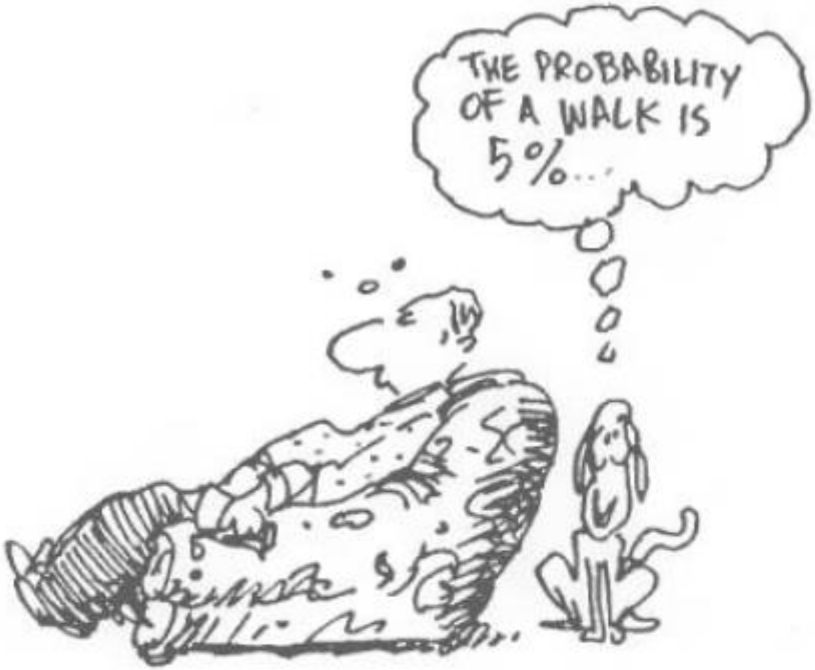
AN INTRODUCTION TO PROBABILITY

Probability, odds, outcomes and events

Nihan Acar-Denizli, Pau Fonseca

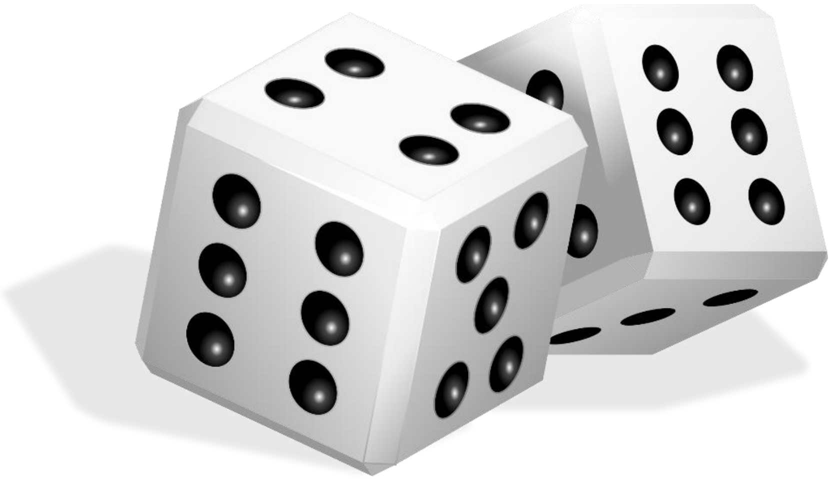
Probability

Probability is a numeric measure of the likelihood of an event that will occur.



Probability

- On a random phenomenon individual outcomes are not certain, but there is a regular distribution of outcomes in the long run.
- The probability of an outcome is its long-term relative frequency.



Definitions

- An experiment that can result in different outcomes, even though it is repeated in the same manner every time, is called a random experiment.
- A single performance of a random experiment is known as a trial.
- The outcomes are possible results of a trial. The individual outcomes of a random experiment are not certain -determined by chance-, but there is a regular distribution of outcomes in the long run.
- The probability of an outcome is its long-term relative frequency.

Probability vs. inference

- If a coin is tossed 100 times, you would expect to see 50 heads.
- However, there is some (non-zero) probability that it only comes up heads 25 times.
- If there were 100 flips, and it only came up heads 25 times, you **might infer** that it wasn't a fair coin.
- Claim + Evidence + Reasoning = Inference.

What are the possible outcomes?

Make a list of possible outcomes, then find probability for each outcome.

- **Sample space** is the set of all possible outcomes. It is denoted by S.
- **Events** are specific outcomes or set of outcomes in the sample space.

Definitions

- **Simple event** is the most basic outcome of an experiment.
- Two special events are null or impossible event (it is never an outcome) and true or universal event (the entire sample space).
- **Theoretical Probability**: n(E)/n(S), (# of favorable outcomes)/(total # of outcomes).

