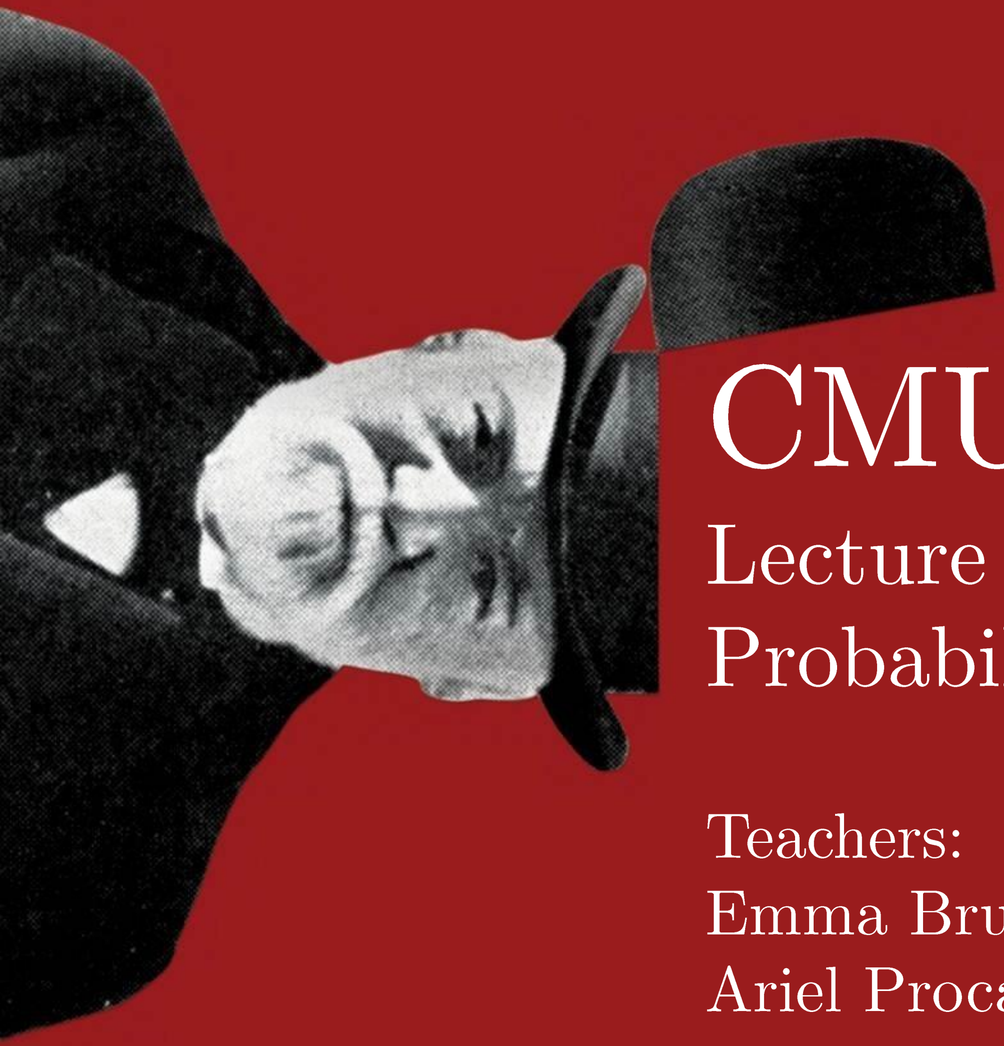


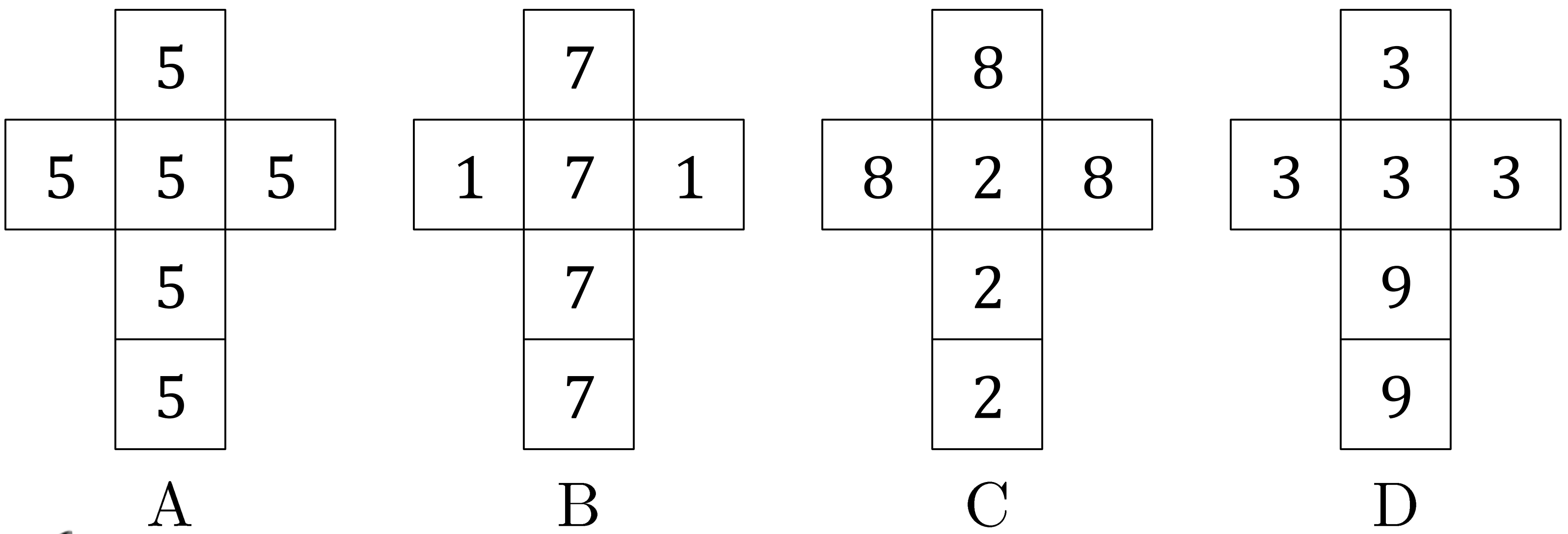
CMU 15-381

Lecture 7: Probability

Teachers:
Emma Brunskill
Ariel Procaccia (this time)

Gambling 101

- You choose a die first, I choose second
- We both throw; higher number wins
- Which die would you choose?



Gambling 101

- Antoine Gombaud (1607-1684) made history for being a loser

I will roll a die four times; I win if I get a 1

- After a while no one would take the bet
- $1-\left(\frac{5}{6}\right)^{4}=0.518$

Gambling 101

- Gombaud invented a new scam:

> I will roll two dice 24 times;
> I win if I get a double 1

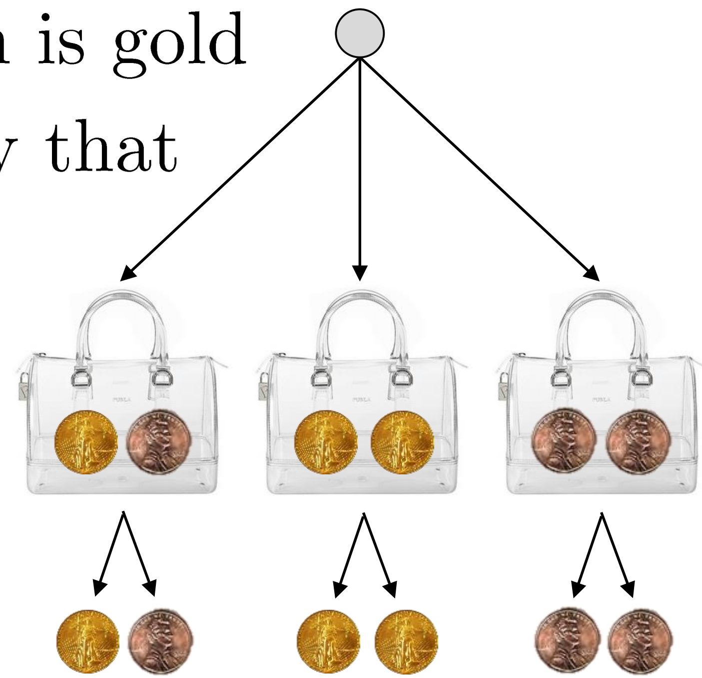
- Why was he losing money?
- $1 - \left(\frac{35}{36}\right)^{24} = 0.491$
- Gombaud wrote to Pascal and Fermat, who subsequently created probability theory

Gambling 101

Morale of the story:
Analyzing gambling is not a side-benefit of probability theory; probability theory was created to analyze gambling!

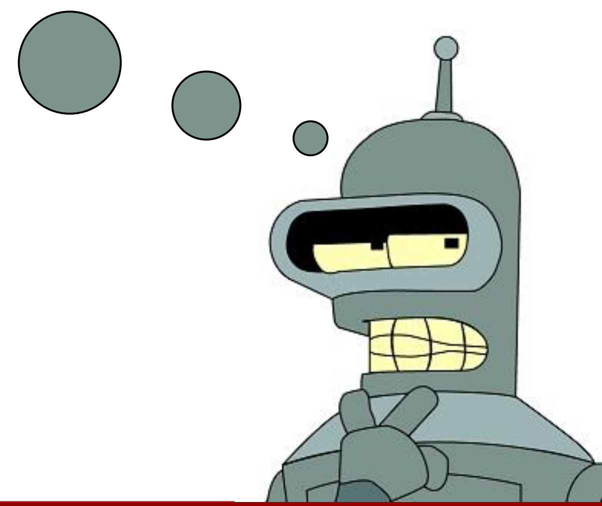
Pennies and gold

- Three bags contain two gold coins, two pennies, and one of each
- Bag is chosen at random, and one coin from it is selected at random; the coin is gold
- Poll 1: What is the probability that the other coin is gold?
    1. 1/6
    2. 1/3
    3. 2/3
    4. 1



Language of probability

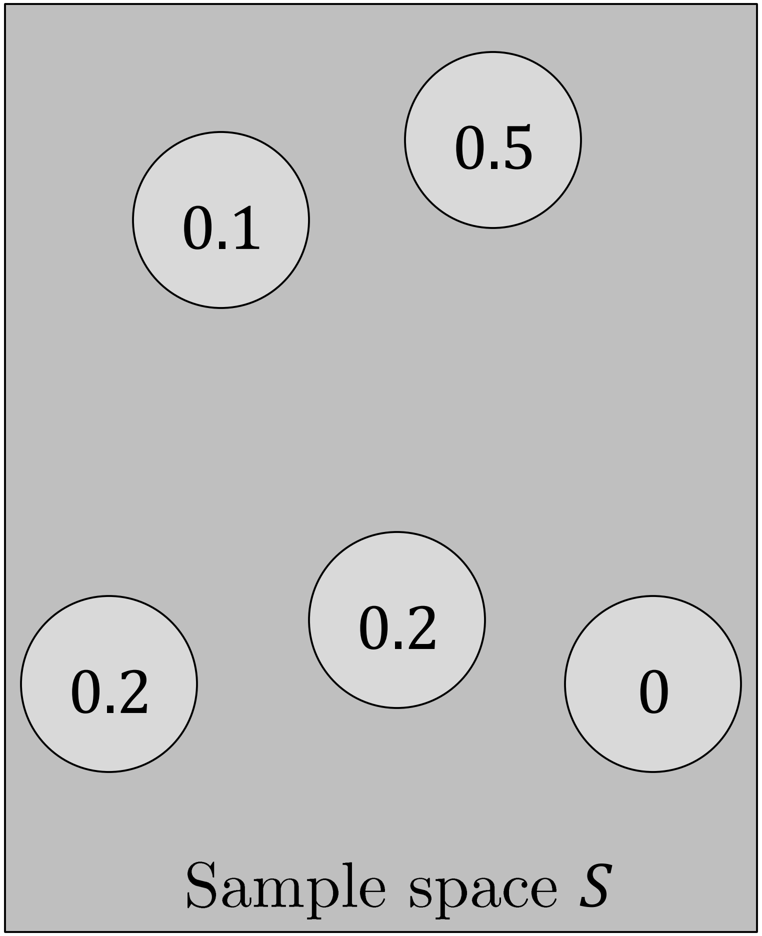
Probability can be counterintuitive; we need a formal language!



