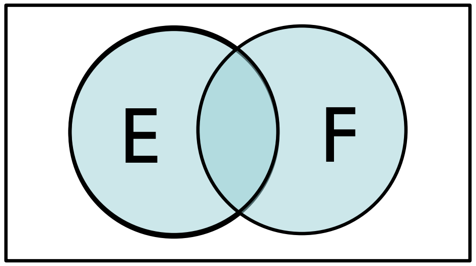
$1 = \Pr(S) = \Pr(E \cup \overline{E}) = \Pr(E) + \Pr(\overline{E})$

If $E \subseteq F$, then $\Pr(E) \leq \Pr(F)$

$\Pr(F) = \Pr(E) + \Pr(F - E) \geq \Pr(E)$

$\Pr(E \cup F) = \Pr(E) + \Pr(F) - \Pr(EF)$

inclusion-exclusion



$\Pr(E) \leq 1$

exercise

And many others

Sample space: S = set of all potential outcomes of experiment

E.g., flip two coins: S = {(H,H), (H,T), (T,H), (T,T)}

Events: **E** $\subseteq$ **S** is an arbitrary subset of the sample space

≥1 head in 2 flips: E = {(H,H), (H,T), (T,H)} S =

Probability:

A function from subsets of S to real numbers – $\Pr: 2^S \to \mathbb{R}$

Probability Axioms:

Axiom 1 (Non-negativity): $0 \le \Pr(E)$

Axiom 2 (Normalization): $\Pr(S) = 1$

Axiom 3 (Additivity): $EF = \varnothing \Rightarrow \Pr(E \cup F) = \Pr(E) + \Pr(F)$

Simplest case: sample spaces with equally likely outcomes.

Coin flips: S = {Heads, 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}

Pr(each outcome) = $\frac{1}{|S|}$

In that case,

$$\Pr(E) = \frac{\text{number of outcomes in } E}{\text{number of outcomes in } S} = \frac{|E|}{|S|}$$

And, conveniently, we've already studied counting

Why? Axiom 3 plus fact that E = union of singletons in E

Roll two 6-sided dice. What is Pr(sum of dice = 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) }

Side point: *S* is small; can write out explicitly, but how would you visualize the analogous problem with 10^3-sided dice?

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

Pr(sum = 7) = |E|/|S| = 6/36 = 1/6.

Roll two 6-sided dice. What is Pr(sum of dice = 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),
(4,1),
(5,1),
(6,1),

E = { (6,1),

Pr(sum = 7) =

SIDEBAR

It's perhaps tempting to try S={2,3,…,12} and E={7} for this problem. This isn't wrong, but note that it doesn't fit the "equally likely outcomes" scenario. E.g., Pr({2})=1/36 ≠ 1/6=Pr({7}). Plus, it's *usually best to make "S" a simple representation of the "experiment" at hand*, e.g., an ordered pair reflecting the 2 dice rolls, rather than a more complex derivative of it, like their sum. The later makes it easy to express *this* event ("sum is 7"), but makes it difficult or impossible to express other events of potential interest ("product is odd," for example).

twinkies and ding dongs

4 Twinkies and 3 DingDongs in a bag. 3 drawn.
What is Pr(one Twinkie and two DingDongs drawn) ?

Ordered: (S: ordered triples with 3 of 7 distinguishable objects)

- Pick 3, one after another: |S| = 7 • 6 • 5 = 210
- Pick Twinkie as either 1st, 2nd, or 3rd item:
  |E| = (4•3•2) + (3•4•2) + (3•2•4) = 72
- Pr(1 Twinkie and 2 DingDongs) = 72/210 = 12/35.

Unordered: (S: *un*ordered triples with 3 of 7 distinguishable objects)

- Grab 3 at once: |S| $= \binom{7}{3} = 35$
- |E| $= \binom{4}{1}\binom{3}{2} = 12$
- Pr(1 Twinkie and 2 DingDongs) = 12/35.

Exercise: a 3rd way – S is ordered list of 7, E is "1st 3 OK"; same answer?

birthdays

What is the probability that, of n people, none share the same birthday?