The steps in analysing a random experiment

1. Identify the sample space

- Its elements are mutually exclusive and collectively exhaustive
- Might be finite or infinite: Toss a coin vs Get a 'head' tossing a coin
- Usually many choices are possible, some of them best suited than others to compute desired probabilities

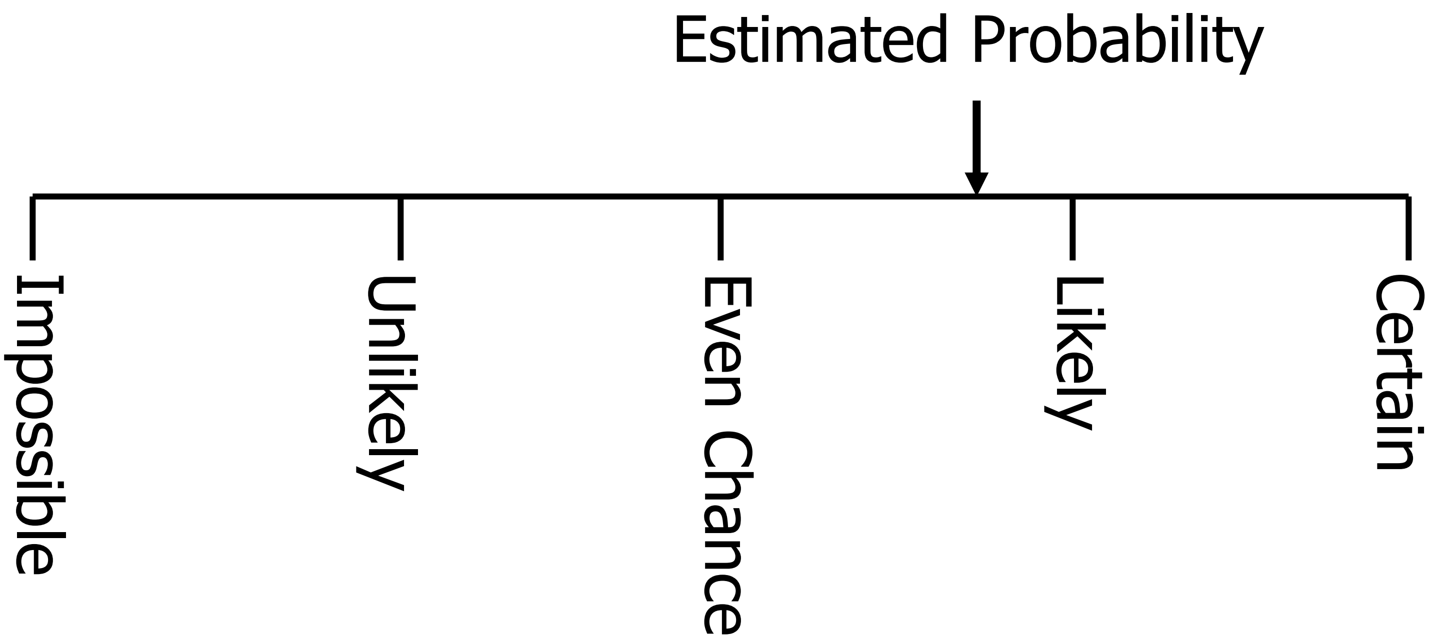
2. Assign probabilities (Classical method, Relative Frequency method, Subjective method)

3. Identify the events of interest

4. Compute desired probabilities

Probability Scale

You can draw a scale from Impossible to Certain:



What is the probability of scoring less than 5 when rolling a fair 6-sided die?

Introduction to Probability

- Rather than using a descriptive scale, we use a **numerical scale.**
- Probability is a numerical measure of the likelihood that an event will occur.
- Probability values are always assigned on a scale from 0 to 1.
- A probability near 0 indicates an event is very unlikely to occur.
- A probability near 1 indicates an event is almost certain to occur.

Useful facts about probability

- Probability cannot be less than 0 or greater than 1.
- All possible outcomes together must have probability 1.
- Probability of an event occurring is 1 minus probability that it does not occur.
- If two events have no outcomes in common, probability that one or the other occurs is sum of their individual probabilities.

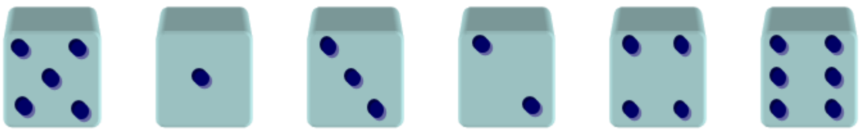
Example

- You roll a six-sided dice whose sides are numbered from 1 through 6. Find the probability of:

a) rolling a 4:
   - number of ways to roll a 4/number of ways to roll the dice=1/6

b) rolling an odd number
   - number of ways to roll an odd number/number of ways to roll the dice=3/6=1/2

c) rolling a number less than 7
   - number of ways to roll less than 7/number of ways to roll the dice=6/6=1



Example

What is the probability of scoring less than 5 when rolling a fair 6-sided die?

1. Identify the sample space: $\Omega = \{1,2,3,4,5,6\}$

2. Assign probabilities of each elemental event: 1/6

3. Identify the events of interest: A= {1,2,3,4}

4.Compute desired probabilities.

