Discrete probability

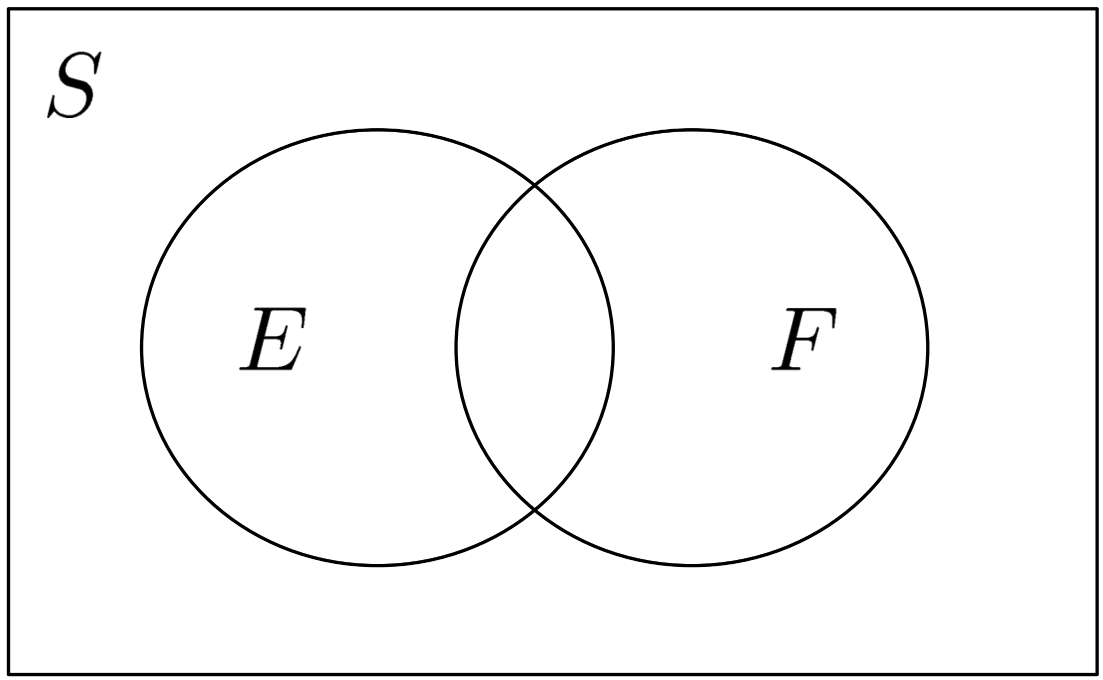
Sample space: S is the set of all possible outcomes of an experiment (Ω in your text book–Greek uppercase omega)

Coin flip:	$S = \{\text{Heads}, \text{Tails}\}$
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:	$S = \{ x : x \in R, 0 \leq x \leq 24 \}$

Events: $\mathbf{E} \subseteq \mathbf{S}$ is some subset of the sample space

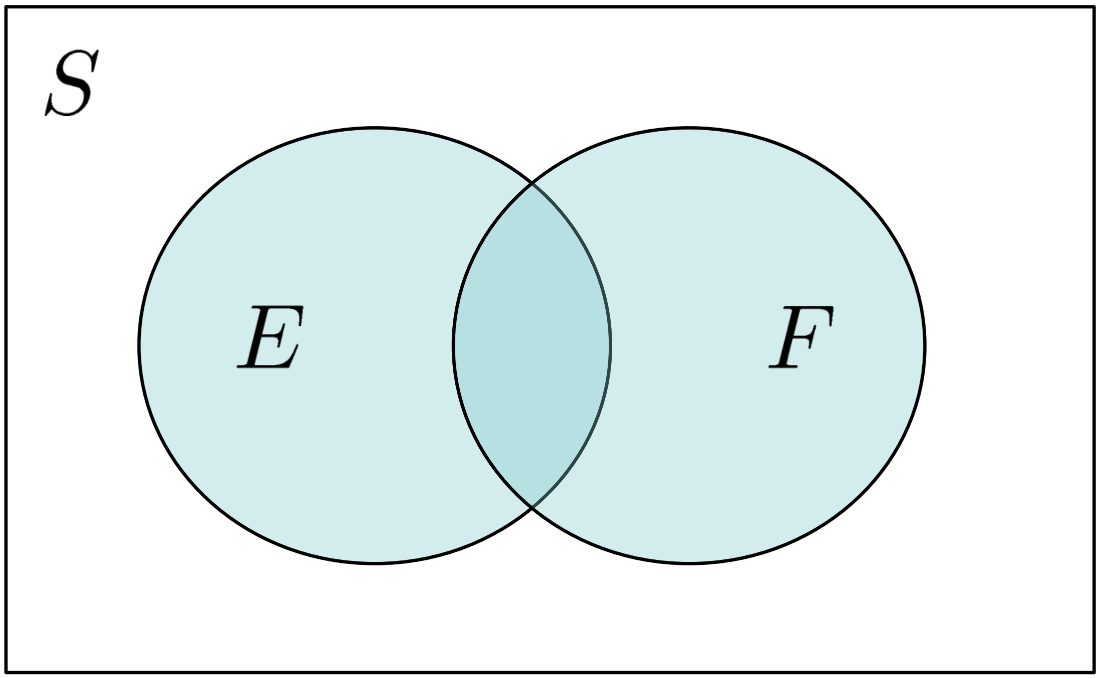
Coin flip is heads:	E = {Head}
At least one head in 2 flips:	E = {(H,H), (H,T), (T,H)}
Roll of die is 3 or less:	E = {1, 2, 3}
# emails in a day < 20:	$E = \{ x : x \in Z,\ 0 \leq x < 20\}$
Wasted day (>5 YT hrs):	$E = \{ x : x \in R,\ x > 5 \}$

