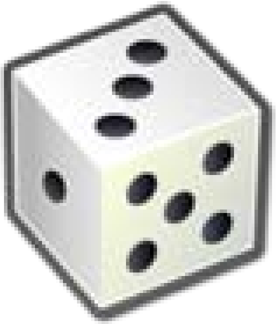
AN INTRODUCTION TO PROBABILITY

Probability, odds, outcomes and events

Probability?

- On a random phenomenon individual outcomes are not certain, but there is a regular distribution of outcomes in the long run.
- A single performance of the experiment is known as a trial.
- 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.

What are the possible outcomes?

- Want to make a list of possible outcomes in a trial, and then find probability for each outcome
- **Sample space** is the set of all possible outcomes in a single performance of the experiment.
- **Events** are specific outcomes or set of outcomes in the sample space
- Two special events: *null or impossible event* (it is never an outcome) and *true or universal event* (the entire sample space).
  - Algebra of events

Definitions

- Trial – any observation or measurement of a random phenomenon experiment (the outcome cannot be predicted with certainty)
- Simple event – the most basic outcome of a trial
- Outcomes– possible results of a trial
- Sample Space – the set of all possible outcomes (i.e., the collection of all simple events
- Theoretical Probability: n(E)/n(S), (# of favorable outcomes)/(total # of outcomes)

Outline of the steps in analysing a random experiment:

- 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
- Assign probabilities
- Identify the events of interest
- Compute desired probabilities

Example: Roll a fair die

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

- Identify the sample space: $\Omega = \{1,2,3,4,5,6\}$
- Assign probabilities: each elemental event $1/6$
- Identify the events of interest? $A = \{1,2,3,4\}$
- Compute desired probabilities?
  - $P(A)=P(\{1\})+ P(\{2\})+ P(\{3\})+ P(\{4\})=4/6=2/3$

Properties and algebra of events

- Sample spaces are formally described using sets and operators are the classical in Set Theory:
- Complement of event A denoted as $\overline{A}$ or A^C or $\neg A$
- Intersection of events A and B: contains outcomes belonging simultaneous to A and B $A \cap B$
- Union of events A and B: A U B contains outcomes belonging either to A or B or both.

Properties and algebra of events

Properties:

- 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: $A \cup \varnothing = A$ i $A \cap \Omega = A$.
- Complementation laws : $A \cup \neg A = \Omega$ i $A \cap \neg A = \varnothing$.
- De Morgan's laws $\overline{A} \bigcup \overline{B} = \overline{A \bigcap B}$

$$\overline{A} \bigcap \overline{B} = \overline{A \bigcup B}$$

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.

$$P:\ \wp(\Omega) \rightarrow [0,1]$$ (Domain set of all events)

  - For any event A, $P(A) \geq 0$
  - $P(\Omega) = 1$
  - $P(A \cup B) = P(A) + P(B)$, whenever A and B are mutually exclusive events
- 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.
- Probability of an event occurring is 1 minus probability that it does not occur.
- P(¬A)=1-P(A)

$$A \in \wp(\Omega) \Rightarrow \overline{A} \in \wp(\Omega) \Rightarrow A \cup \overline{A} = \Omega \ \text{ i } A \cap \overline{A} = \varnothing$$

$$\text{By } 3 \Rightarrow P(\Omega) = P\left(A \cup \overline{A}\right) = P(A) + P\left(\overline{A}\right)$$

$$\text{By } 2 \Rightarrow P(\Omega) = 1$$

$$\text{Thus}, P(\Omega) = 1 = P(A) + P\left(\overline{A}\right) \Rightarrow P\left(\overline{A}\right) = 1 - P(A)$$

- As a consequence, $P(\varnothing)=1-P(\Omega)=0$

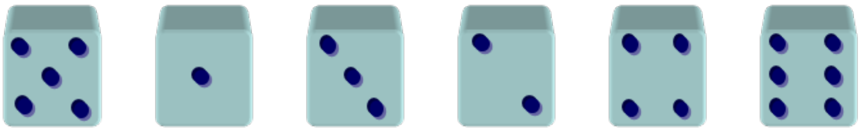
Example: all elemental events have equally probability

- You roll a fair six-sided die 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 die=1/6

