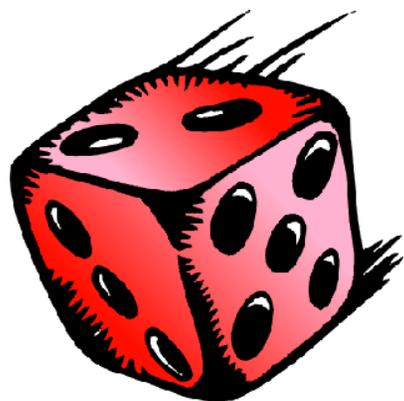

3. Discrete Probability



CSE 312
Autumn 2013
W.L. Ruzzo



Probability theory:

“an aberration of the intellect”

and

“ignorance coined into science”

– John Stuart Mill

Change to Syllabus

Introducing: The Daily Puzzler

One short problem

Due at start of lecture, MWF

Lightly Graded: 😊 😐 or 😞

Electronic turnin only

Due 10/4: What is the probability of getting
“2 pairs” when dealt 5 cards in poker?

(will be on web later today)

Sample space: S is the set of all possible outcomes of an experiment (often Ω in text books—Greek uppercase omega)

Coin flip: $S = \{\text{Heads, Tails}\}$

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: $S = \{x : x \in \mathbb{Z}, x \geq 0\}$

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

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

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

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

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

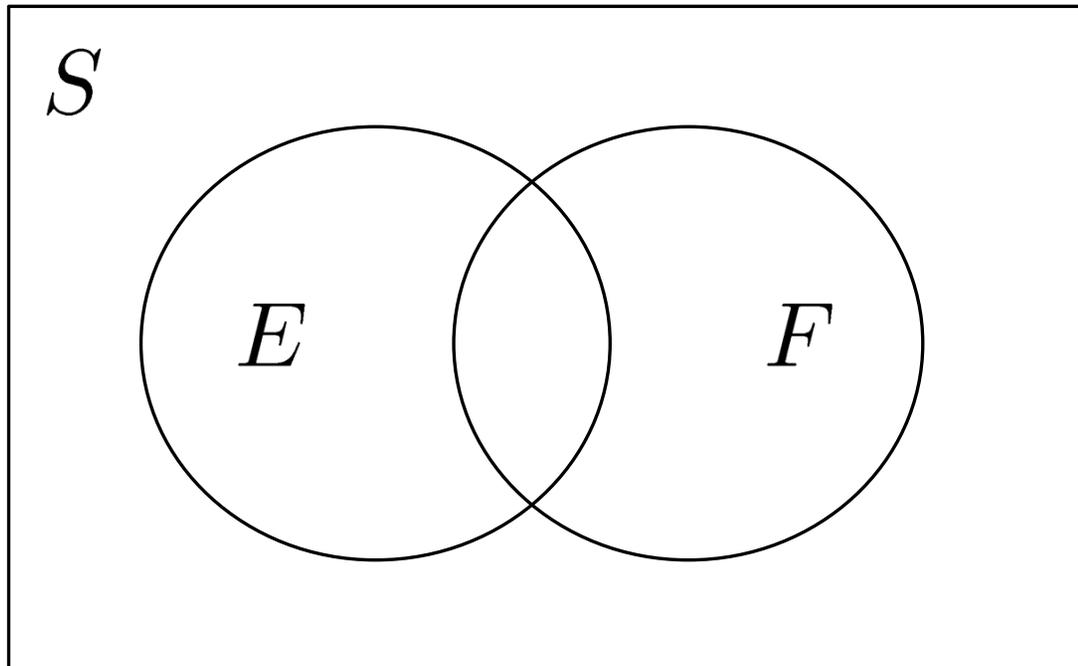
emails in a day < 20 : $E = \{x : x \in \mathbb{Z}, 0 \leq x < 20\}$

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

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

set operations on events

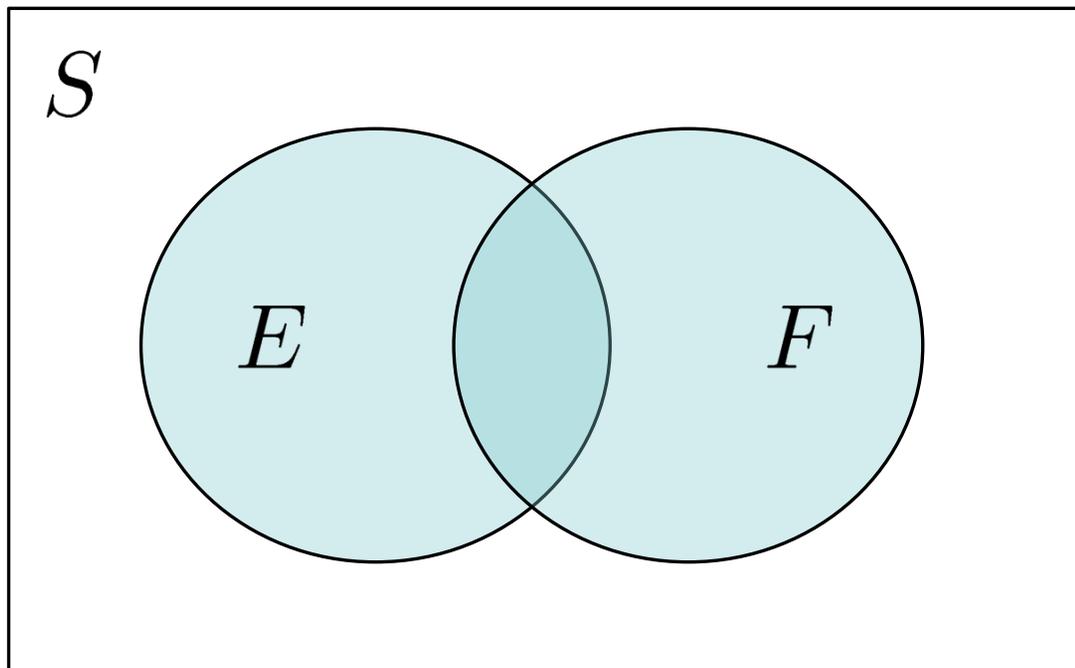
E and F are events in the sample space S