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 $E \cup F$

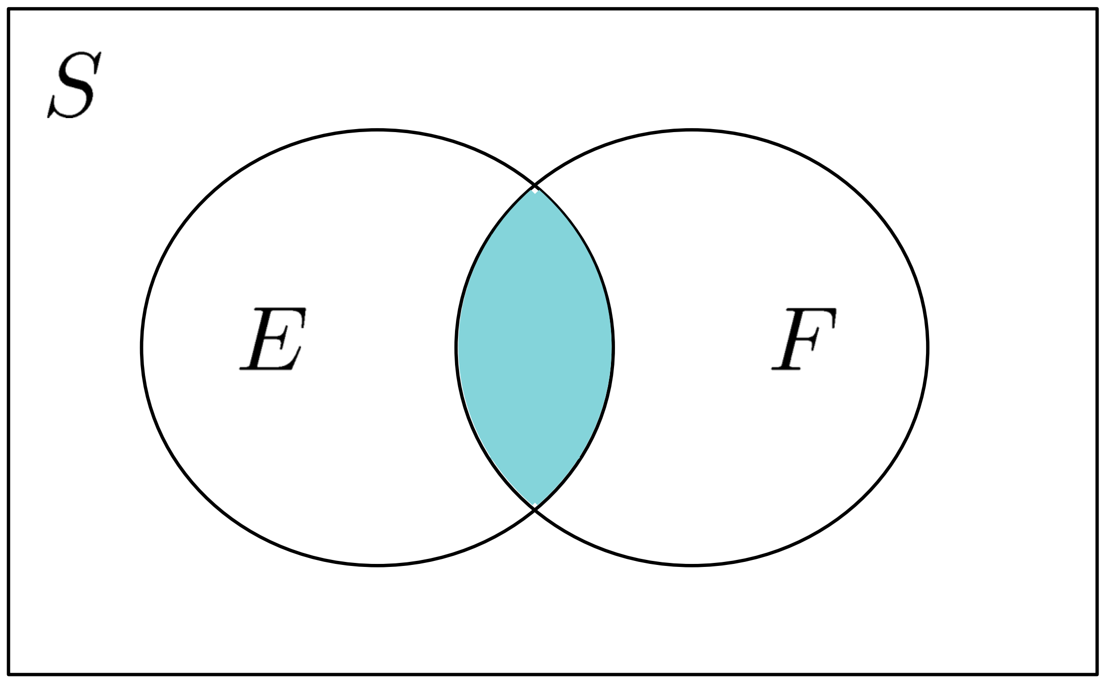


$S = \{1,2,3,4,5,6\}$
outcome of one die roll

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

E and F are events in the sample space S

Event "E AND F", written $E \cap F$ or EF

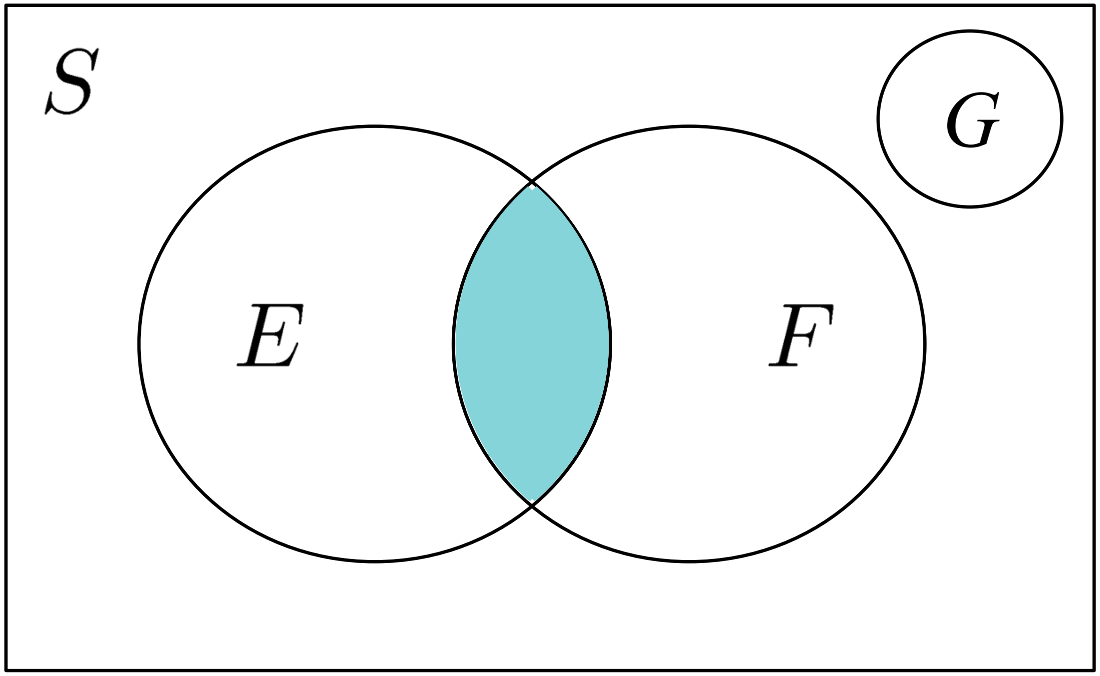


S = {1,2,3,4,5,6}
outcome of one die roll

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

E and F are events in the sample space S

EF = ∅ ⇔ E,F are "mutually exclusive"