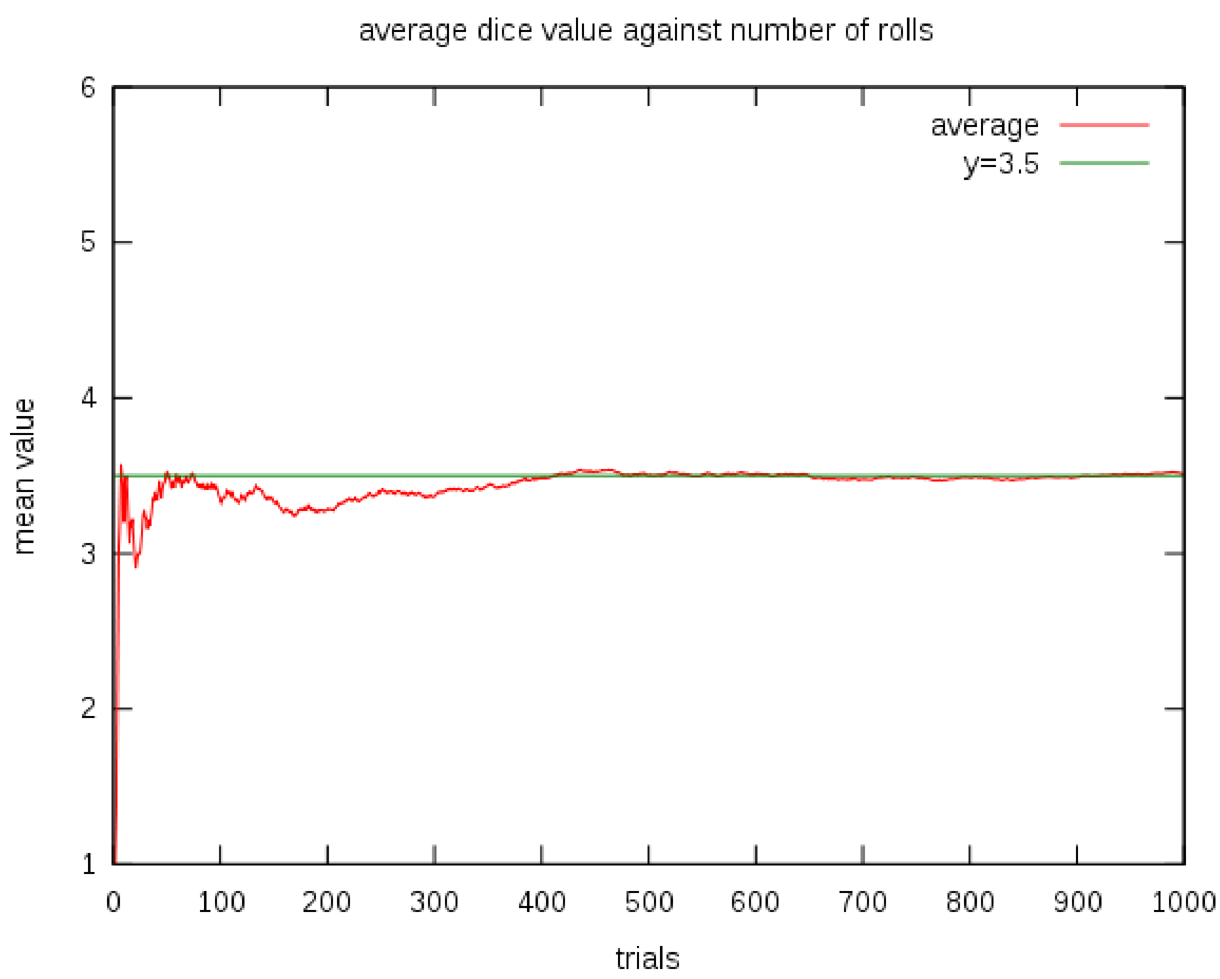
P(A) = P({1}) + P({2}) + P({3}) + P({4}) = 4/6 = 2/3.

Terminology

- <u>Empirical probability</u>: P(E) = (# of times event E occurred)/(# of times experiment was performed)

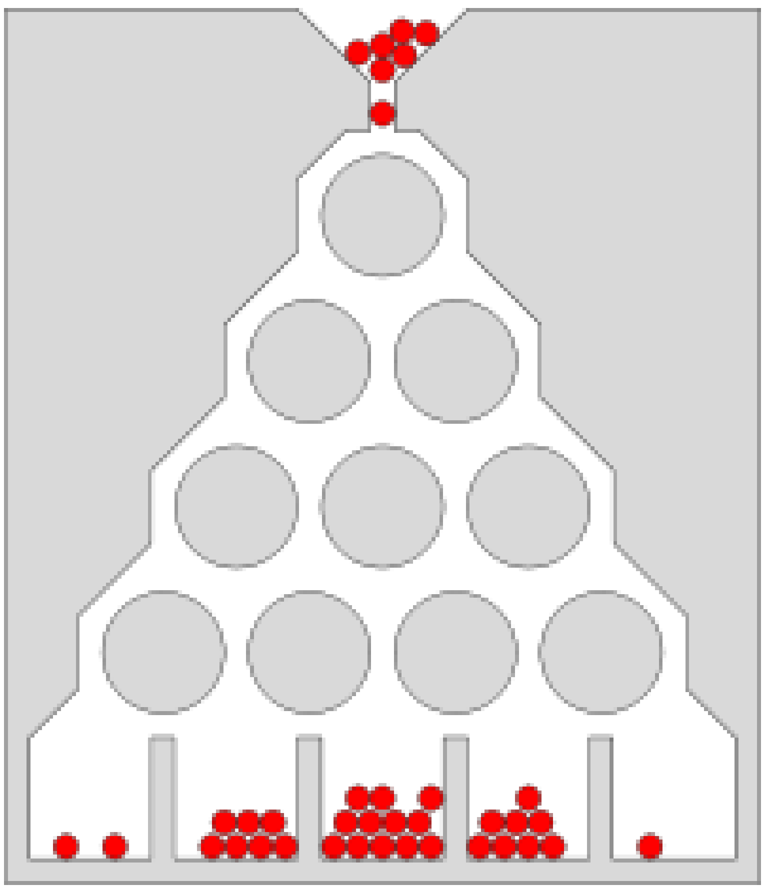
  This is defined by experimentation
  - Did: P(face card) = 12/52=0.2307….
  - P(not a face card)=1-0.2307=0.7693…=40/52
- <u>Law of Large Numbers (Law of averages):</u> An experiment is repeated more and more times, the proportion of outcomes favorable to any particular event will come closer and closer to the theoretical probability of that event.