LANGUAGE OF PROBABILITY

- The sample space is a finite set of elements S
- A probability distribution p assigns a non-negative real probability to each element, such that

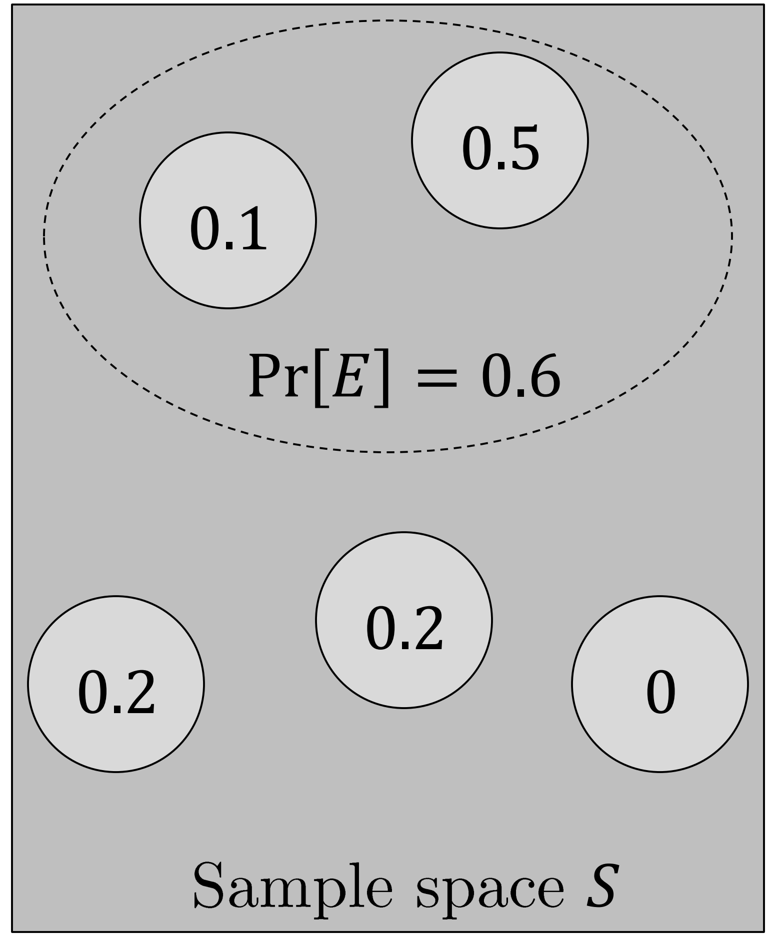
$$\sum_{x \in S} p(x) = 1$$



Language of probability

- An event is a subset $E \subseteq S$
- $\Pr[E] = \sum_{x \in E} p(x)$
- If each element $x \in S$ has equal probability, the distribution is uniform:

$$\Pr[E] = \sum_{x \in E} p(x) = \frac{|E|}{|S|}$$

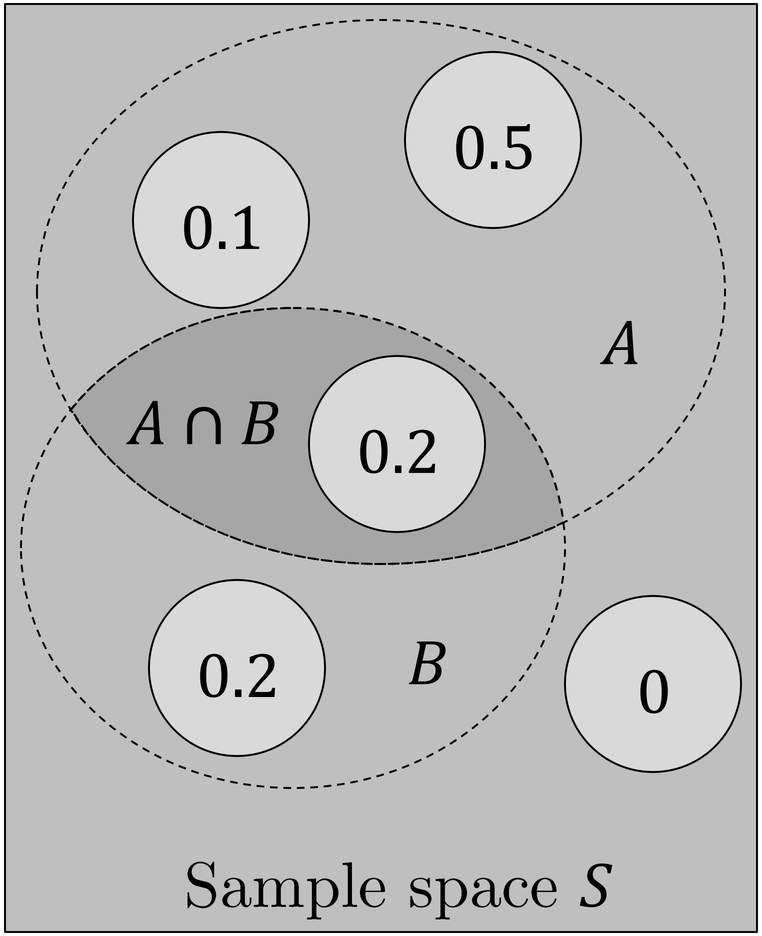


Language of probability

- We roll a white die and black die
- $S = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),$
  $(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),$
  $(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),$
  $(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),$
  $(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),$
  $(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}$
- Poll 2: Probability that the sum is 7 or 11?
  1. $1/9$
  2. $2/9$
  3. $3/9$
  4. $4/9$