set operations on events

E and F are events in the sample space S

Event “ E OR F ”, written $E \cup F$



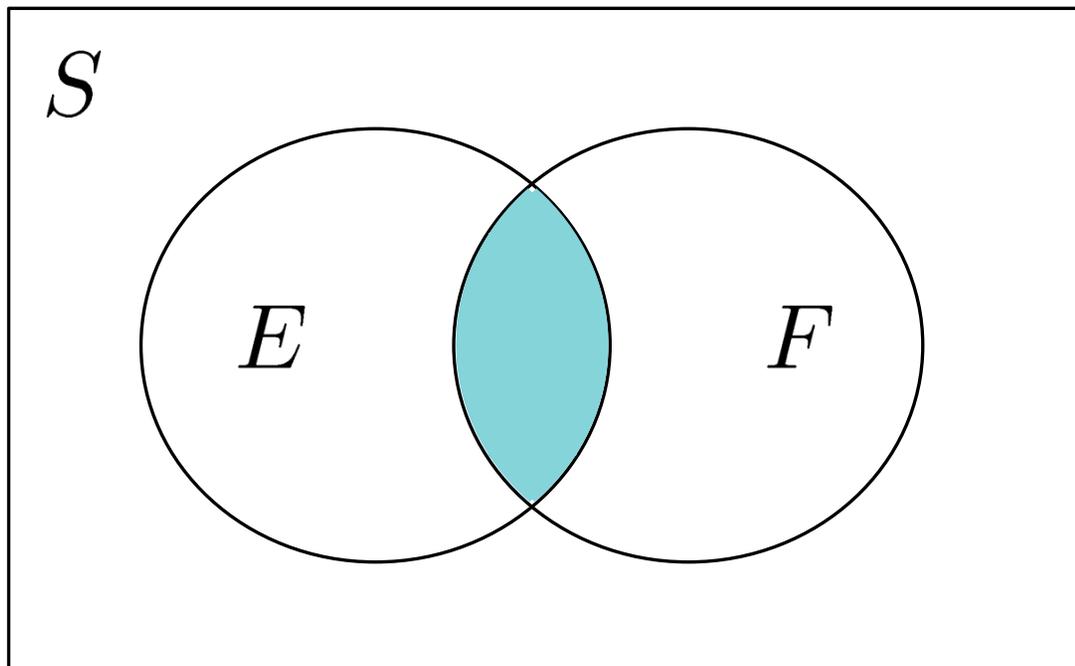
$S = \{1,2,3,4,5,6\}$
outcome of one die roll

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

set operations on events

E and F are events in the sample space S

Event “ E AND F ”, written $E \cap F$ or EF



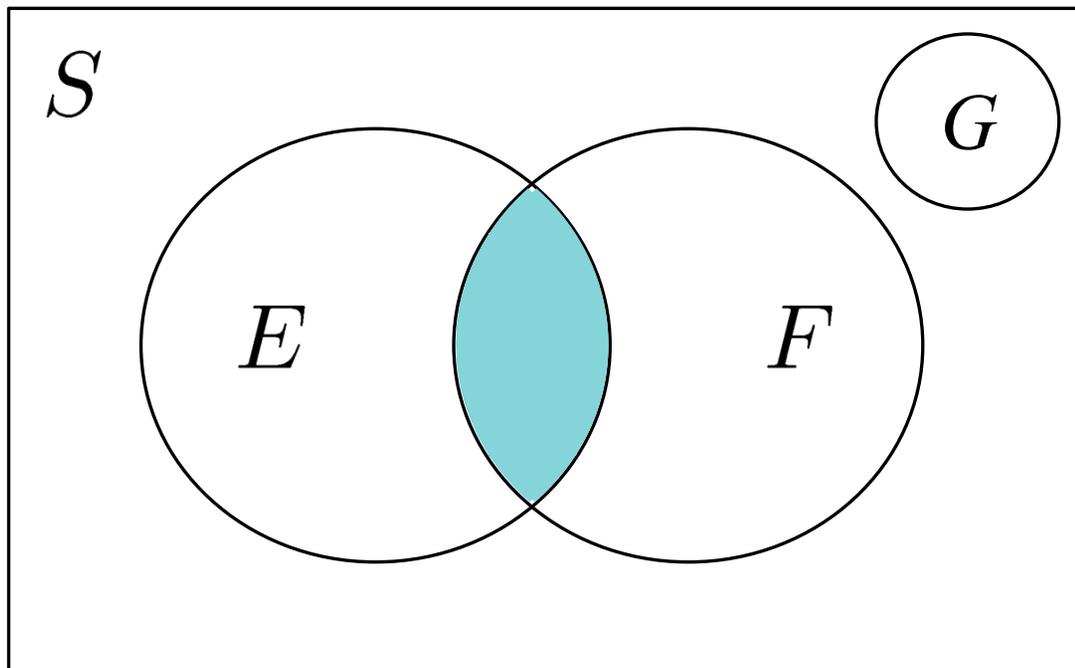
$S = \{1,2,3,4,5,6\}$
outcome of one die roll