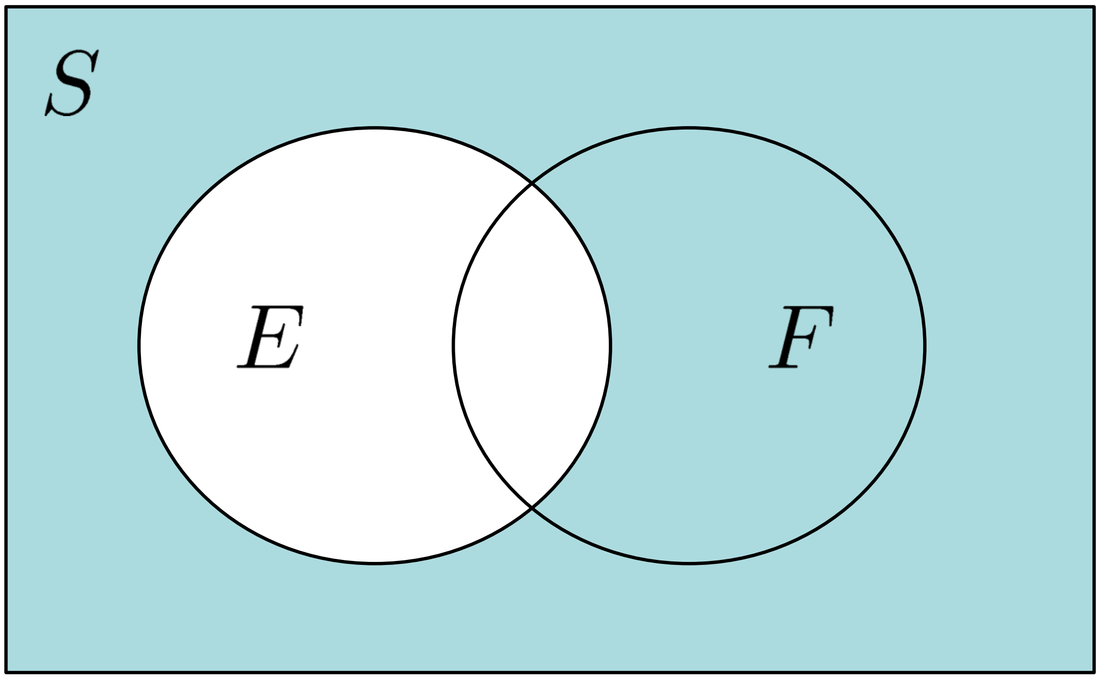


S = {1,2,3,4,5,6}
outcome of one die roll

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 $\neg E$

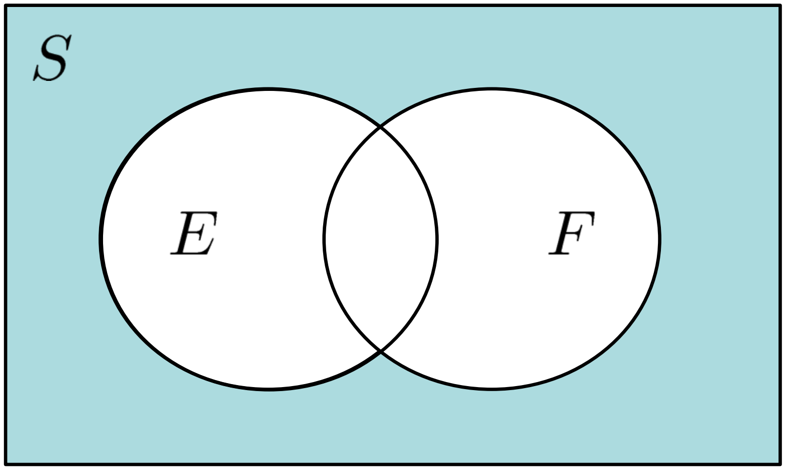


S = {1,2,3,4,5,6}
outcome of one die roll

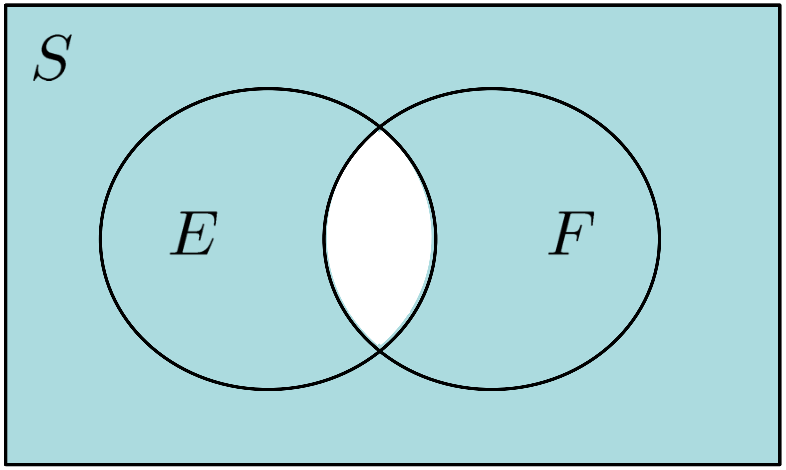
$E = \{1, 2\}$ $\neg E = \{3, 4, 5, 6\}$

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

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

Axiom 1: $0 \leq \Pr(E) \leq 1$

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

Axiom 3: If E and F are mutually exclusive ($EF = \varnothing$), then

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

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

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

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

