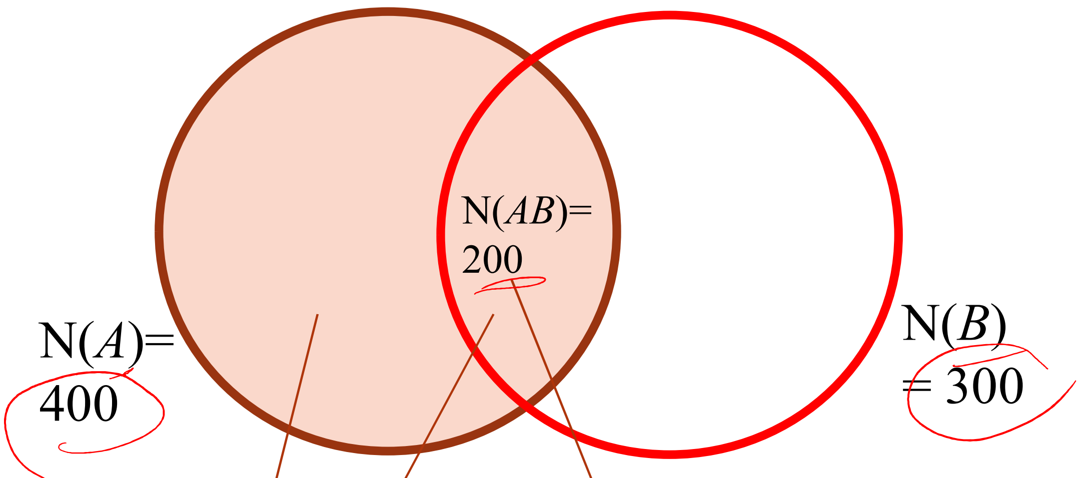
Joint probability

- The probability that events *A* and *B* both occur (in the same experiment) is called the **joint probability** of the events *A* and *B*. This is another word for the probability of the *intersection* of *A* and *B*.

Conditional and Joint probability: what's the difference?

Conditional probability is when other happens for given one happen. Joint is when both one and other happen together and it is having a larger sample space.



- Let the original sample space be *S*.
- In computing P(*B*|*A*), you assume that *A* has already occurred.
- Therefore your new sample space *S*'' for computing P(*B*|*A*) contains only those events which lie in *A*.
- For computing P(*AB*), the sample space is the entire *S*.
- Now can you compute P(*A*|*B*)?

Conditional probability

This area divided by the total sample space size is the joint probability of *A* and *B*

$P(A|B) = \frac{200}{300} = \frac{2}{3}$

Joint probability

The ratio of the smaller area to the larger area (area of the full circle) is the conditional probability of *B* given *A*.

$P(AB)$

$A \cap B = \frac{200}{N(A) + N(B) - N(AB)} = \frac{200}{500} = \frac{2}{5}$

If $P(A|B) > P(AB)$? Yes

Yes, because AB has larger sample space.

Conditional and joint probability: what's the difference

- Consider two consecutive rolls of a die. Given that the first die produced a 3, what's the probability that the sum of the two throws does not exceed 7?

A

B

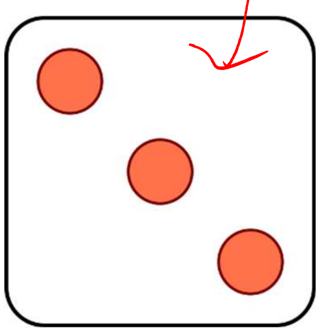
- Solution:

\# of combinations s.t

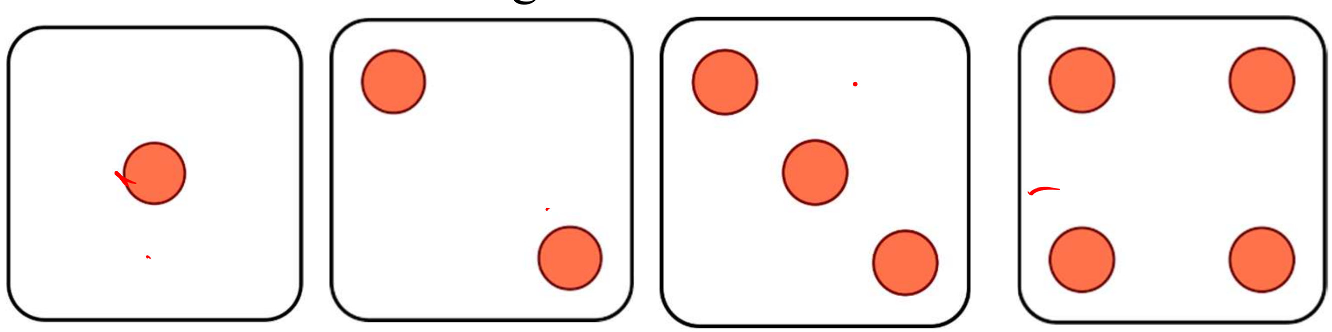
  - A = event that first throw produced a 3. P(A) = 1/6.
  - B = event that sum does not exceed 7. $P(B) = \frac{21}{36}$
  - Joint probability P(AB) = 4/36.
  - Conditional probability: P(B|A) = P(AB)/P(A) = (4/36)/(1/6)=2/3.

We accepted that first one has occurred with probability 1, then our focus shifts on the second one, how is that behaving.

1st roll



2nd roll given that 1st roll is a 3



Example

India has a literacy rate of 74%. The state of Kerala has a literacy rate of 94% and constitutes 2.8% of India's population.

K

L

What is the probability that:

- A randomly chosen Indian person is literate $.74$ $P(L)$
- A randomly chosen Indian person is from Kerala $.028$ $P(K)$
- A randomly chosen person from Kerala is literate $.94$ $P(L|K)$
- A randomly chosen Indian person is from Kerala and is literate $P(KL) = P(L|K)P(K) = 0.94 \times 0.028$
- A randomly chosen Indian person is from Kerala if you knew already that (s)he was literate

$$P(K|L) = \frac{P(L|K)P(K)}{P(L)}$$

Example

India has a literacy rate of 74%. The state of Kerala has a literacy rate of 94% and constitutes 2.8% of India's population.

What is the probability that:

- A randomly chosen Indian person is literate $P(L)=0.74$
- A randomly chosen Indian person is from Kerala $P(K)=0.028$
- A randomly chosen person from Kerala is literate $P(L|K)=0.94$
- A randomly chosen person is from Kerala and is literate
  $P(K, L)=P(L|K)P(K)=0.94*0.028$
- A randomly chosen person is from Kerala if you knew already that (s)he was literate $P(K|L)=P(K, L)/P(L)=0.94*0.028/0.74$

Independence of events

- If A and B are independent, the occurrence of A has no bearing on the probability of occurrence of B (and vice versa). Events are independent if the occurrence of one doesn't affect the occurrence of other.
- Examples of independent events:
  - Outcomes from two dice rolls
  - Height of a person and the last digit of their mobile phone
  - Rainfall tomorrow in Mumbai and number of winning ticket lottery
- Example of dependent events
  - Is it raining in morning & Is it sunny in the afternoon
  - Two pulls of balls from a bag without replacement.
  - Preparation for a test & marks.

Independence

Two events A and B are **independent** if:

$$P(A) = P(A|B)$$

Intuitive Definition:

Knowing that event B happened doesn't change our belief that A happens.

With independence, we can simplify the chain rule:

$$P(AB) = P(A \cap B) = P(A|B) \cdot P(B)$$
$$= P(A) \cdot P(B)$$

You can also show this ^ to prove independence

joint probability becomes the product of individual probability.