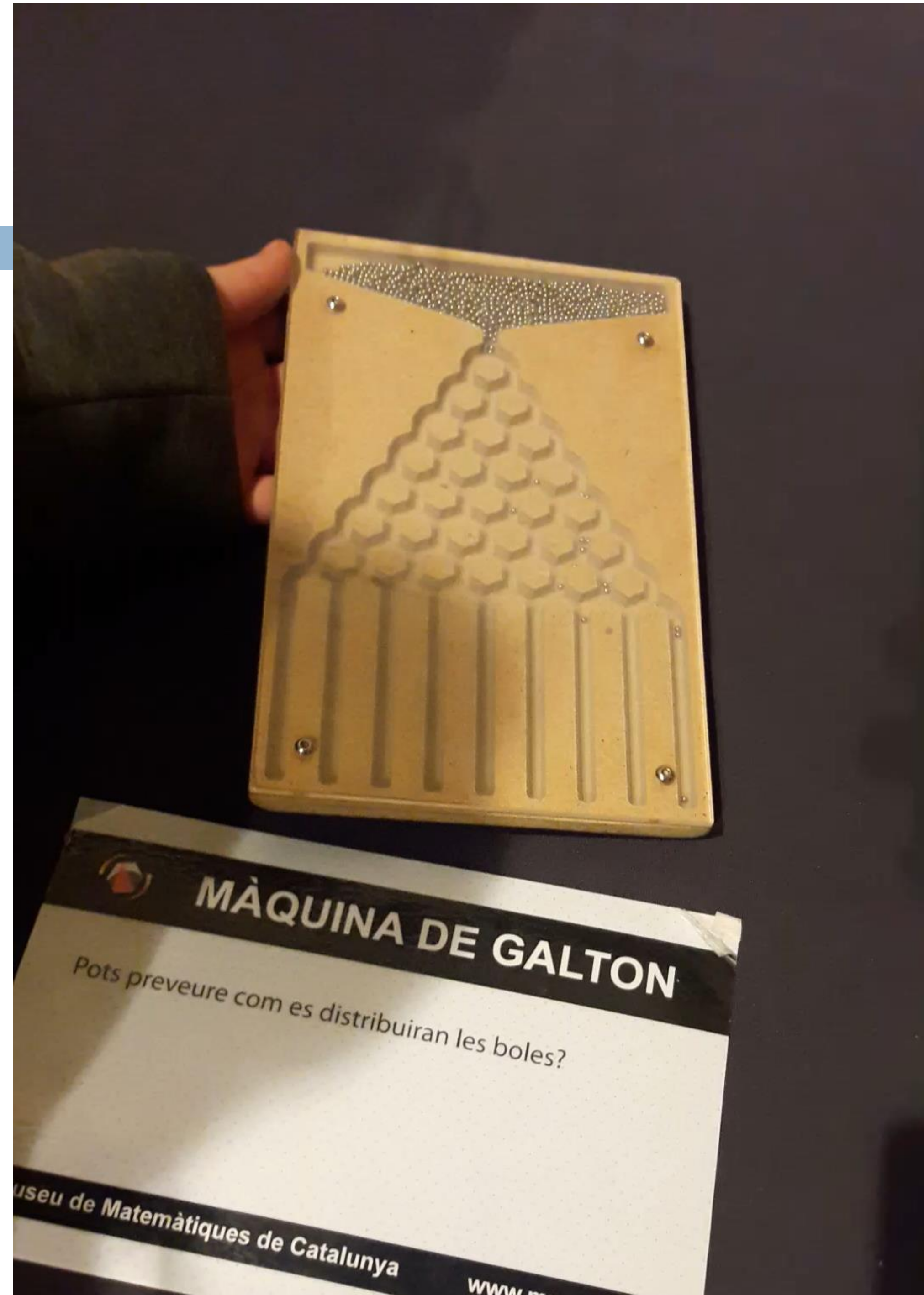
Law of large numbers



Galton box

https://galtonboard.com/probabilityexamplesinlife





Probability questions

- "I have two children. One is a boy, and one is as a….." What is the chance that I have two boys?
  - Ans: 1/3: outcomes-- BB, BG, GB
- "I have two children. The older is a boy, and ….." What is the chance that I have two boys?
  - Ans: ½

Combining probabilities

- What happens if we need to calculate the probability of one event occurring **and** another event occurring?

- What is the probability of rolling a 5 with a die and tossing a TAIL with a coin?

- How many outcomes are there?

Combining probabilities

- 12 possible outcomes.

Our outcome

Roll	Dice	Coin
1	1	T
2	2	T
3	3	T
4	4	T
5	5	T
6	6	T
7	1	H
8	2	H
9	3	H
10	4	H
11	5	H
12	6	H

Combining probabilities

- The probability of rolling a 5 is 1/6.
- The probability of throwing a TAIL is ½.
- So P(5 and TAIL) is

  $1/6 \times ½ = 1/12$.