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

set operations on events

E and F are events in the sample space S

$EF = \emptyset \Leftrightarrow E, F$ are “mutually exclusive”



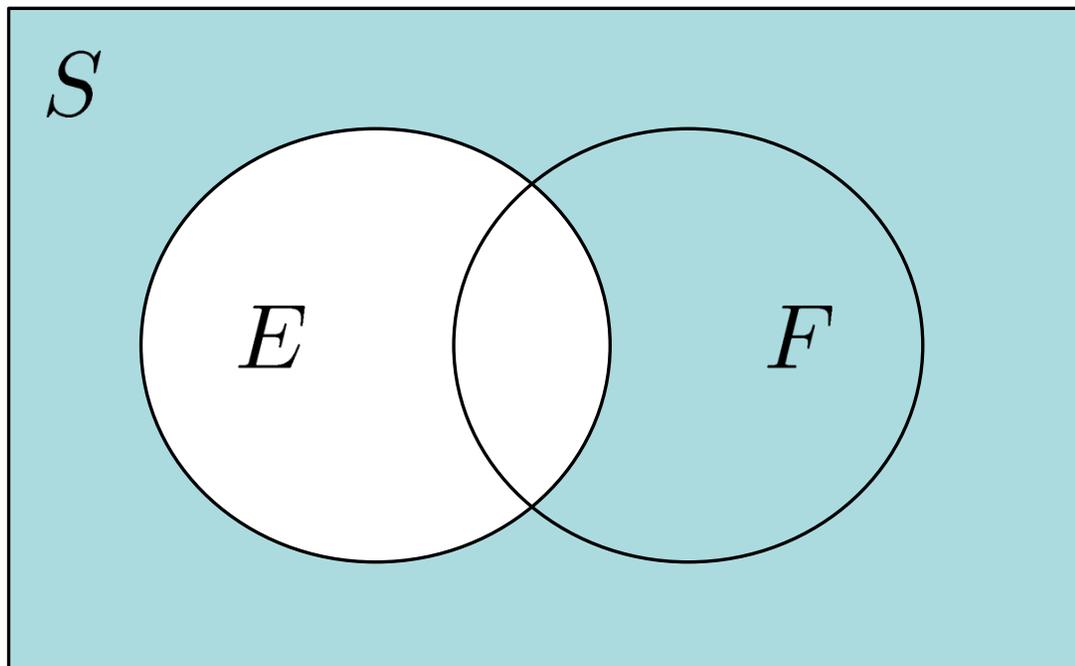
$S = \{1, 2, 3, 4, 5, 6\}$
outcome of one die roll

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

set operations on events

E and F are events in the sample space S

Event “not E ,” written \bar{E} or $\neg E$

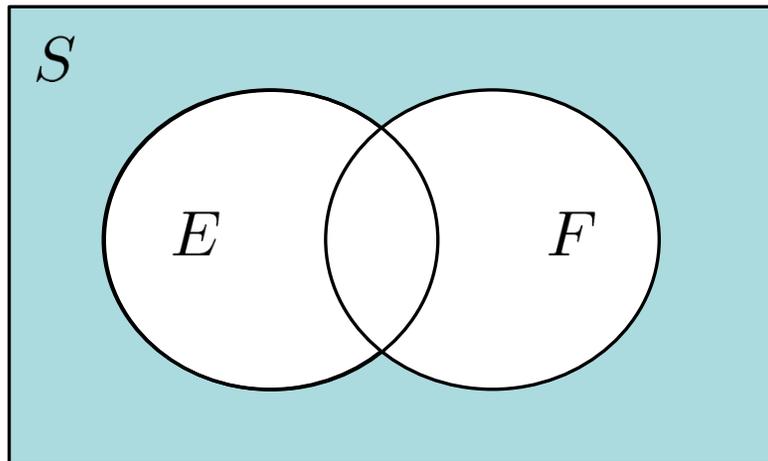


$S = \{1, 2, 3, 4, 5, 6\}$
outcome of one die roll

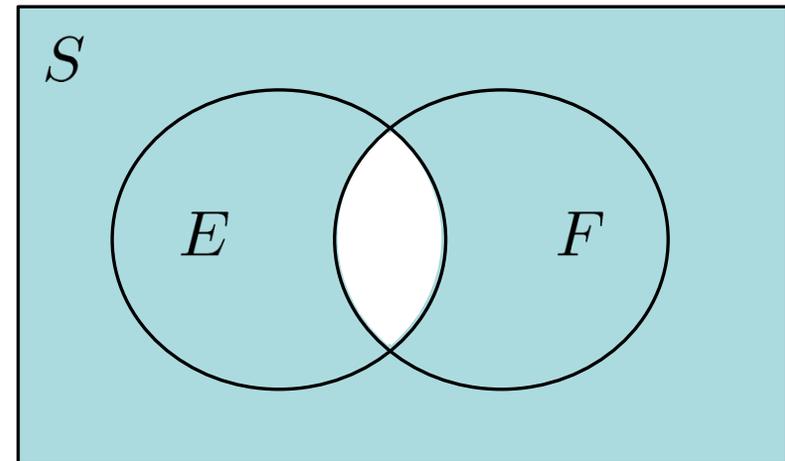
$E = \{1, 2\}$ $\neg E = \{3, 4, 5, 6\}$

DeMorgan's Laws

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



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



axioms of probability

Intuition: Probability as the relative frequency of an event

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

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, \dots, E_n of mutually exclusive events,

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

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

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

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

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

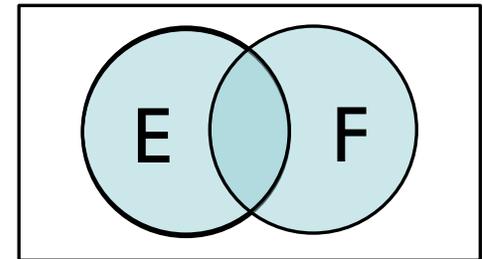
$$\Pr(E \cup F) = \Pr(E) + \Pr(F) - \Pr(EF)$$

inclusion-exclusion

$$\Pr(E) \leq 1$$

exercise

And many others



equally likely outcomes

Simplest case: sample spaces with equally likely outcomes.