$$\Pr(\bar{E}) = \Pr(S) - \Pr(E) \text{ because } S = E \cup \bar{E}$$

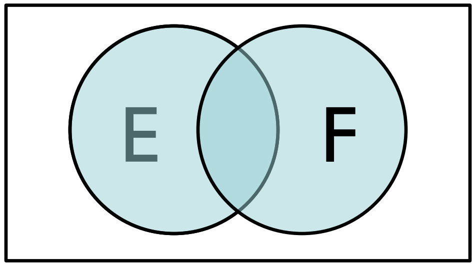
- If $E \subseteq F$, then $\Pr(E) \leq \Pr(F)$

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

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

inclusion-exclusion formula

- And many others



equally likely outcomes

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) = $\frac{1}{|S|}$

uniform distribution

In that case,

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

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) }

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

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

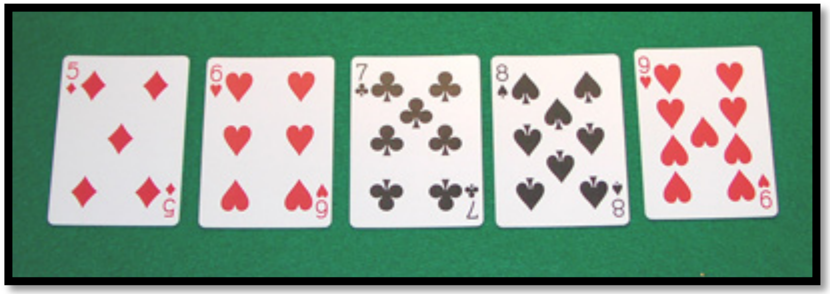
poker hands

any straight in poker

Consider 5 card poker hands.