Independence of more than two events

- We say that n > 2 events are mutually independent if and only if for every subset A of k ≤ n events, we have:

$$P\left(\bigcap_{i=1}^{k} P(A_i)\right) = \prod_{i=1}^{k} P(A_i)$$

$n = 3 : A_1, A_2, A_3$

- Example: three events ~~A, B, C~~. To show that they are independent we need to show that:

$$P(A_1 A_2 A_3) = P(A_1)\,P(A_2)\,P(A_3)$$

$$P(A_1 A_2) = P(A_1)\,P(A_2)$$

$$P(A_1 A_3) = P(A_1)\,P(A_3)$$

$$P(A_2 A_3) = P(A_2)\,P(A_3)$$

Example 3.8.c. Two fair dice are thrown. Let E_7 denote the event that the sum of the dice is 7. Let F denote the event that the first die equals 4 and let T be the event that the second die equals 3. Now it can be shown (see Problem 36) that E_7 is independent of F and that E_7 is also independent of T; but clearly E_7 is not independent of FT [since $P(E_7|FT) = 1$]. ■

$$\left.\begin{array}{l} E_7 \perp\!\!\!\perp T \\ E_7 \perp\!\!\!\perp F \end{array}\right\} \not\Rightarrow E_7 \perp\!\!\!\perp TF$$

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]