Birthday problem

- How many people are needed in a room so that the probability that there are two people whose birthdays are the exactly the same day is roughly ½?
- How many pairs of dates do we have?

  365x365

- How many pairs which are guaranteed that to people are not sharing the date:
  - 365x364; (364*365)/(365*365) = 364/365=0.9972…

Birthday problem

- 3, no sharing the date :
  - (365*364*363)/(365*365*365)=0.9917;
  - P(sharing)=1-0.9917…= 0.0082…
- 4, no sharing the date :
  - (365*364*363*362)/(365)^4 = 0.01635;
- Pattern for the probability

P(no sharing the date): $\frac{365!/(365-n)!}{365^n}$

Birthday problem

# people in room	P (2 sharing birthday)
5	0.027
10	0.117
20	0.411
23	0.507
50	0.970
70	0.994
80	0.99991
90	0.999993

Birthday problem

- Thus, at 23 people, the probability of
  P(2 people sharing the birthday) = 0.5072

- We there are 9 people,
  P(not sharing)=0,9, is it true?

Set-Theoretic Point of View of Probability

Basic Properties and Algebra of Events

- Sample spaces are formally described using sets and operators in the classical Set Theory.
- Complement of event A denoted as A^c or A.
- Intersection of events A and B $(A \cap B)$ contains outcomes belonging simultaneous to A and B.
- Union of events A and B $(A \cup B)$ contains outcomes belonging either to A or B or both.

Basic Properties and Algebra of Events

- Commutative laws: $A \cup B = B \cup A$ and $A \cap B = B \cap A$.
- Associative laws : $A \cup (B \cup C) = (A \cup B) \cup C$. Same for $\cap$.
- Distributive laws :

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

- Identity laws: $A \cup \emptyset = A$ and $A \cap \emptyset = \emptyset$.
- Complementation laws : $A \cup \bar{A} = \Omega$ and $A \cap \bar{A} = \emptyset$.
- De Morgan's laws:

$$\bar{A} \cup \bar{B} = \overline{A \cap B}.$$
$$\bar{A} \cap \bar{B} = \overline{A \cup B}.$$

Basic Properties and Algebra of Events

$P: \boldsymbol{\mathcal{P}} \to [0,1]$ (Domain set of all events)

- For any event A, P(A) > 0.
- P(Ω)=1.
- P(A ∪ B) =P(A)+P(B), whenever A and B are mutually exclusive events.
- The probability values assigned to each experimental outcome (sample point) must be between 0 and 1.
  - A probability near 0 indicates an event is very unlikely to occur.
  - A probability near 1 indicates an event is almost certain to occur.

Set-Theoretic Point of View of Probability

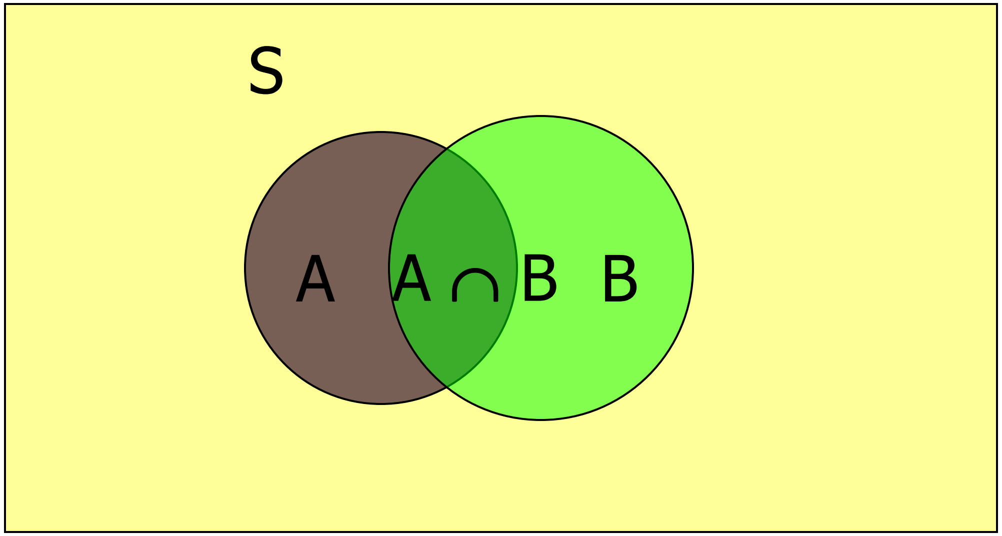
- Consider a set S. For each subset X of S, we associate a number