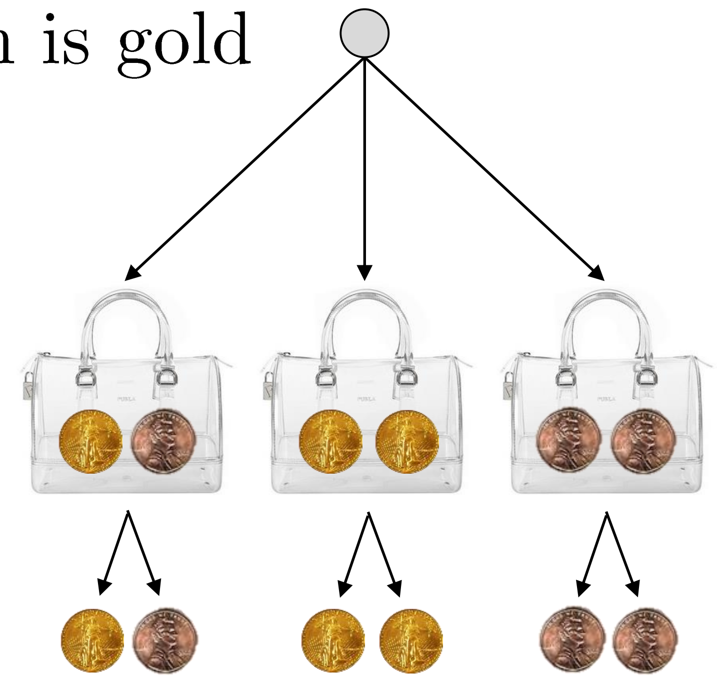
Conditional probability

- The probability of event A given event B is defined as
  $$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$
- Think of it as the proportion of $A \cap B$ to B



Pennies and gold, revisited

- Three bags contain two gold coins, two pennies, and one of each
- Bag is chosen at random, and one coin from it is selected at random; the coin is gold
- G_i: coin $i \in \{1,2\}$ is gold
- $\Pr[G_1] = \frac{1}{2}, \Pr[G_1 \cap G_2] = \frac{1}{3}$
- $\Pr[G_2 | G_1] = \frac{1/3}{1/2} = \frac{2}{3}$



The Chain Rule

- For A and B to occur, B must occur, and A must occur given that B occurred; formally:

The Chain Rule:

$$\Pr[A \cap B] = \Pr[B] \times \Pr[A|B]$$

- Applying iteratively:

$$\Pr[A_1 \cap \cdots \cap A_n] = \Pr[A_1] \times \Pr[A_2|A_1] \times \cdots \Pr[A_n|A_1, \cdots, A_{n-1}]$$

BAYES' RULE

- $\Pr[B] \times \Pr[A|B] = \Pr[A \cap B] = \Pr[A] \times \Pr[B|A]$
- Therefore,

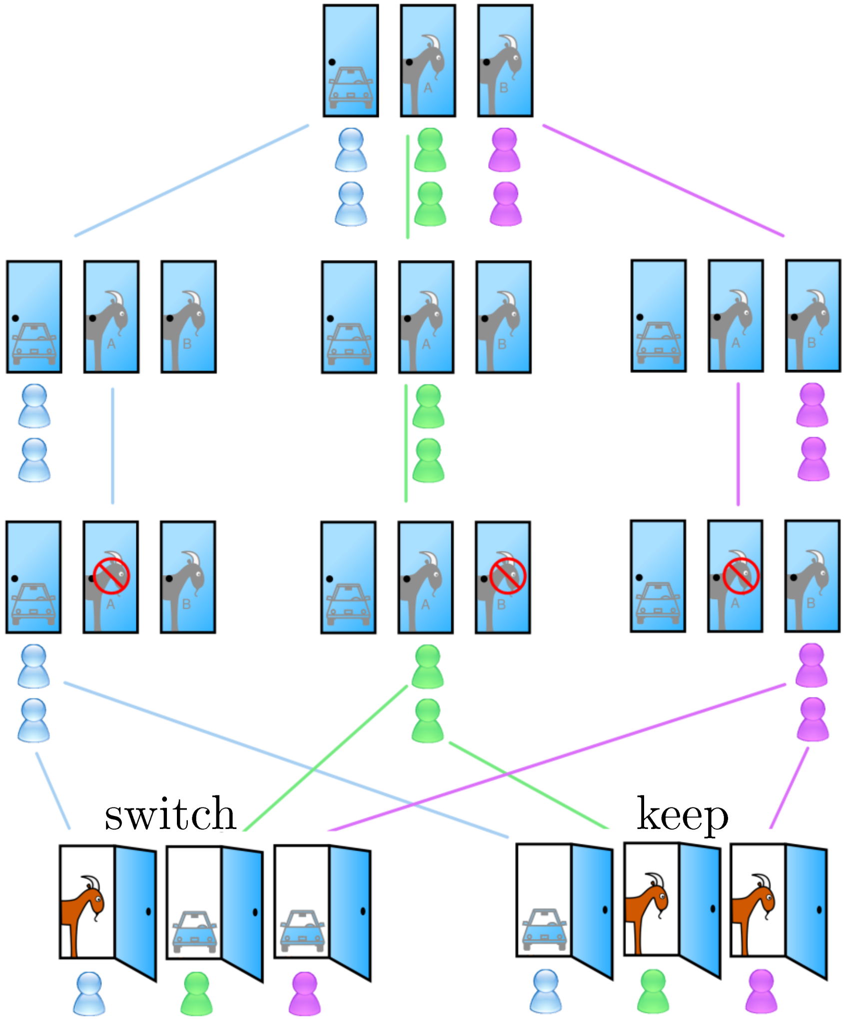
Bayes' Rule:

$$\Pr[A|B] = \frac{\Pr[A]\Pr[B|A]}{\Pr[B]}$$

Monty Hall problem

- Announcer hides prize behind one of three doors at random
- You choose a door
- Announcer opens a door with no prize
- Should you stay with your choice or switch?

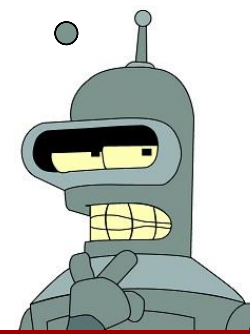
Monty Hall problem



Monty Hall problem

- Choose door 1, door 2 opens
- $\Pr[P_3|O_2] = \frac{\Pr[P_3]\Pr[O_2|P_3]}{\Pr[O_2]}$
- $\Pr[P_3] = \frac{1}{3}, \Pr[O_2|P_3] = 1,$
  $\Pr[O_2] = 1/2$
- Therefore, $\Pr[P_3|O_2] = 2/3$
- Poll 3: Assuming there are five doors, what is the probability of winning when switching?
  1. $3/15$
  2. $4/15$
  3. $5/15$
  4. $6/15$

$$\Pr[A|B] = \frac{\Pr[A]\Pr[B|A]}{\Pr[B]}$$



INDEPENDENCE

- Events A and B are independent if and only if $\Pr[A|B] = \Pr[A]$
- Poll 4: Which of the following events are independent when rolling black die and white die?
  1. Black die is 1, white die is 1
  2. Sum is 2, sum is 3
  3. Black die is 1, product is 2
  4. Black die is 1, sum is 2