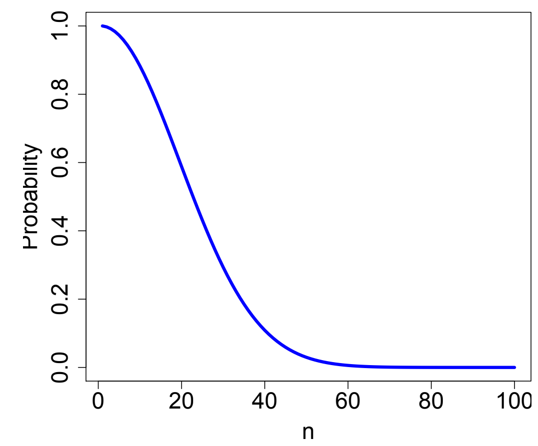
What are S, E??

$|S| = (365)^n$

$|E| = (365)(364)(363)\cdots(365-n+1)$

Pr(no matching birthdays) = |E|/|S|

$= (365)(364)\ldots(365-n+1)/(365)^n$



Some values of n…

n = 23: Pr(no matching birthdays) < 0.5

n = 77: Pr(no matching birthdays) < 1/5000

n = 90: Pr(no matching birthdays) < 1/162,000

n = 100: Pr(no matching birthdays) < 1/3,000,000

n = 150: Pr(…) < 1/3,000,000,000,000,000

n = 366?

Pr = 0

Above formula gives this, since

$(365)(364)\ldots(365-n+1)/(365)^n == 0$

when n = 366 (or greater).

Even easier to see via pigeon hole principle.

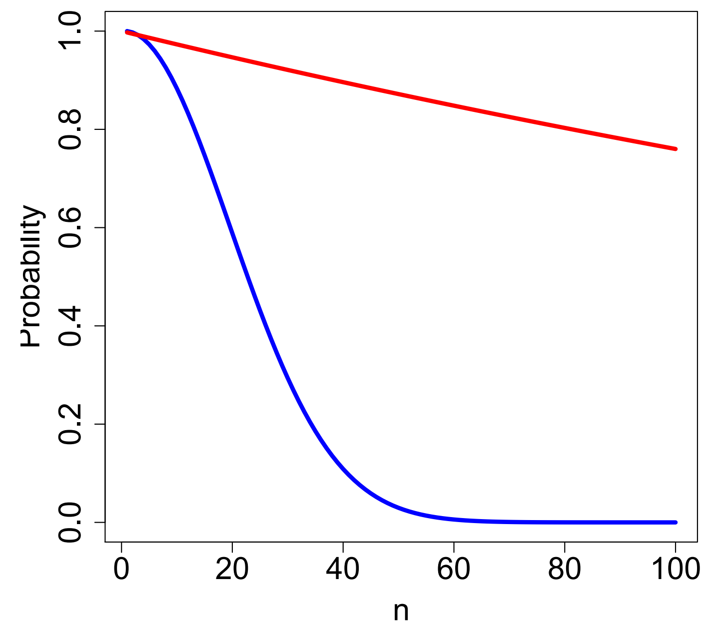
What is the probability that, of n people, none share the same birthday as <u>you</u>?

$|S| = (365)^n$

$|E| = (364)^n$

Pr(no birthdays = yours)