Coin flips: $S = \{\text{Heads, Tails}\}$

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

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

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

In that case,

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

Why? Axiom 3 plus fact that $E = \text{union of singletons in } E$

rolling two dice

Roll two 6-sided dice. What is $\Pr(\text{sum of dice} = 7)$?

$$S = \{ (1,1), (1,2), (1,3), (1,4), (1,5), (1,6), \\ (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), \\ (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), \\ (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), \\ (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), \\ (6,1), (6,2), (6,3), (6,4), (6,5), (6,6) \}$$

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

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

$$\Pr(\text{sum} = 7) = |E|/|S| = 6/36 = 1/6.$$

twinkies and ding dongs



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 triple with 3 of 7 distinguishable objects)

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

Unordered: (S unordered triple 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

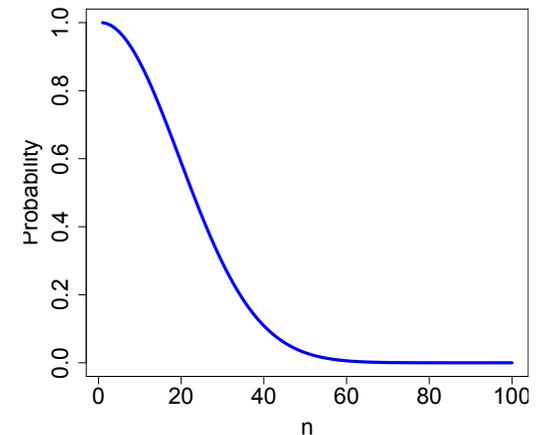


What is the probability that, of n people, none share the same birthday?

$$|S| = (365)^n$$

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

$$\begin{aligned}\text{Pr}(\text{no matching birthdays}) &= |E|/|S| \\ &= (365)(364)\cdots(365-n+1)/(365)^n\end{aligned}$$



Some values of n ...

$$n = 23: \text{Pr}(\text{no matching birthdays}) < 0.5$$

$$n = 77: \text{Pr}(\text{no matching birthdays}) < 1/5000$$

$$n = 100: \text{Pr}(\text{no matching birthdays}) < 1/3,000,000$$

$$n = 150: \text{Pr}(\dots) < 1/3,000,000,000,000,000$$

$n = 366?$

$Pr = 0$

Above formula gives this, since

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

when $n = 366$ (or greater).

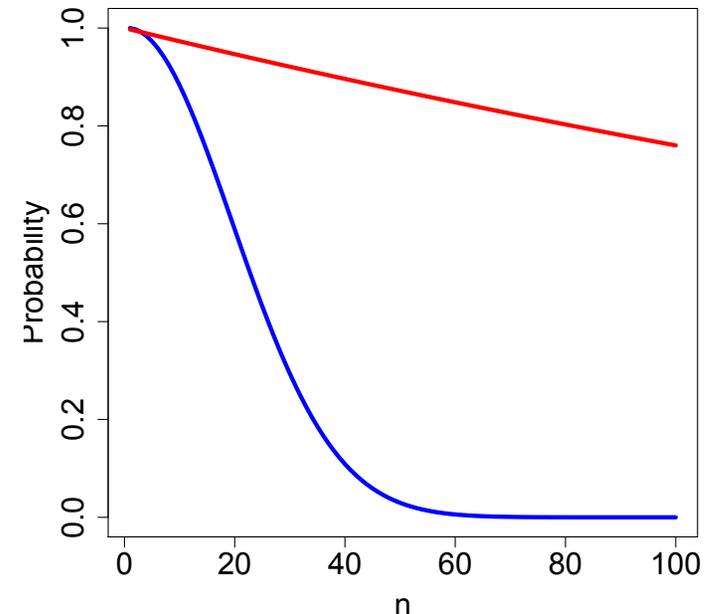
Even easier to see via pigeon hole principle.

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

$$|S| = (365)^n$$

$$|E| = (364)^n$$

$$\begin{aligned} \text{Pr}(\text{no birthdays} = \text{yours}) \\ = |E|/|S| = (364)^n/(365)^n \end{aligned}$$



Some values of n ...