A "straight" is 5 consecutive rank cards of any suit

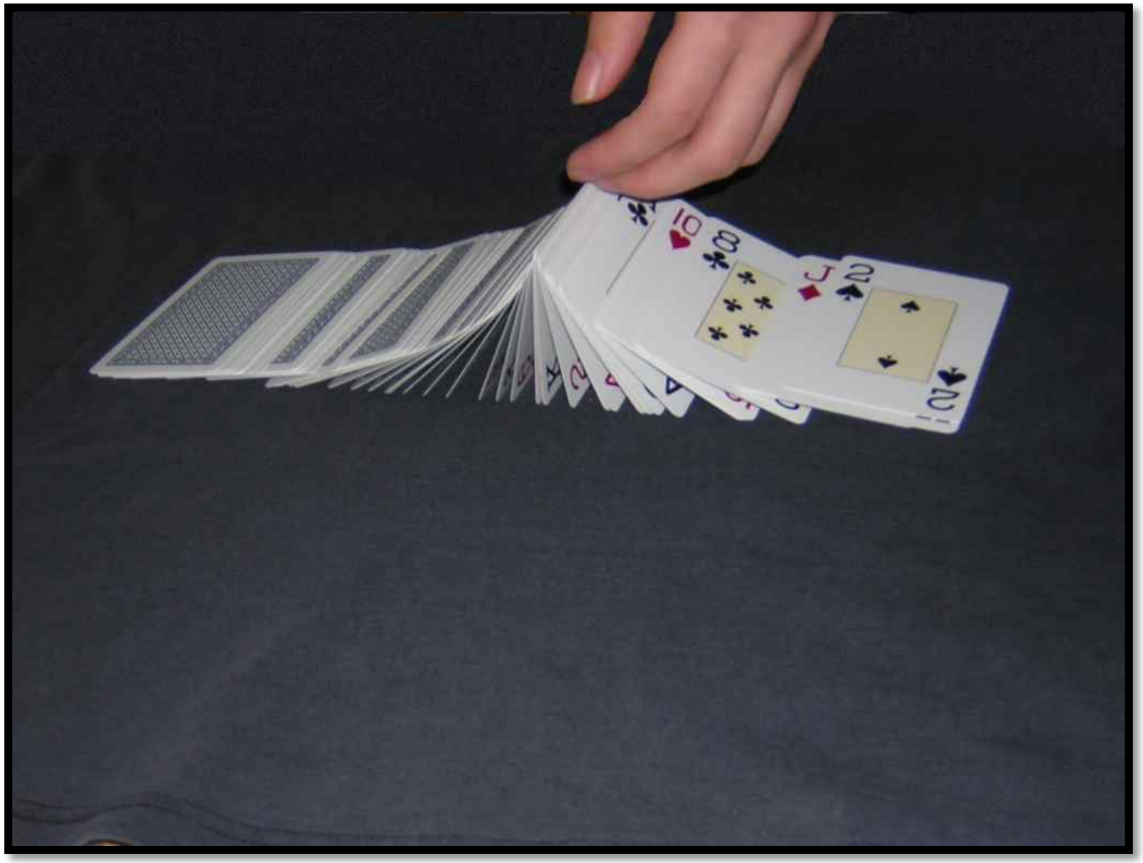
What is Pr(straight) ?



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

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

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



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) ?

Case 1: Take Ace of Spades out of deck

Shuffle remaining 51 cards, add ace of spades after first ace

|S| = 52! (all cards shuffled)

|E| = 51! (only 1 place ace of spades can be added)