$= |E|/|S| = (364)^n/(365)^n$



Some values of n…

n = 23: Pr(no matching birthdays) ≈ 0.9388

n = 90: Pr(no matching birthdays) ≈ 0.7812

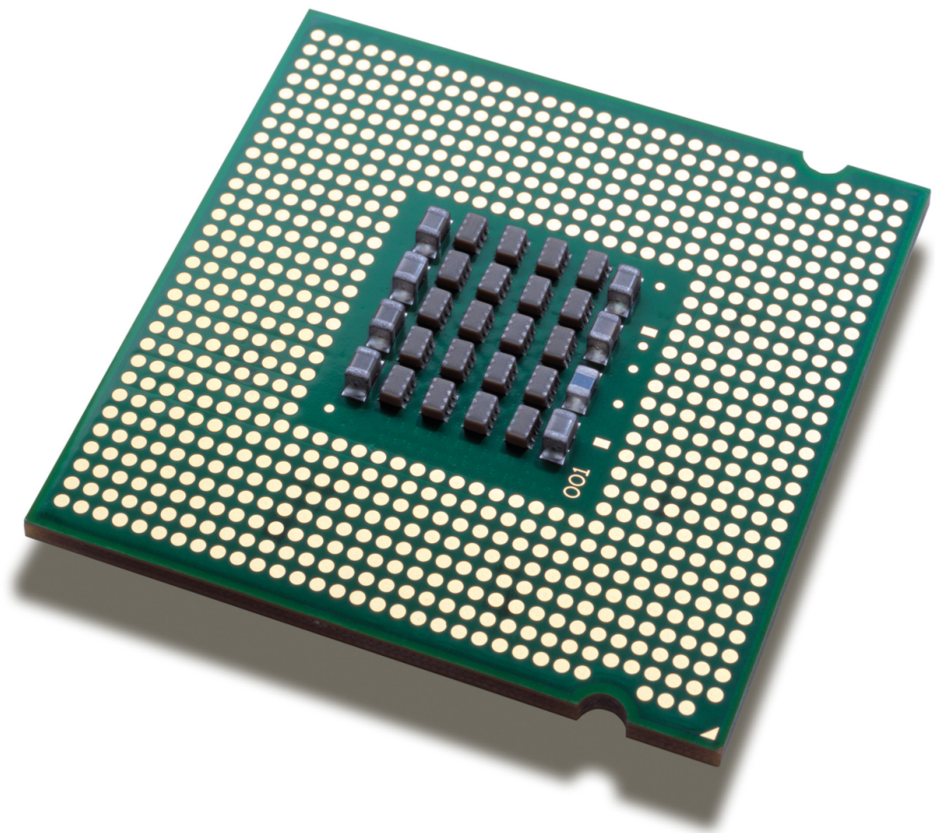
n = 253: Pr(no matching birthdays) ≈ 0.4995

Exercise: p^n is not linear, but red line looks straight. Why?

Q: If you hash 23 entries into a hash table with 365 buckets, what is the chance that there will be no collisions?

A: <1/2 *even when the hash table is >93% empty!*

chip defect detection



n chips manufactured, one of which is defective

k chips randomly selected from n for testing

What is Pr(defective chip is in k selected chips) ?

|S| = $\binom{n}{k}$

|E| = $\binom{1}{1}\binom{n-1}{k-1}$

Pr(defective chip is among k selected chips)

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

n chips manufactured, one of which is defective

k chips randomly selected from n for testing

What is Pr(defective chip is in k selected chips) ?

Different analysis:

- Select k chips at random by permuting all n chips and then choosing the first k.
- Let E_i = event that i^{th} selected chip is defective.
- Events $E_1, E_2, \ldots, E_k$ are mutually exclusive
- $\Pr(E_i) = 1/n$ for $i=1,2,\ldots,k$
- Thus Pr(defective chip is selected)

$$= \Pr(E_1) + \cdots + \Pr(E_k) = k/n.$$

n chips manufactured, *two* of which are defective

k chips randomly selected from n for testing

What is Pr(a defective chip is in k selected chips) ?

$|S| = \binom{n}{k}$ $|E|$ = (1 chip defective) + (2 chips defective)

$$= \binom{2}{1}\binom{n-2}{k-1} + \binom{2}{2}\binom{n-2}{k-2}$$

Pr(a defective chip is in k selected chips)

$$= \frac{\binom{2}{1}\binom{n-2}{k-1} + \binom{2}{2}\binom{n-2}{k-2}}{\binom{n}{k}}$$

n chips manufactured, *two* of which are defective

k chips randomly selected from n for testing

What is Pr(a defective chip is in k selected chips) ?

Another approach:

Pr(a defective chip is in k selected chips) = 1-Pr(none)

Pr(none):

$$|S| = \binom{n}{k}, |E| = \binom{n-2}{k}, Pr(\text{none}) = \frac{\binom{n-2}{k}}{\binom{n}{k}}$$

Pr(a defective chip is in k selected chips) = $1 - \frac{\binom{n-2}{k}}{\binom{n}{k}}$

(Same as above? Check it!)

poker hands

poker hands

5 card poker hands (ordinary 52 card deck, no jokers etc.)

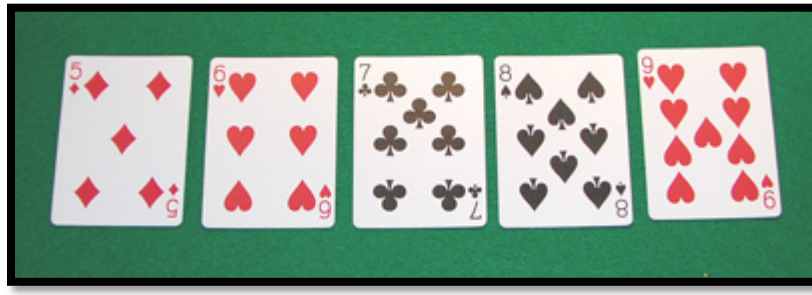
flush, 1 pair, 3 of a kind, 2 pairs, full house, …

Sample Space?