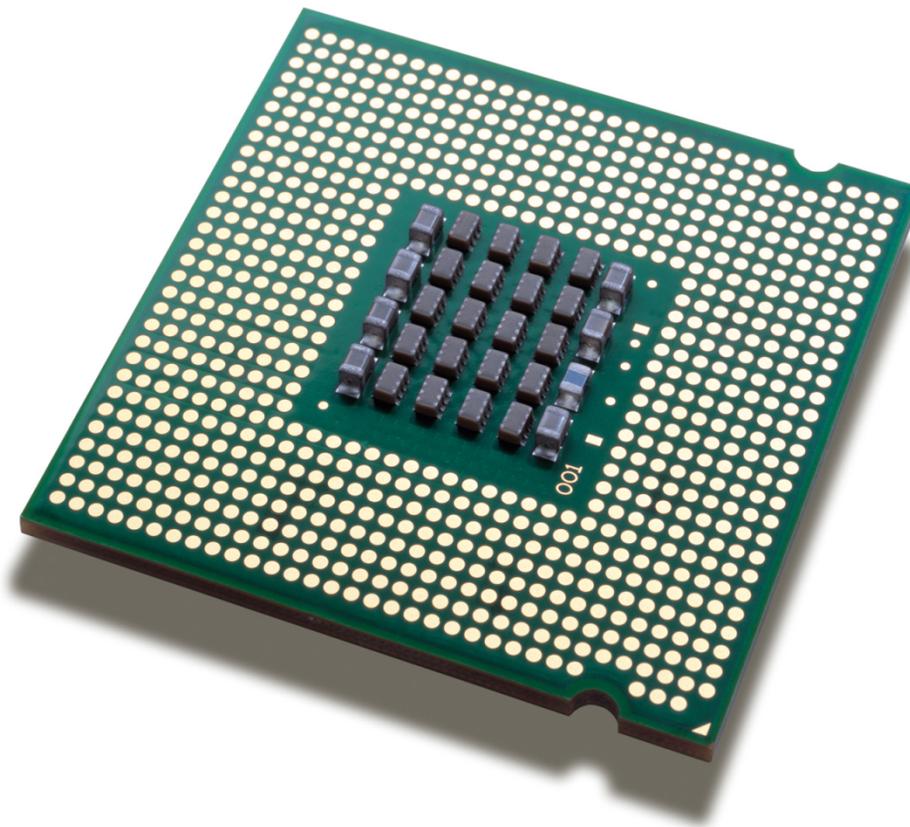
$$n = 23: \quad \text{Pr}(\text{no matching birthdays}) \approx 0.9388$$

$$n = 77: \quad \text{Pr}(\text{no matching birthdays}) \approx 0.8096$$

$$n = 253: \quad \text{Pr}(\text{no matching birthdays}) \approx 0.4995$$

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

chip defect detection



chip defect detection

n chips manufactured, one of which is defective
k chips randomly selected from n for testing

What is $\Pr(\text{defective chip is in } k \text{ selected chips})$?

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

$\Pr(\text{defective chip is among } k \text{ selected chips})$

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

n chips manufactured, one of which is defective
k chips randomly selected from n for testing

What is $\Pr(\text{defective chip is in } k \text{ selected chips})$?

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} chip is defective.
- Events E_1, E_2, \dots, E_k are mutually exclusive
- $\Pr(E_i) = 1/n$ for $i=1,2,\dots,k$
- Thus $\Pr(\text{defective chip is selected})$
 $= \Pr(E_1) + \dots + \Pr(E_k) = k/n.$

chip defect detection

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)** ?

$$\begin{aligned} |S| &= \binom{n}{k} & |E| &= (\text{1 chip defective}) + (\text{2 chips defective}) \\ & & &= \binom{2}{1} \binom{n-2}{k-1} + \binom{2}{2} \binom{n-2}{k-2} \end{aligned}$$

Pr(a defective chip is in k selected chips)

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

chip defect detection

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):

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

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

(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:

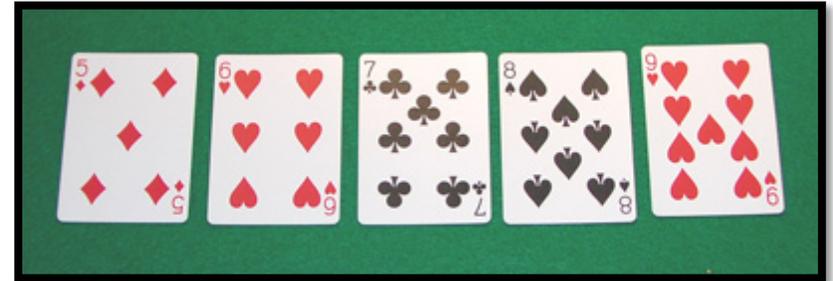
A♥	2♥	3♥	...	10♥	J♥	Q♥	K♥
A♣	2♣	3♣	...	10♣	J♣	Q♣	K♣
A♦	2♦	3♦	...	10♦	J♦	Q♦	K♦
A♠	2♠	3♠	...	10♠	J♠	Q♠	K♠

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

any straight in poker

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 $\Pr(\text{straight})$?