The birthday paradox

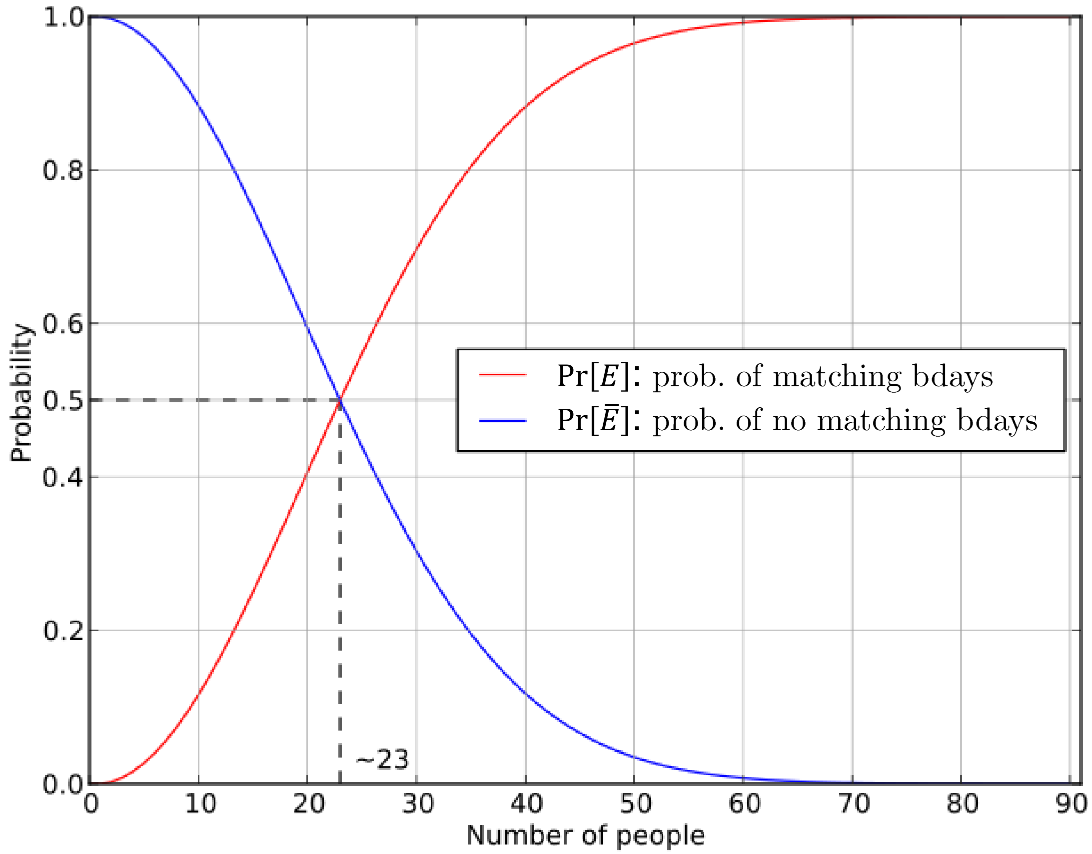
- m people in a room; suppose all birthdays are equally likely (excluding Feb 29); what is the probability that two people have the same birthday?
- $S = \{1, \dots, 365\}^m$, sample $\vec{x} = (x_1, \dots x_m)$
- $E = \{\vec{x} \in S \mid \exists i, j, \text{s.t. } x_i = x_j\}$
- Let A_i be the event that person i's birthday differs from the birthdays of $1, \dots, i-1$
- $\bar{E} = A_1 \cap \cdots \cap A_m$
- Using the chain rule:

$$\Pr[\bar{E}] = \Pr[A_1] \times \Pr[A_2 | A_1] \times \cdots \Pr[A_m | A_1, \cdots, A_{m-1}]$$

The birthday paradox

- $A_1 \cap \cdots \cap A_{i-1}$ means first $i-1$ students had different birthdays
- $i-1$ out of 365 occupied when ith birthday is chosen
- $\Pr[A_i | A_1, \ldots, A_{i-1}] = \frac{365-(i-1)}{365} = 1 - \frac{i-1}{365}$
- $\Pr[\bar{E}] = 1 \times \left(1 - \frac{1}{365}\right) \times \cdots \times \left(1 - \frac{m-1}{365}\right)$
- $\Pr[E] = 1 - \Pr[\bar{E}]$

The birthday paradox



The birthday paradox

- Poll 5: What is the probability that two people have the same birthday if there are 730 people?
  1. 1/2
  2. 0.75
  3. 0.99999999999997
  4. 1

Joint Probability Distribution

	.05	.2	0	.1
	.1	0	.1	0
	0	.1	.05	.1
	.1	0	.1	0

The Sum Rule

- How do we answer the question "What is the probability of a red car of any type?"

The Sum Rule:

$$\Pr[X] = \sum_{Y} \Pr[X \cap Y]$$

- $\Pr[X]$ is known as the marginal probability of X

Joint Probabilities:

				
	.05	.2	0	.1
	.1	0	.1	0
	0	.1	.05	.1
	.1	0	.1	0

Probability of color given type:

				
	.2	.667	0	.5
	.4	0	.4	0
	0	.333	.2	.5
	.4	0	.4	0

Summary

- Definitions / facts
  - Language of probability
  - Conditional probability
  - Independence
- Three useful rules:
  - Chain Rule
  - Bayes' Rule
  - Sum Rule