Imagine sorted tableau of cards, pick 5:

A♥	2♥	3♥	…	10♥	J♥	Q♥	K♥
A♣	2♣	3♣	…	10♣	J♣	Q♣	K♣
A♦	2♦	3♦	…	10♦	J♦	Q♦	K♦
A♠	2♠	3♠	…	10♠	J♠	Q♠	K♠

$|S| = \binom{52}{5}$



Consider 5 card poker hands.

A "straight" is 5 consecutive rank cards ignoring suit (Ace low or high, but not both. E.g., A,2,3,4,5 or 10,J,Q,K,A)

What is Pr(straight) ?

S as on previous slide, |S| = $\binom{52}{5}$

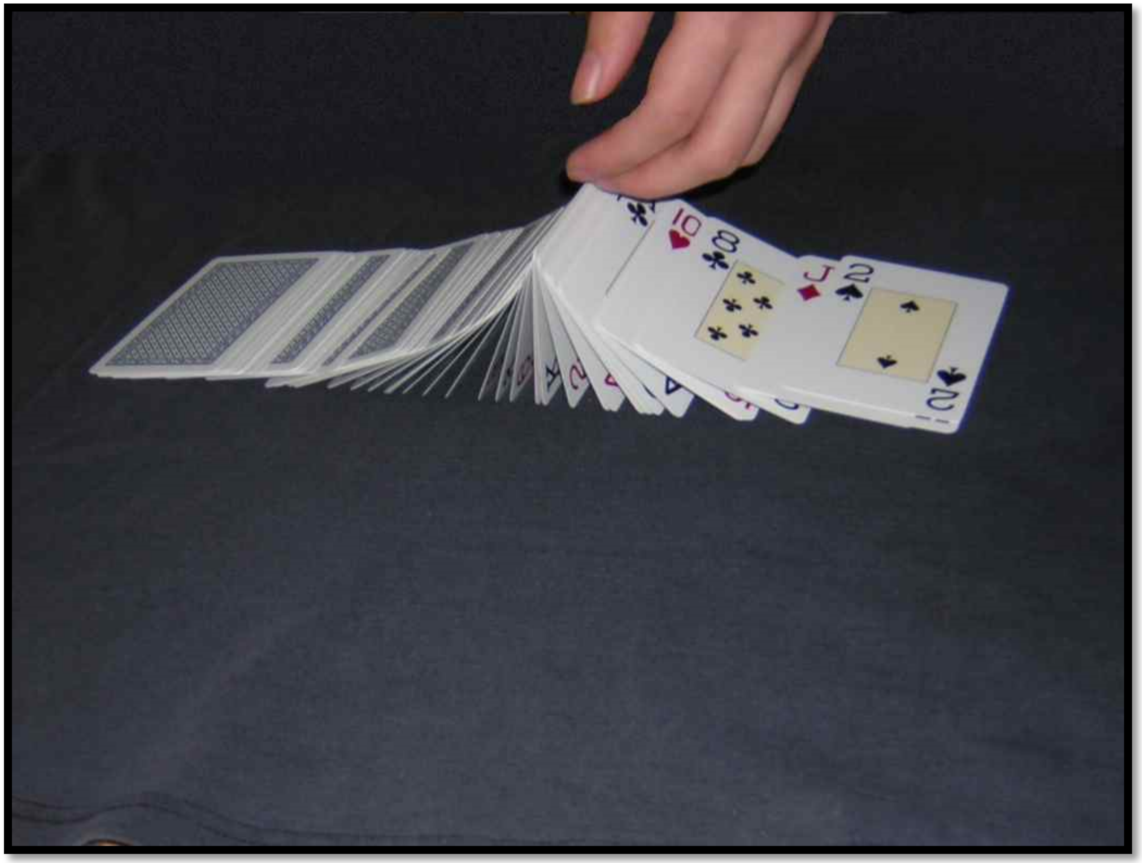
What's E?