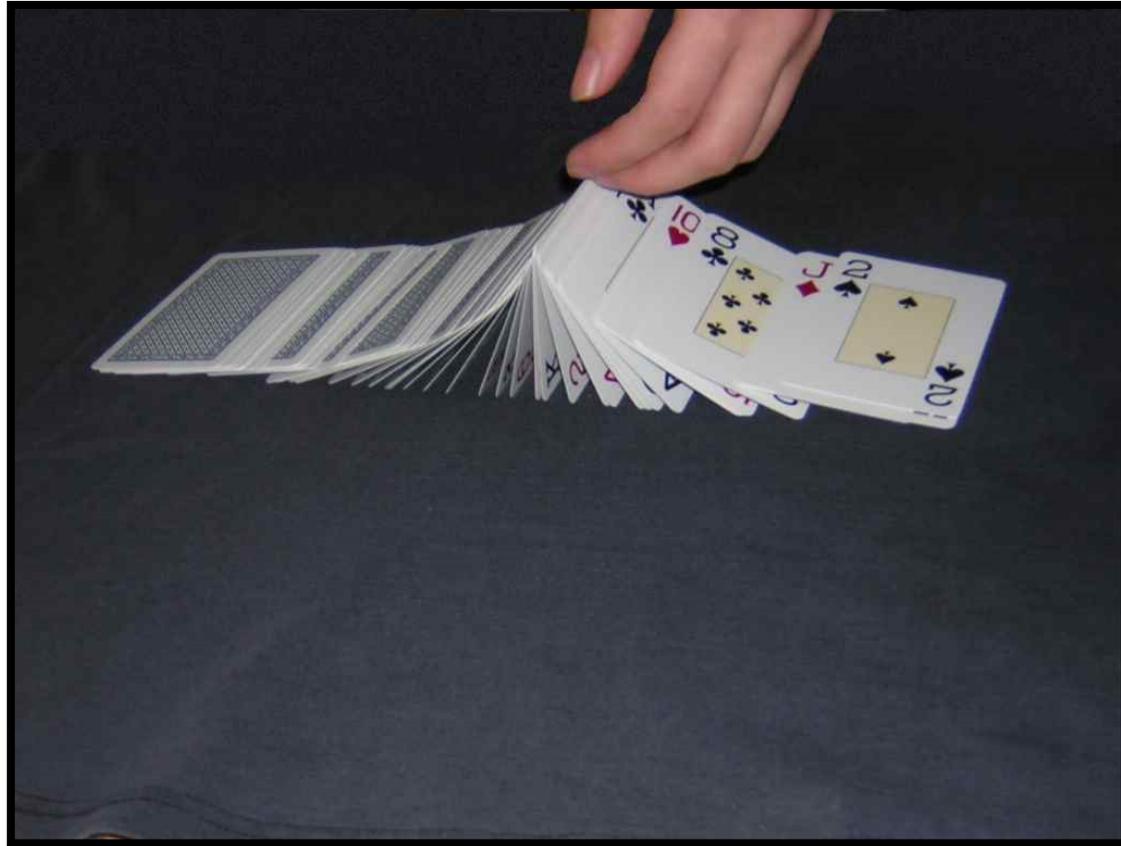
S as on previous slide, $|S| = \binom{52}{5}$

What's E?

E = Pick a col A, 2, ... 10, then 1 of 4 in next 5 cols (wrapping K→A)

$$|E| = 10 \cdot \binom{4}{1}^5 \quad \Pr(\text{straight}) = \frac{10 \binom{4}{1}^5}{\binom{52}{5}} \approx 0.00394$$

card flipping



52 card deck. Cards flipped one at a time.

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

$\Pr(\text{next card} = \text{ace of spades}) < \Pr(\text{next card} = 2 \text{ of clubs}) ?$

Maybe, Maybe Not ...

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

Event 1: Next = Ace of Spades.

Remove A_{\spadesuit} , shuffle remaining 51 cards, add A_{\spadesuit} after first Ace

$|E_1| = 51!$ (only 1 place A_{\spadesuit} can be added)

Event 2: Next = 2 of Clubs

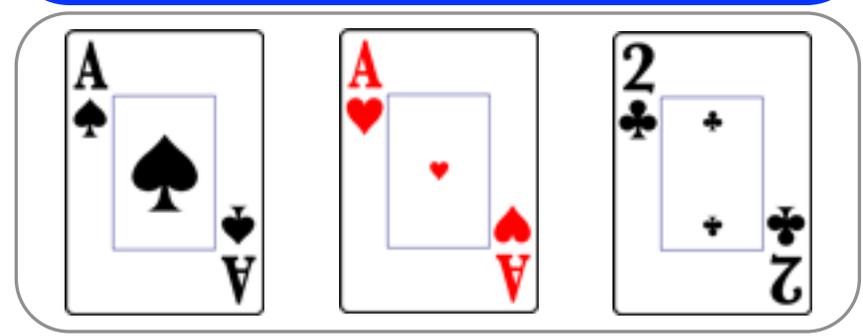
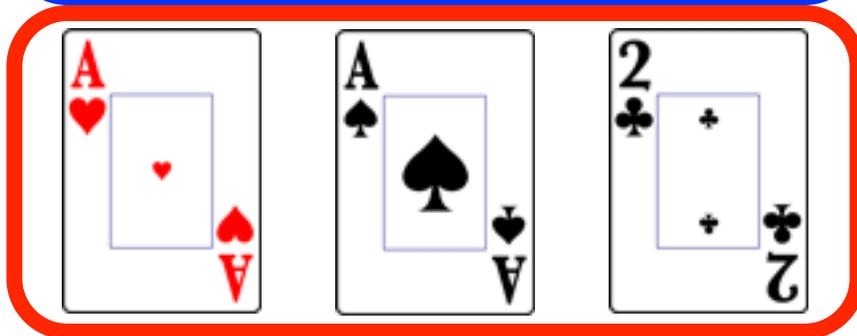
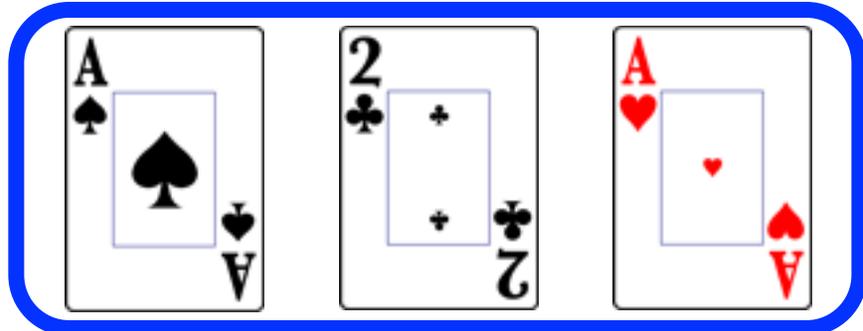
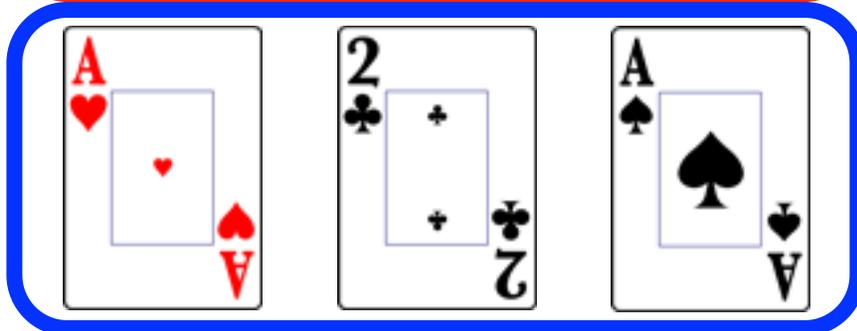
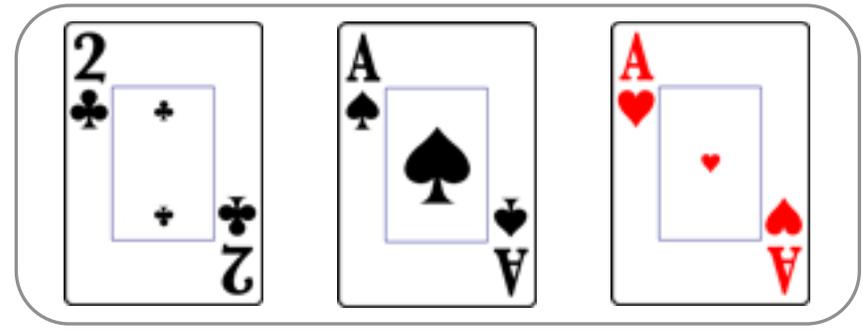
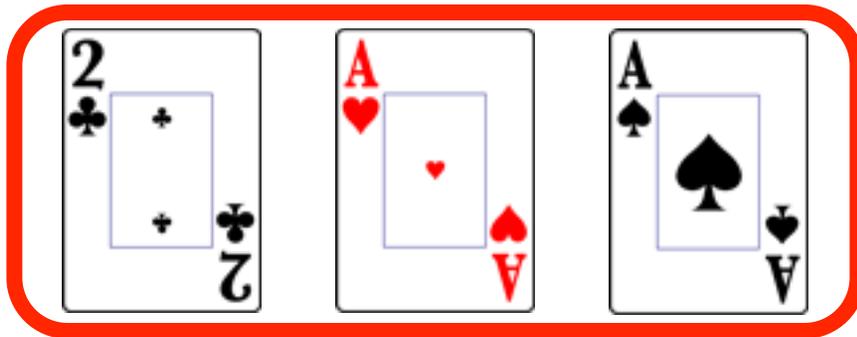
Do the same thing with 2_{\clubsuit} ; E_1 and E_2 have same size