  $0 \le P(X) \le 1$, such that

  $P(\varnothing) = 0, \quad P(S) = 1,$

  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Venn Diagram



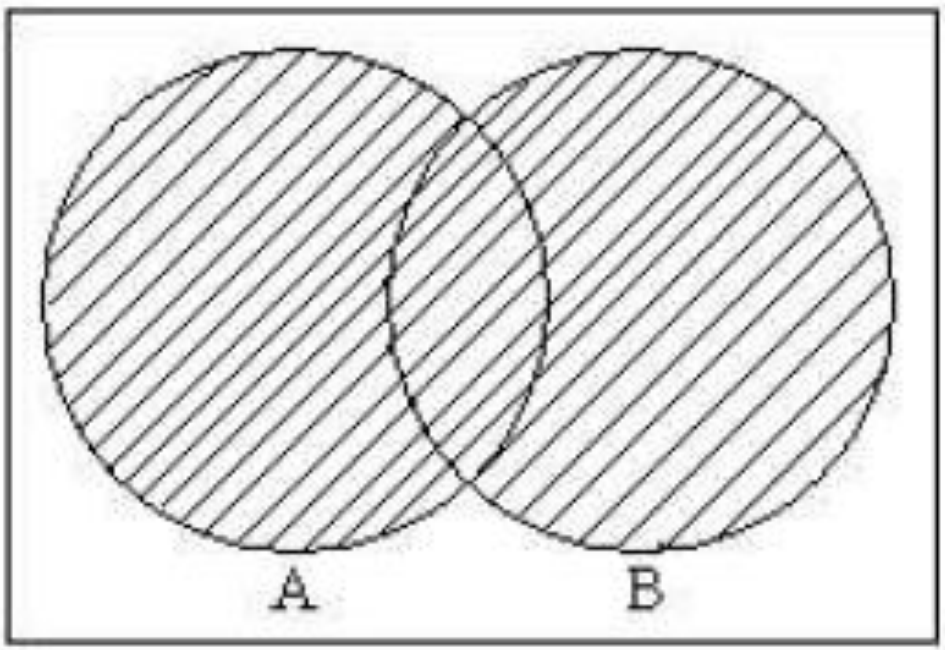
∪: union, or

∩: intersection, and

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

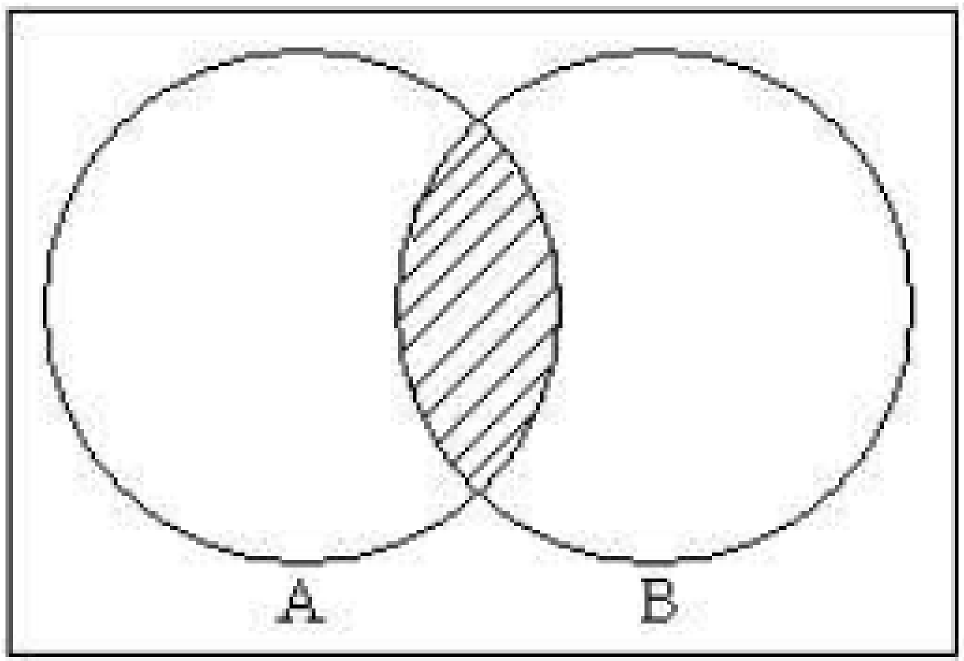
Union of sets

- $A \cup B = \{s \in S : s \in A \text{ or } s \in B\}$



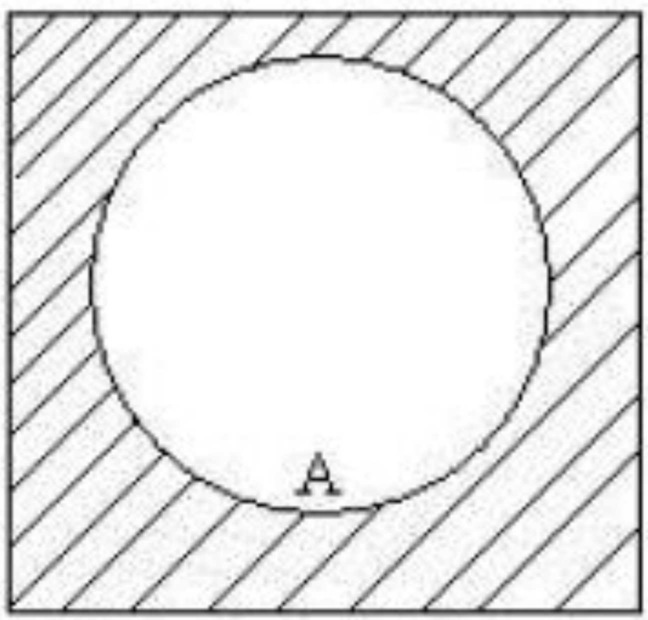
Intersection of Sets

- $A \cap B = AB = \{s \in S : s \in A \text{ and } s \in B\}$



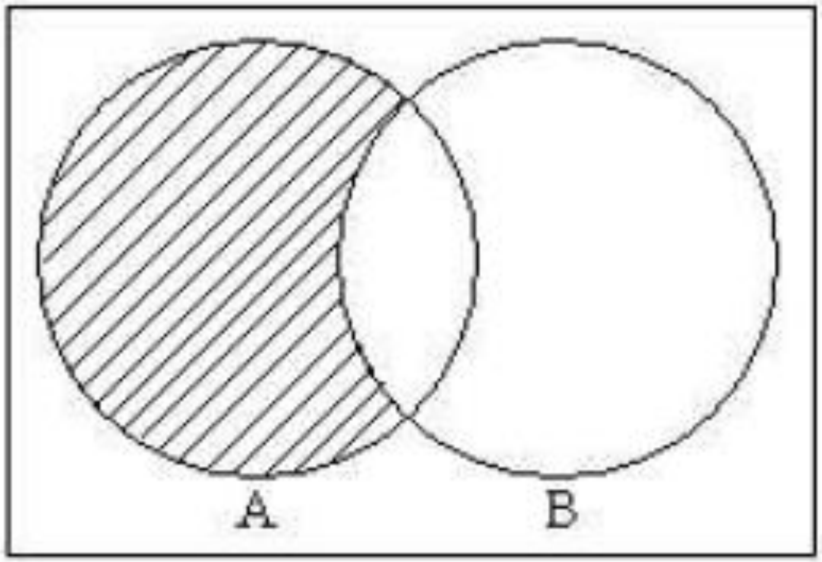
Complement

- $A^c = \{s \in S : s \notin A\}$



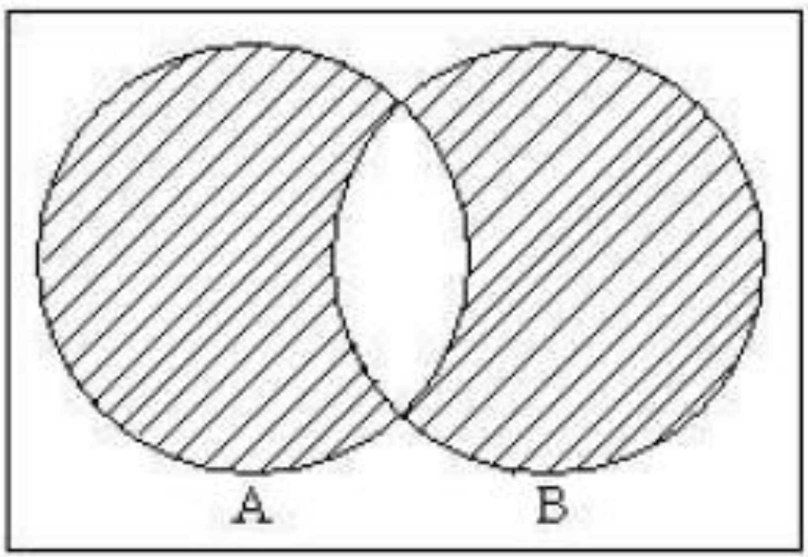
Difference

- $A \backslash B = A - B = \{s \in S : s \in A \text{ and } s \notin B\} = A \cap B^c$



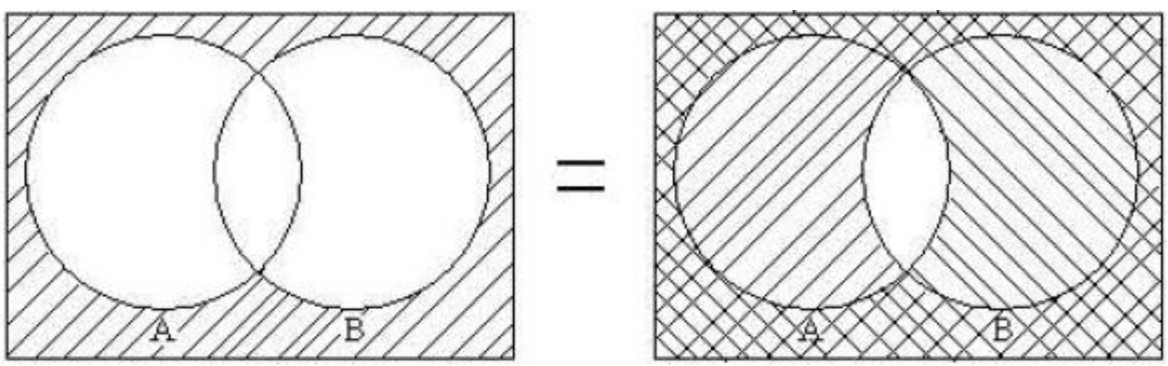
Symmetric Difference

- $A \Delta B = \{s \in S : (s \in A \text{ and } s \notin B) \text{ or } (s \in B \text{ and } s \notin A)\} = (A \cap B^c) \cup (B \cap A^c)$



Properties of Set Operations

- $A \cup B = B \cup A$
- $(A \cup B) \cup C = A \cup (B \cup C)$
- The same for intersections
- Associative rule: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
- $(A \cup B)^c = A^c \cap B^c$



Properties of Set Operations

- $(A \cap B)^c = A^c \cup B^c$
  - $s \in (A \cap B)^c = s \notin (A \cap B)$
  - $s \notin A$ or $s \notin B = s \in A^c$ or $s \in B^c$
  - $s \in (A^c \cup B^c)$

Mutual Exclusion

- If two events (subsets) A and B cannot happen simultaneously, i.e.,

  $A \cap B = \varnothing$, we say A and B are mutually exclusive events.