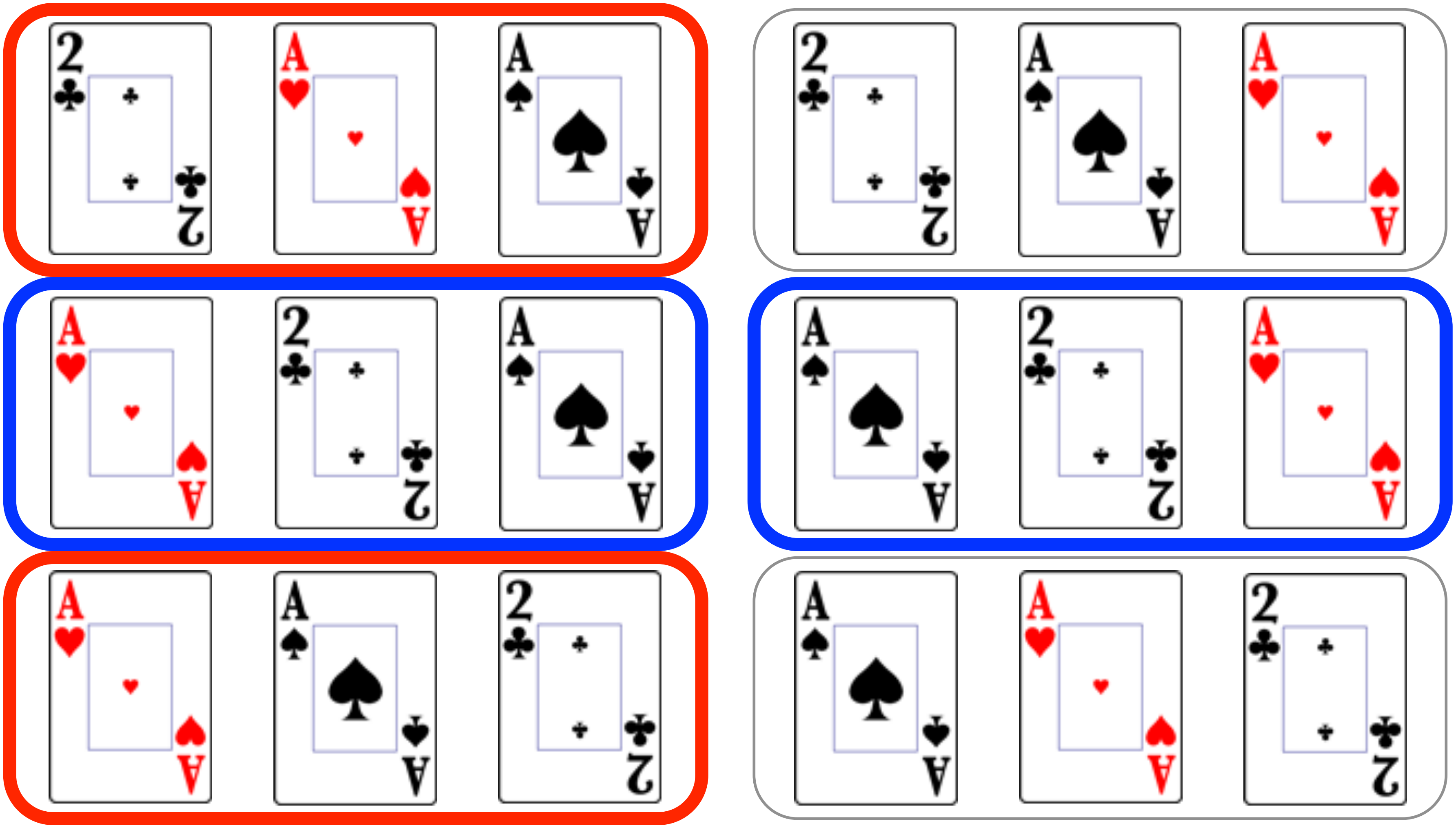
Case 2: Do the same thing with the 2 of clubs

|S| and |E| have same size

So,

Pr(next = Ace of spades) = Pr(next = 2 of clubs) = 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

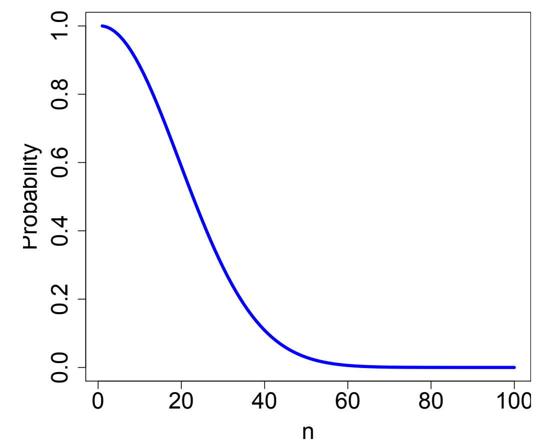
What is the probability that, of n people, none share the same birthday?

$|S| = (365)^n$

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

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

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



Some values of n…

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

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

n = 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 you?

$|S| = (365)^n$

$|E| = (364)^n$

Pr(no birthdays matches yours) = |E|/|S|

$= (364)^n/(365)^n$

Some values of n…

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

n = 77: Pr(no matching birthdays) ≈ 0.8096

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

hats

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

- What is the sample space?
- What is the probability of each outcome?
- What is the event of interest?

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 : set of all permutations where π_i is not equal to i for all i.

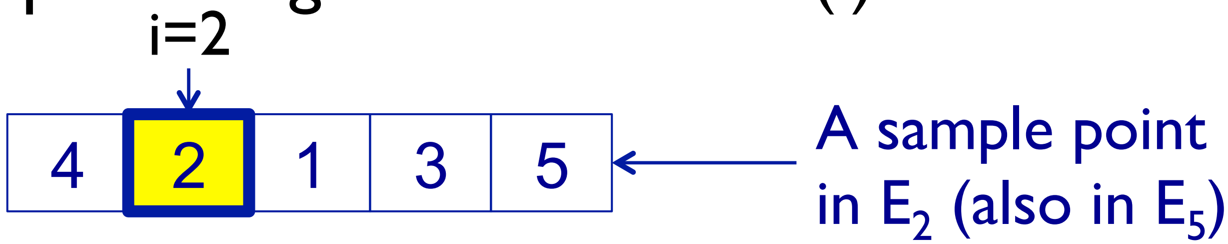
n persons at a party throw hats in middle, select at random. What is Pr(no one gets own hat)?