E = Pick col A, 2, … 10, then 1 of 4 in 5 consecutive cols (wrap K→A)

$|E| = \binom{10}{1}\binom{4}{1}^5$

$\Pr(\text{straight}) = \dfrac{\binom{10}{1}\binom{4}{1}^5}{\binom{52}{5}} \approx 0.00394$

card flipping



52 card deck. Cards flipped one at a time.

After first ace (of any suit) appears, consider next card

Pr(next card = ace of spades) < Pr(next card = 2 of clubs) ?

Maybe, Maybe Not …

S = all permutations of 52 cards, |S| = 52!

Event 1: Next = Ace of Spades.

Remove A♠, shuffle remaining 51 cards, add A♠ after first Ace

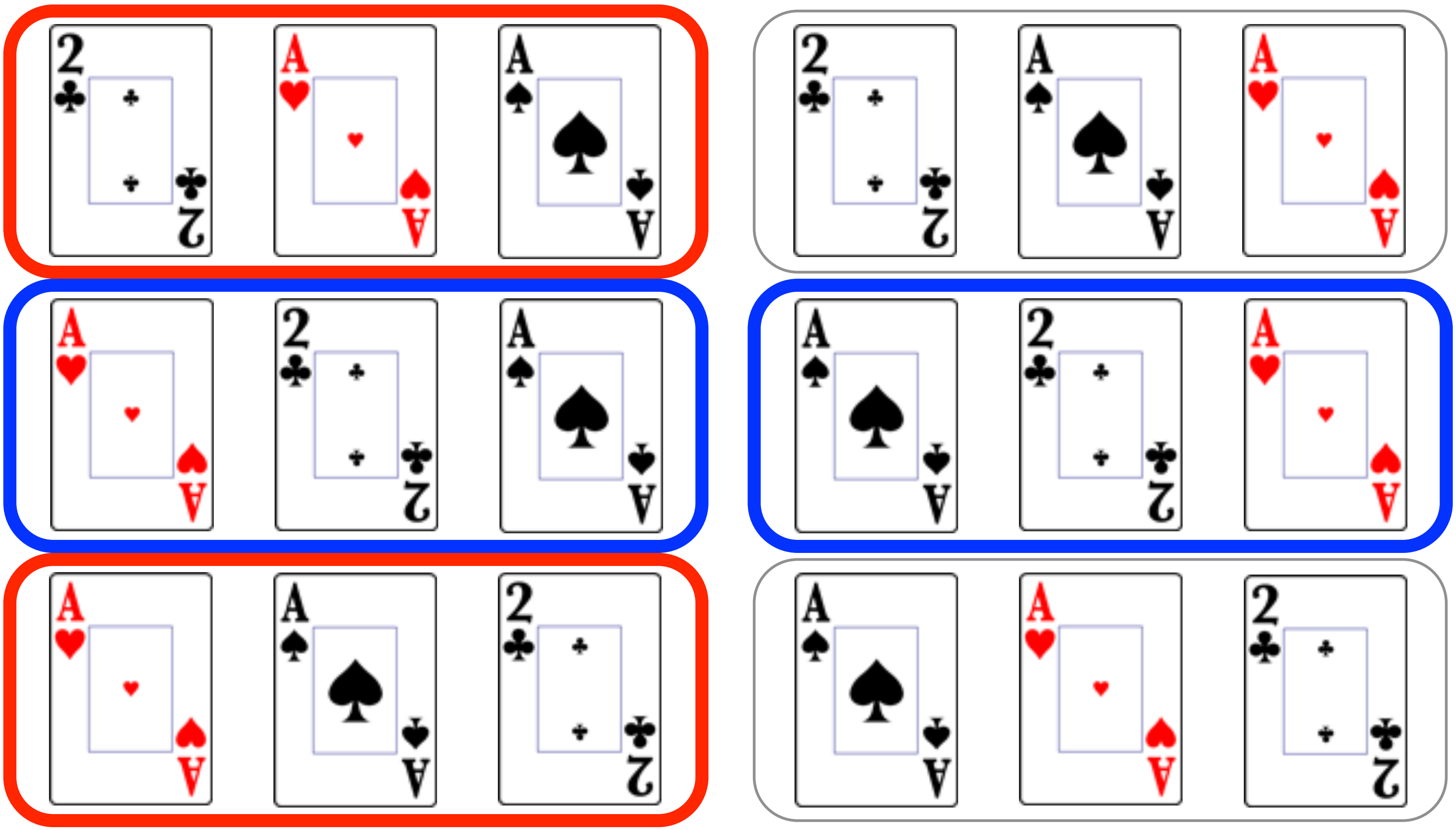
$|E_1| = 51!$ (only 1 place A♠ can be added)