}

$P(E_7) = \frac{1}{6}$ ✓

$P(F) = \frac{1}{6}$ ✓

$P(T) = \frac{1}{6}$

$P(E_7 \cap F) = \frac{1}{36}$

$P(E_7 F) = P(E_7)P(F)$ ✓

$P(E_7 T) = P(E_7)P(T)$

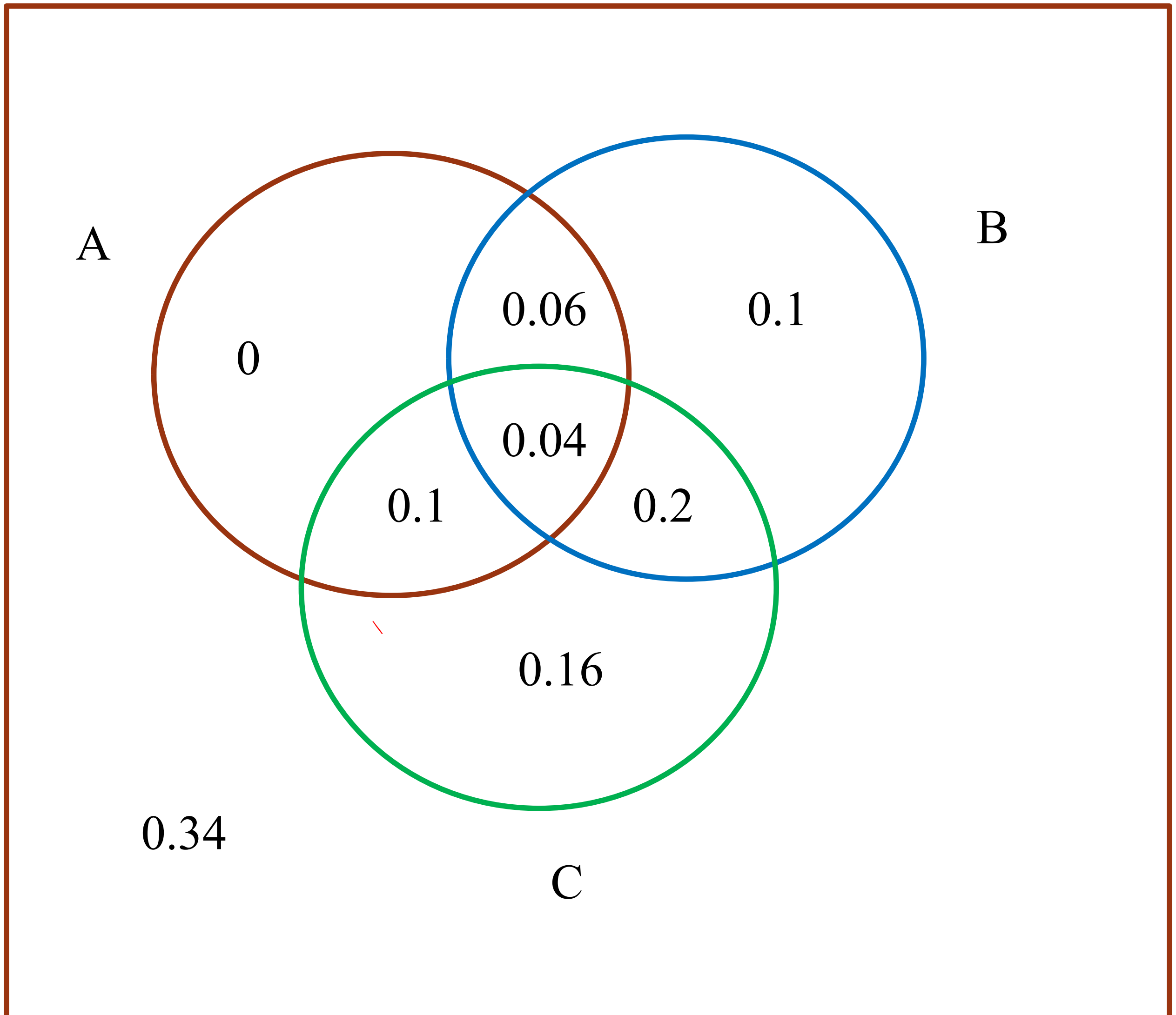
$P(E_7 | FT) = 1 \neq P(E_7)$

Only n-way independence does not suffice

- Note that only n-way independence of events does not imply that every pair of events are independent.

- Example: See next slide

http://www.engr.mun.ca/~ggeorge/MathGaz04.pdf



$$P(ABC) = 0.04 = P(A)P(B)P(C) = (0.2)(0.4)(0.5)$$

$$P(AB) = 0.1 \neq P(A)P(B)$$

Independence versus Mutual Exclusion

- If A and B are mutually exclusive, then $P(AB) = 0$.

- If A and B are independent, then $P(AB) = P(A)P(B) \neq 0$.

- The two are usually not the same! In fact, for mutually exclusive events, the occurrence of one *does* have an effect on that of the other.

Independence and mutual exclusion

If A is independent of B, can we say that A is independent of B^c? Yes,

Yes:

$$P(AB) = P(A)P(B) \quad \because A \perp\!\!\!\perp B$$

$$P(A) = P(AB) + P(AB^c) \quad [\text{Law of total probability}]$$

$$= P(A)P(B) + P(AB^c)$$

$$\Rightarrow P(AB^c) = P(A) - P(A)P(B)$$

$$= P(A)[1 - P(B)] =$$

$$= P(A)P(B^c)$$

$$\Rightarrow A \perp\!\!\!\perp B^c$$

The Core Probability Toolkit

The Law of Total Probability

$S = \bigcup_{i=1} B_i$

$B_i \cap B_j = \phi$

$$P(E) = P(E \text{ and } F) + P(E \text{ and } F^C)$$

$$P(E) = P(E|F)\,P(F) + P(E|F^C)\,P(F^C)$$

$$P(E) = \sum_{i=1}^{n} P(E \text{ and } B_i) = \sum_{i=1}^{n} P(E|B_i)\,P(B_i)$$

Bayes' Theorem

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E)}$$

$$P(B|E) = \frac{P(E|B) \cdot P(B)}{P(E|B) \cdot P(B) + P(E|B^C) \cdot P(B^C)}$$

Definition of Conditional Probability

$$P(E|F) = \frac{P(E \text{ and } F)}{P(F)}$$

Axiom 1: $0 \le P(E) \le 1$

Axiom 2: $P(S) = 1$

$$P(E^C) = 1 - P(E)$$

De Morgan's Laws

$(A \text{ or } B)^C = A^C \text{ and } B^C$

$(A \text{ and } B)^C = A^C \text{ or } B^C$

Axiom 3: If E and F are mutually exclusive, then $P(E \text{ or } F) = P(E) + P(F)$

Otherwise, use Inclusion-Exclusion:

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

Chain Rule

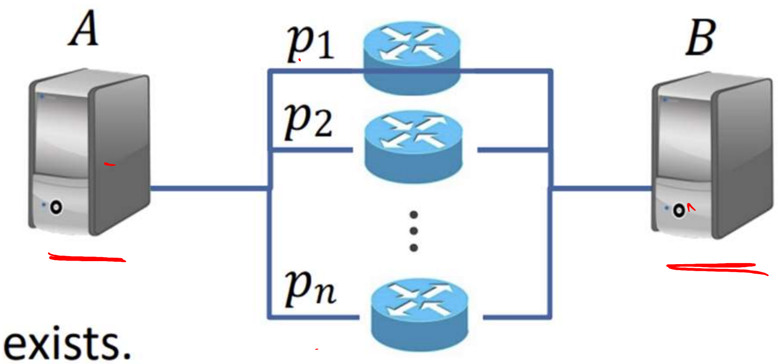
$$P(E \text{ and } F) = P(E|F) \cdot P(F) = P(F|E) \cdot P(E)$$

Independence

$$P(E|F) = P(E)$$

$$P(E \text{ and } F) = P(E)\,P(F)$$

Practice: Network Reliability

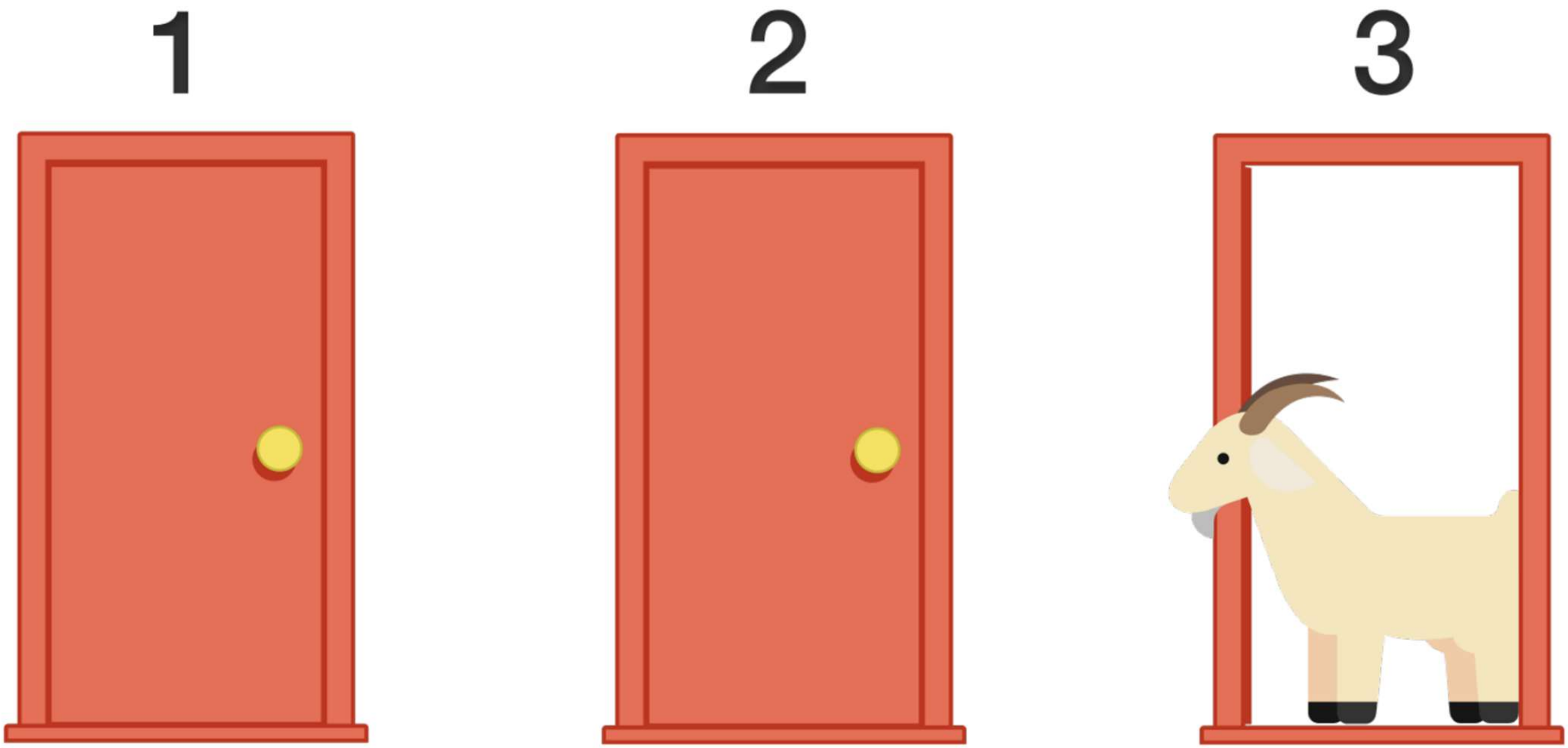


Consider the following parallel network:

- n independent routers, which each are working with probability p_i $(1 \le i \le n)$

Let E be the event that a working path from A to B exists. What is $P(E)$?

$$P(E) = 1 - \prod_{i=1}^{n} (1 - p_i)$$



The Monty Hall Problem

The Monty Hall Problem

Behind one door is a prize (equally likely for each door).

Behind the other two doors are goats.

How to play:

1. We choose a door.
2. Host opens 1 of the other 2 doors, revealing a goat.
3. We are given an option to switch to the other door.

Note: If we don't switch,
P(win) = 1/3

We are comparing
P(win) vs. P(win|switch)

Should we switch?

$P(\text{win}) = P(\text{win} \mid \text{switch}$

$P(\text{prize in } 1) = P(\text{prize in } 2)$

Let's Find P(win | switch)

Paul Erdos

Without loss of generality, let's pick door A (out of doors A,B,C).

1/3 1/3 1/3

<u>A = prize</u>

<u>B = prize</u>

<u>C = prize</u>

Let's Find P(win | switch)

Without loss of generality, let's pick door A (out of doors A,B,C).

1/3 1/3 1/3

A = prize

- Host opens B or C
- We switch
- We **lose**

P(win | A prize, pick A, switch) = 0

B = prize

- Host must open C
- We switch to B
- We **win**

P(win | B prize, pick A, switch) = 1

C = prize

- Host must open B
- We switch to C
- We **win**

P(win | C prize, pick A, switch) = 1

P(win | pick A, switch) = P(win | A prize, pick A, switch) * P(A prize) +
P(win | B prize, pick A, switch) * P(B prize) +
P(win | C prize, pick A, switch) * P(C prize)

= 1/3 * 0 + 1/3 * 1 + 1/3 * 1 = 2/3

You should switch!

The Birthday Paradox!

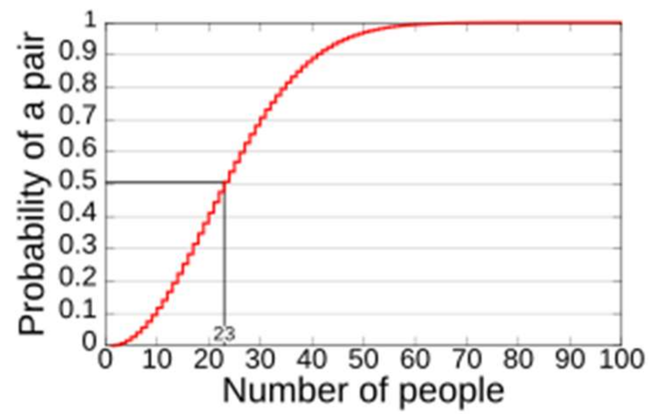
- Given n people in a room, what should be the least value of n such that the probability that at least 2 people in the room share the same birthday is greater than or equal to 99.9%?

- Each person can have his/her birthday on any of the 365 days. For n people, there are 365^n outcomes.

- The number of outcomes resulting in no two people sharing a birthday is (365)(364)(363)…(365-n+1).

The Birthday Paradox!

- So required probability is
- 1-(365)(364)(363)…(365-n+1)/(365)n ≥ 0.999 (given)
- This is satisfied for n as small as 70.
- For n = 20, it is around 41%.
- For n = 23 it is close to 50%
- For n = 40, it is around 89%.



- For more information see the [wikipedia article on the birthday paradox.]

Conclusions

- Reasoning about probabilities is tricky
- Important to carefully analyze the sample space, and conditioning variable