- For mutually exclusive events,

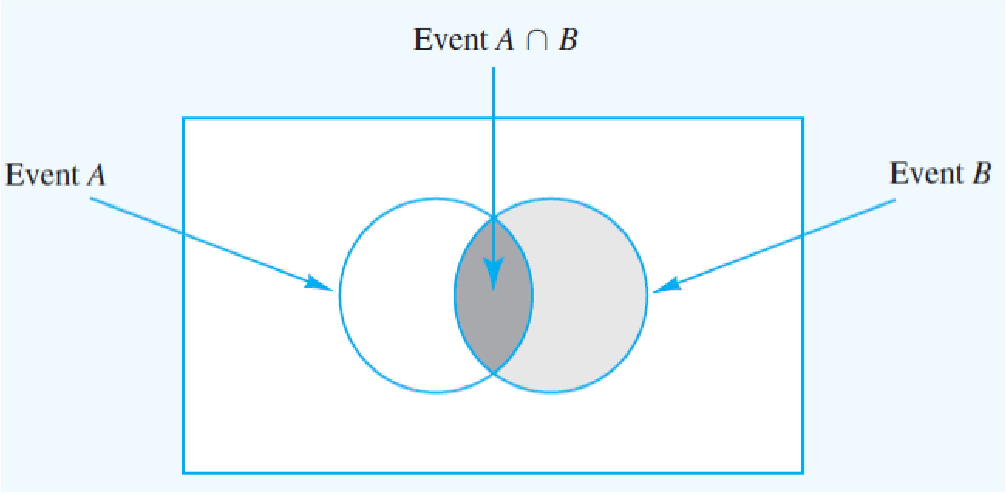
  $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$

Conditional Probability

The probability of event A given the condition that event B has ocurred. The probability of event A given that event B has already happened is,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

where $P(B) > 0$.



Conditional probabilities

- Example: 2 fair coins are tossed –at least one is a head -- find the probability that both are heads:
  - A= 1 head; B = both heads;
  - $P(B|A) = P(A \cap B)/P(A) = (1/4)/(3/4) = 1/3$

Independence

- If $P(A|B) = P(A)$, then we say A is **independent** of B.

- Equivalently,

  $P(A \cap B) = P(A)\,P(B)$, if A and B are independent.

Bayes's Theorem

- This theorem gives the relationship between P(A|B) and P(B|A):

$$P(A \mid B) = P(B \mid A)\frac{P(A)}{P(B)}$$

This equation forms the basis for Bayesian statistical analysis.

Two Schools

- Frequentists: Fraction of times an event occurs if it is repeated "N" times.
- Bayesians: A probability is a degree of belief.

Selection on a set

Variations

Permutations

Combinations

… and with repetition

Selection on a set

- Given a set A of n elements, how many different ways can we select r elements of A?
- Keep in mind:
  - If the **size** of the selection coincides with the number of elements ($r = n$).
  - or whether to take into account the **order** of items or not.
  - or if allowed **repetitions** of elements or not.

Variations

- n : number of elements
- r : number of elements in the groups
- $r < n$, order matters
- $V_{n,r}$ are the number of r ordered groups of different elements that can be formed with n elements.
  - $V_{n,r}$ = n (n-1) (n-2) ... (n-r +1)
  - Example: <u>five-letter</u> words you do not repeat any letter (26 letters in the alphabet)

Combinations

- n : number of elements
- r : number of elements in the groups
- $r < n$, order doesn't matter
- $C_{n,r}$ is the number of r groups of different elements can be formed with n elements.

$$\mathrm{C}_{n,r} = \frac{V_{n,r}}{P_r} = \frac{n(n-1)(n-2)\ldots(n-r+1)}{r!} = \frac{n!}{r!(n-r)!} = \binom{n}{r}$$

- Example: possible cases to distribute 4 cards from a deck of 40.

Permutations

- n : number of elements
- r : number of elements in the groups
- $r = n$, order matters
- P_n is the number of different ways to sort n elements.
  - $P_n = n!$
  - Example possible ways of ordering the vowels

Variations with repetition

- n : number of elements
- r : number of elements in the groups
- $r < n$, order matters
- $V_{n,r}$ are the number of r ordered groups of different elements that can be formed with n elements.
  - $V_{n,r} = n^r$
  - Example: five-letter words you can repeat any letter (26 letters in the alphabet)

Combinations with repetition

- n : number of elements
- r : number of elements in the groups
- $r < n$, order doesn't matter
- $C_{n,r}$ is the number of r groups of different elements can be formed with n elements.

$$\mathrm{CR}_{n,r} = \binom{n+r-1}{r} = \frac{(n+r-1)!}{r!(n-1)!}$$

- Selection of 4 capsules of coffee out of 10 different type of coffee capsules.

Permutations with repetition

- n : number of elements
- r : number of elements in the groups
- $r = n$, order matters
- P_n is the number of different ways to sort n elements. The first element can be repeated r_1 times, the second r_2 times and the k element r_k times.
- Also $r_1 + r_2 + .. + r_k = n$

$$P_n^{r_1, r_2, \ldots, r_k} = \frac{n!}{r_1! r_2! \ldots r_k!}$$

- Example possible words with the letters of STATISTICS.

To know more

- Part I: Probability and Random and Chapter 2: Probability of **Variables of Probability and Statistics for Computer Scientists** (2014 Ed.)