Event 2: Next = 2 of Clubs

Do the same thing with 2♣; E_1 and E_2 have same size

So, $\Pr(E_1) = \Pr(E_2) = 51!/52! = 1/52$

Ace of Spades: 2/6 — 2 of Clubs: 2/6



Card images from http://www.eludication.org/playingcards.html

Theory is the same for a 3-card deck; $Pr = 2!/3! = 1/3$

hats

n persons at a party throw hats in a pile, select at random. What is Pr(no one gets own hat)?

Pr(no one gets own hat) =
 1 – Pr(someone gets own hat)

Pr(someone gets own hat) = $Pr(\cup_{i=1}^{n} E_i)$, where
E_i = event that person i gets own hat

$$Pr(\cup_{i=1}^{n} E_i) = \Sigma_i P(E_i) - \Sigma_{i<j} Pr(E_i E_j) + \Sigma_{i<j<k} Pr(E_i E_j E_k) \ldots$$

Visualizing the sample space *S*:

People:
Hats:

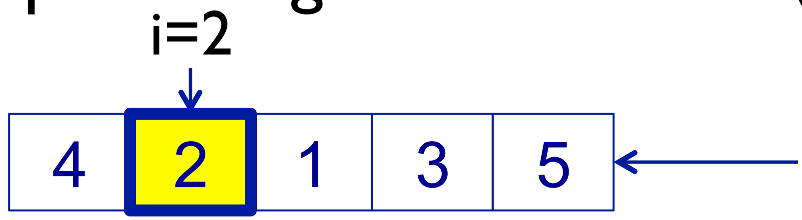
P_1	P_2	P_3	P_4	P_5
H_4	H_2	H_5	H_1	H_3

I.e., a sample point is a *permutation* π of 1, …, n

4	2	5	1	3

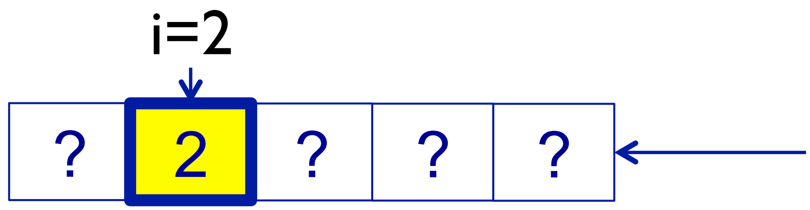
$|S| = n!$

E_i = event that person i gets own hat: $\pi(i) = i$



A sample point in E_2 (also in E_5)

Counting single events:



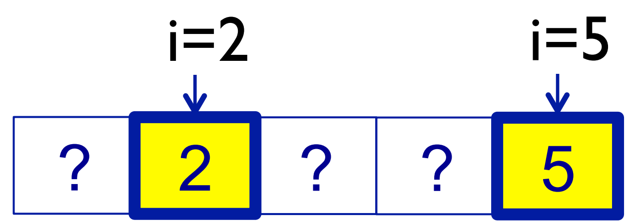
All points in E_2

$|E_i| = (n-1)!$ for all i

Counting pairs:

E_iE_j : $\pi(i) = i$ & $\pi(j) = j$

$|E_iE_j| = (n-2)!$ for all i, j



All points in $E_2 \cap E_5$

n persons at a party throw hats in middle, select at random. What is Pr(no one gets own hat)?

E_i = event that person i gets own hat

$Pr(\cup_{i=1}^{n} E_i) = \Sigma_i P(E_i) - \Sigma_{i<j} Pr(E_i E_j) + \Sigma_{i<j<k} Pr(E_i E_j E_k) \ldots$

Pr(k fixed people get own back) = (n-k)!/n!