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

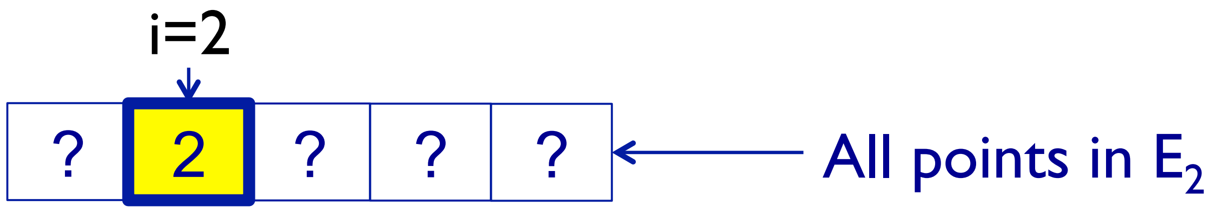
Pr(someone gets own hat) = $\Pr(\cup_{i=1}^{n} E_i)$, where
E_i = event that person i gets own hat

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

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



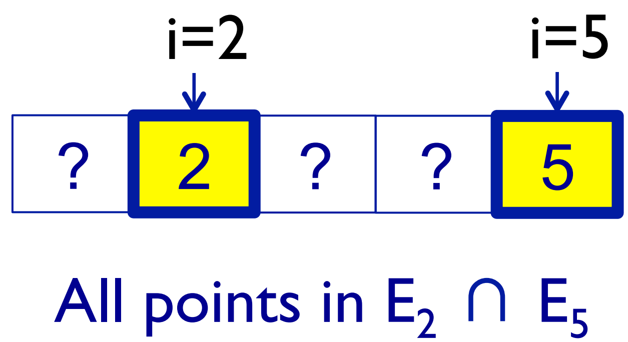
Counting single events:



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

Counting pairs:

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



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

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!

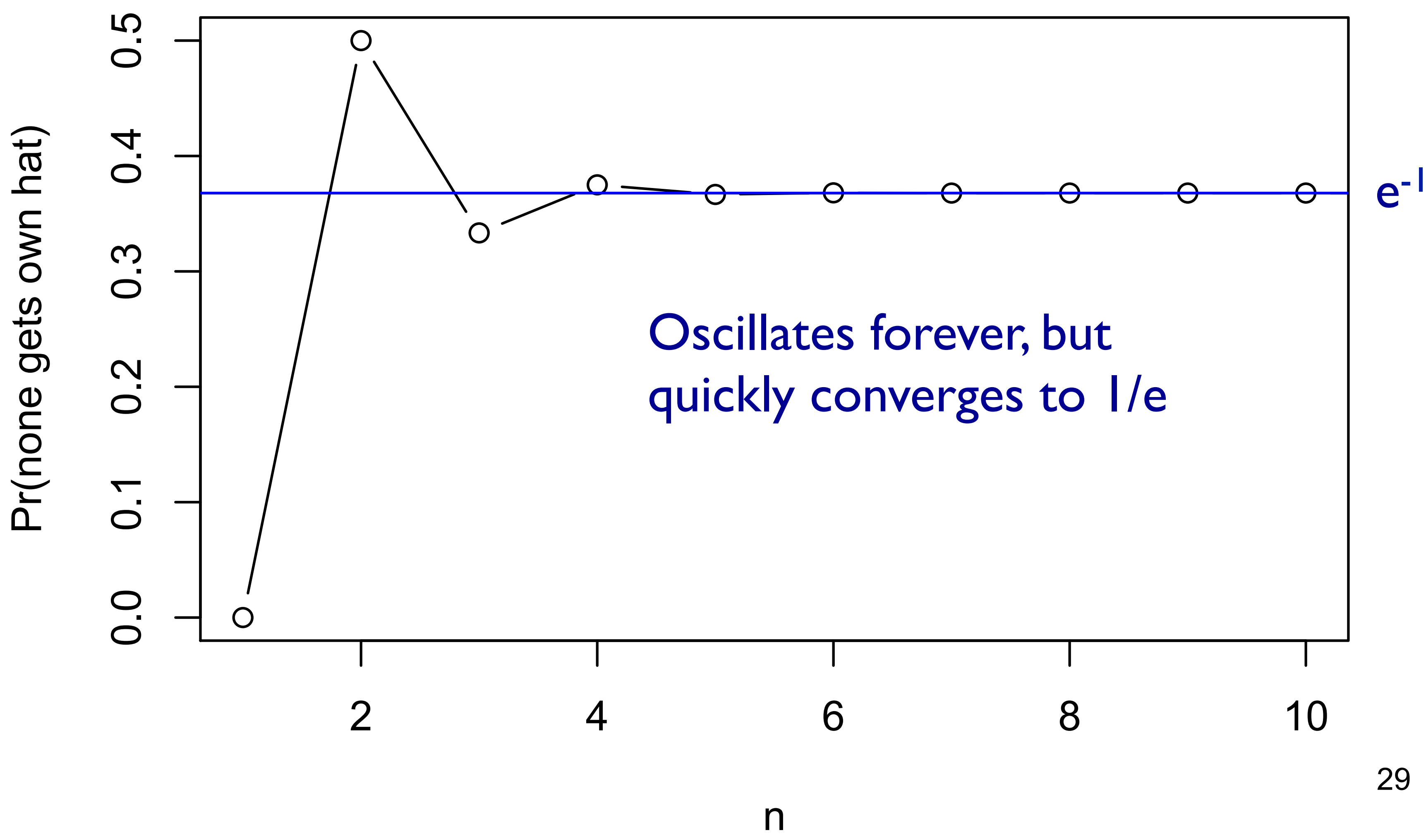
$$\binom{n}{k} \text{ times that } = \frac{n!}{k!(n-k)!} \frac{(n-k)!}{n!} = 1/k!$$