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 middle, select at random. What is $\Pr(\text{no one gets own hat})$?



$\Pr(\text{no one gets own hat}) =$
 $1 - \Pr(\text{someone gets own hat})$

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

$\Pr(\bigcup_{i=1}^n E_i) = \sum_i P(E_i) - \sum_{i < j} \Pr(E_i E_j) + \sum_{i < j < k} \Pr(E_i E_j E_k) \dots$

Visualizing the sample space S :

People:

P_1	P_2	P_3	P_4	P_5
H_4	H_2	H_5	H_1	H_3

Hats:



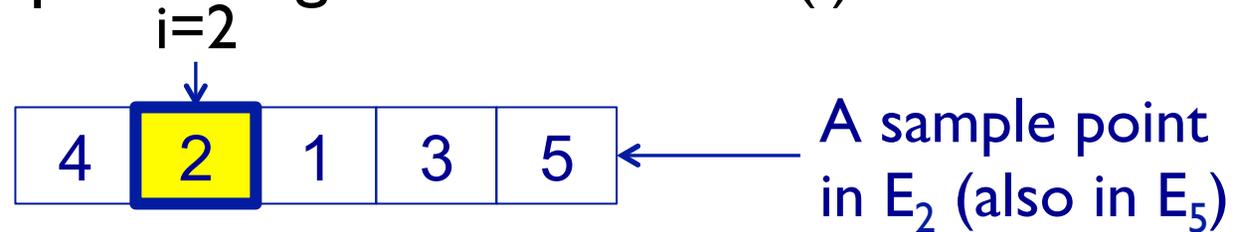
I.e., a sample point is a *permutation* π of $1, \dots, n$

4	2	5	1	3
---	---	---	---	---

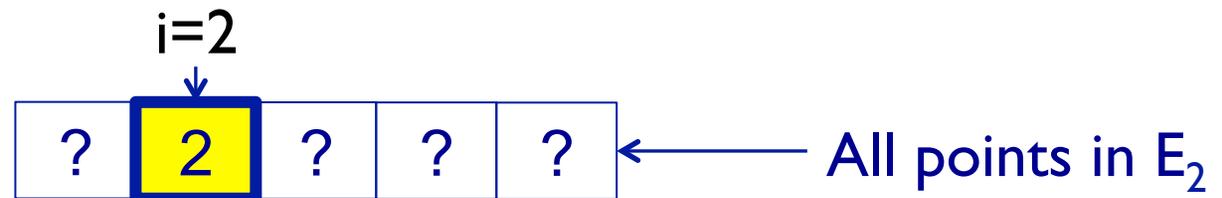
$$|S| = n!$$

hats: events

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



Counting single events:

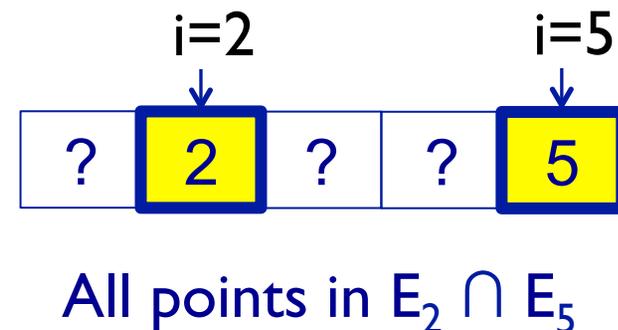


$$|E_i| = (n-1)! \text{ for all } i$$

Counting pairs:

$$E_i E_j : \pi(i) = i \text{ \& } \pi(j) = j$$

$$|E_i E_j| = (n-2)! \text{ for all } i, j$$



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



E_i = event that person i gets own hat

$$\Pr(\bigcup_{i=1}^n E_i) = \sum_i P(E_i) - \sum_{i<j} \Pr(E_i E_j) + \sum_{i<j<k} \Pr(E_i E_j E_k) \dots$$

$$\Pr(k \text{ fixed people get own back}) = (n-k)!/n!$$

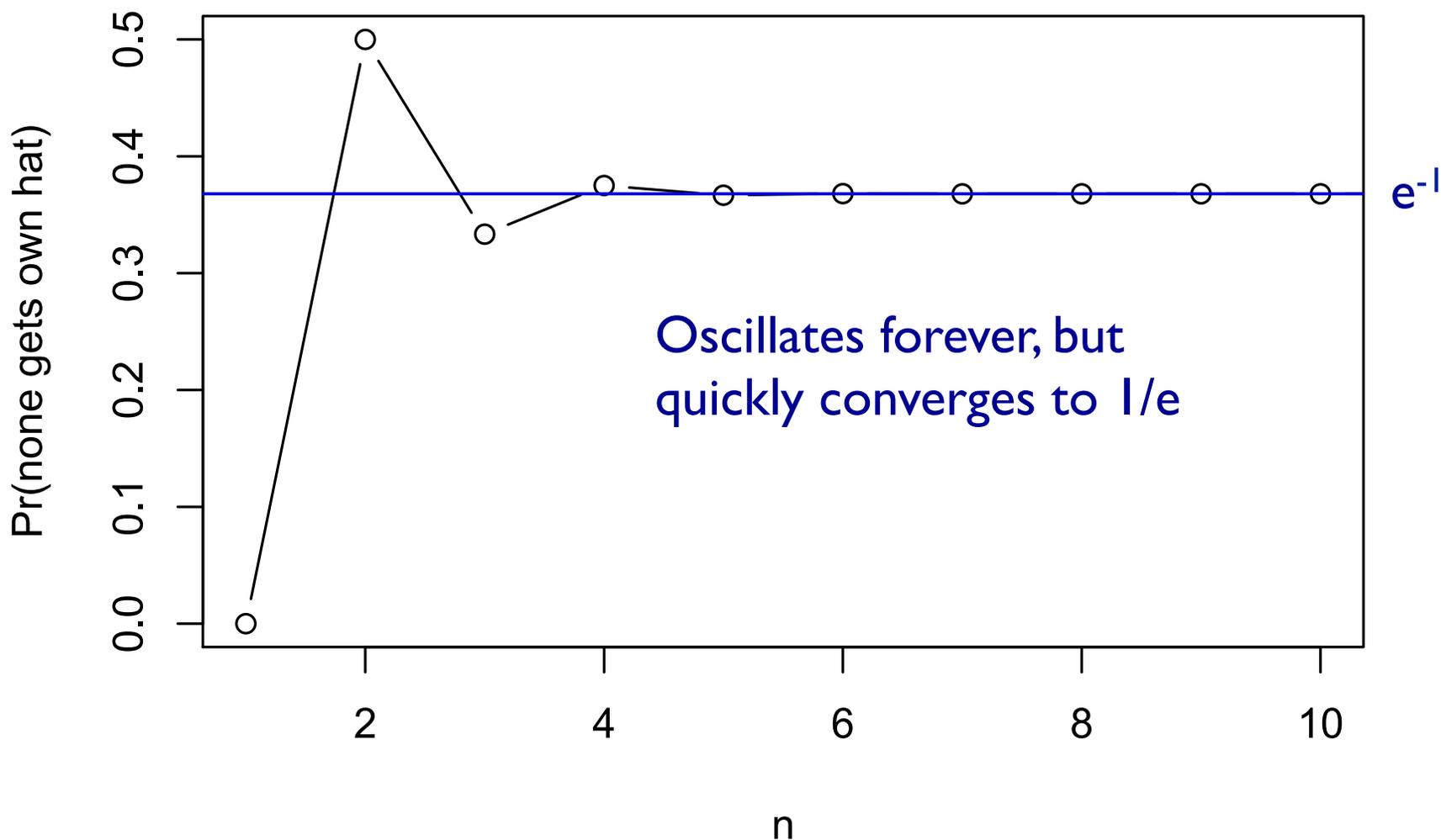
$$\binom{n}{k} \text{ times that} = \frac{n!}{k!(n-k)!} \frac{(n-k)!}{n!} = 1/k!$$

$$\Pr(\text{none get own}) = 1 - \Pr(\text{some do}) =$$

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

$\Pr(\text{none get own}) = 1 - \Pr(\text{some do}) =$

$$1 - \left(1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} \dots + \frac{(-1)^n}{n!} \right) \approx e^{-1} \approx .37$$



Sample spaces } Visualize!
Events

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