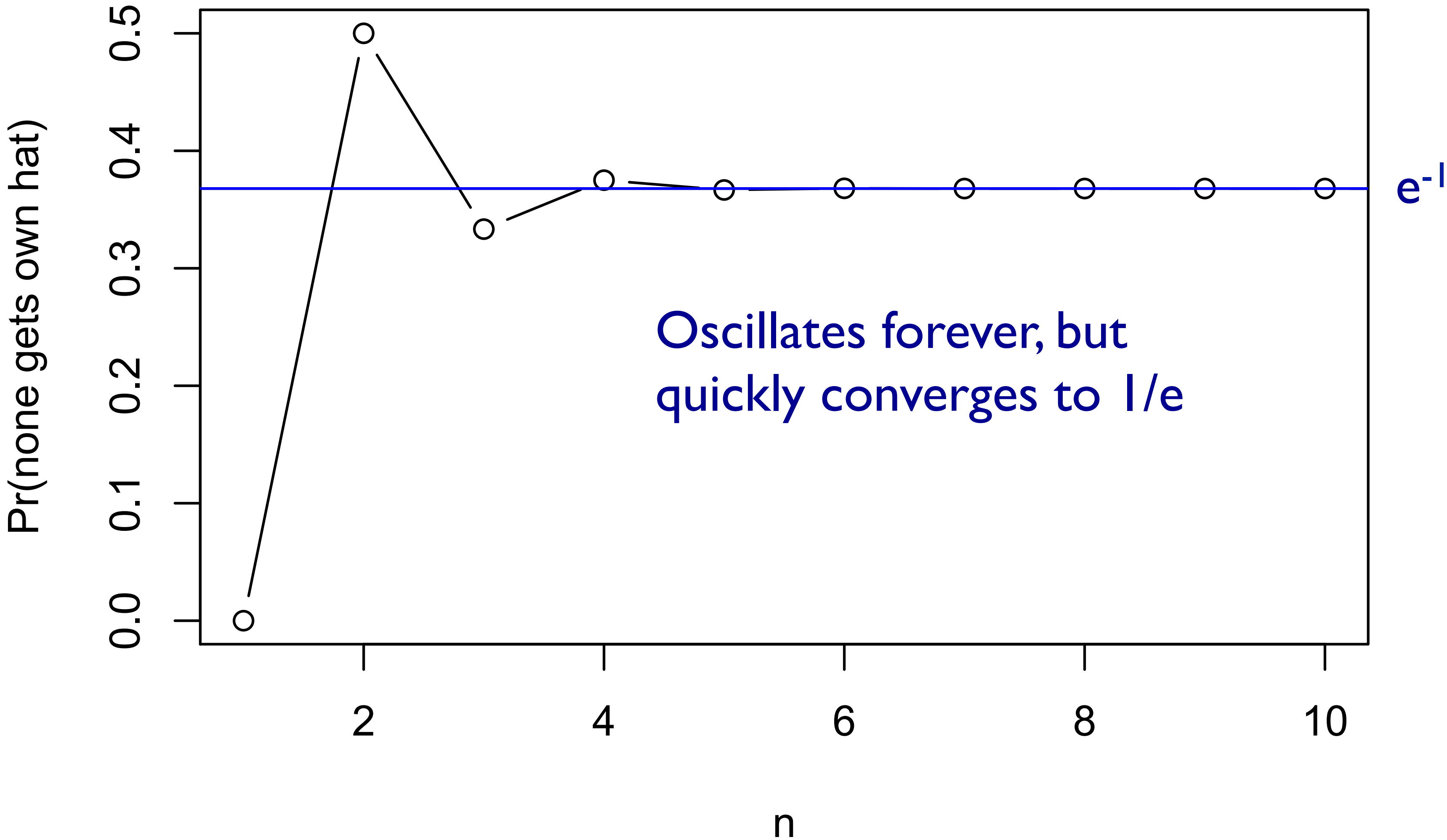
$\binom{n}{k}$ times that = $\frac{n!}{k!(n-k)!} \frac{(n-k)!}{n!} = 1/k!$

Pr(none get own) = 1-Pr(some do) =

$1 - 1/1! + 1/2! - 1/3! + 1/4! \ldots + (-1)^n/n! \approx 1/e \approx .37$

Pr(none get own) = 1 - Pr(some do) =

$1 - 1 + 1/2! - 1/3! + 1/4! \ldots + (-1)^n/n! \approx e^{-1} \approx .37$



Sample spaces
Events

} Visualize!

Set theory

Axioms

Simple identities

Equally likely outcomes (counting)

Examples

- All good for building your skills
- Birthdays is particularly important for applications
- Hats is important as example of inclusion/exclusion