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) =

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

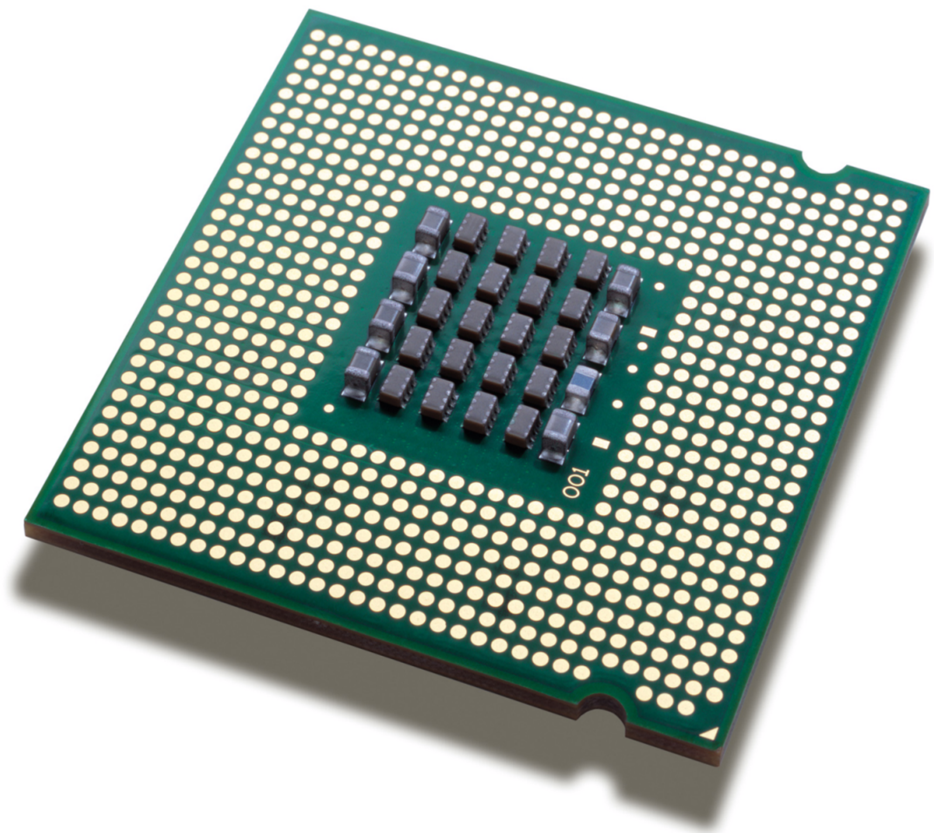


The Monty Hall Problem

- *Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the other, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?*

- Assumptions:
  - The car is equally likely to be behind each of the doors.
  - The player is equally likely to pick each of the three doors, regardless of the car's location
  - After the player picks a door, the host ***must*** open a different door with a goat behind it and offer the player the choice of staying with the original door or switching
  - If the host has a choice of which door to open, then he is equally likely to select each of them.

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 in k selected chips)

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

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} 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!)