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

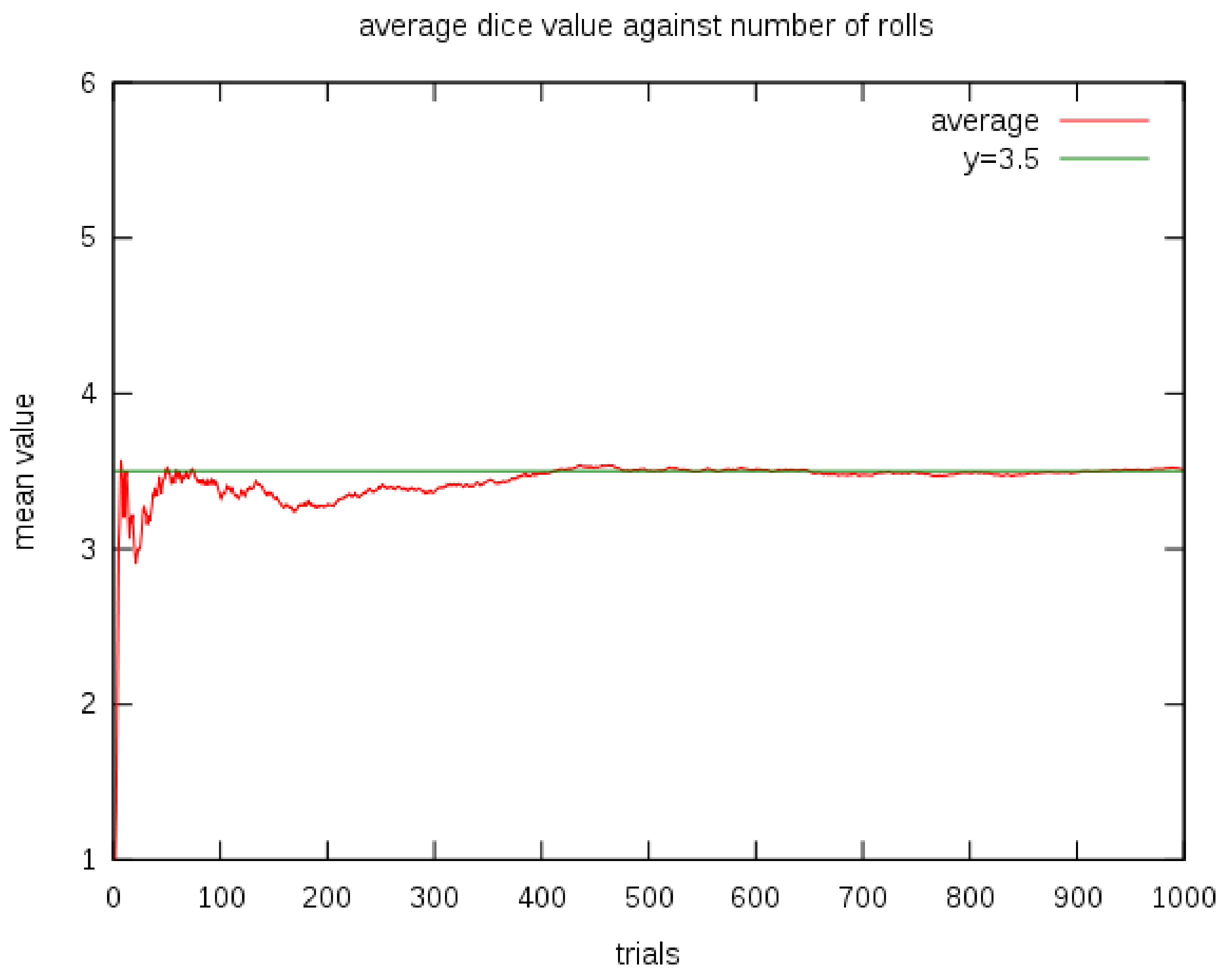
c) rolling a number less than 7
   - number of ways to roll less than 7/number of ways to roll the die=6/6=1

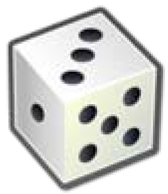


Terminology

- Empirical probability: 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
- Law of Large Numbers (Law of averages): 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





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: Sample Space-- Ω={BB, BG, GB}
- "I have two children. The older is a boy, and ….." What is the chance that I have two boys?
  - Ans: ½ Ω={BB, BG}
- "I have two children" What is the chance that I have two boys?
  - Ans: 1/4: Sample Space-- Ω={BB, BG, GB,GG}

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.
- Convinience notation: elementary events ω_i i=1,12

Our outcome

ω	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

- Ω={1T, 2T, 3T, 4T, 5T, 6T, 1H,2H,3H,4H,5H,6H}

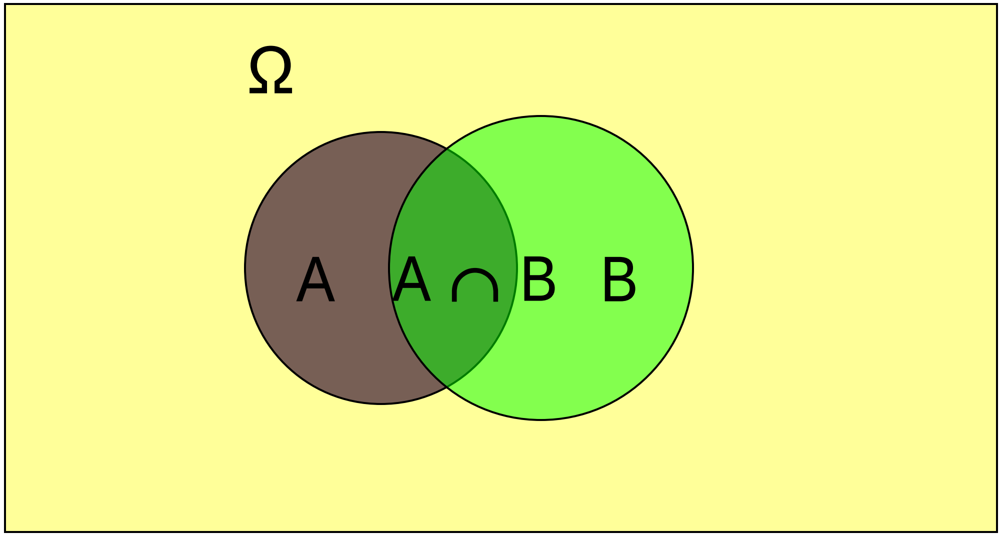
… of in independent events

- Let A be "Rolling a 5", A={ 5T, 5H }
- The probability of rolling a 5?
  - P(A)=P(5T)+P(5H)=1/12+1/12= 1/6
- Let B be " Getting a TAIL", B={1T, 2T, 3T, 4T, 5T, 6T}
- The probability of throwing a TAIL?
  - P(B)=P(1T)+…+P(6T)=1/12+…+1/12= 6/12=1/2
- Let C be an event "Rolling a 5 and getting a TAIL"
- The probability of C ?
  - P(C)= P({5 and TAIL})=1/12
  - Since $C=A\cap B$, $P(C)=P(A \cap B)=P(A)P(B)$ =1/6 x ½ = 1/12
  - **Does this property always hold? NO**

Computing probabilities

- General addition Rule:
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Venn Diagram



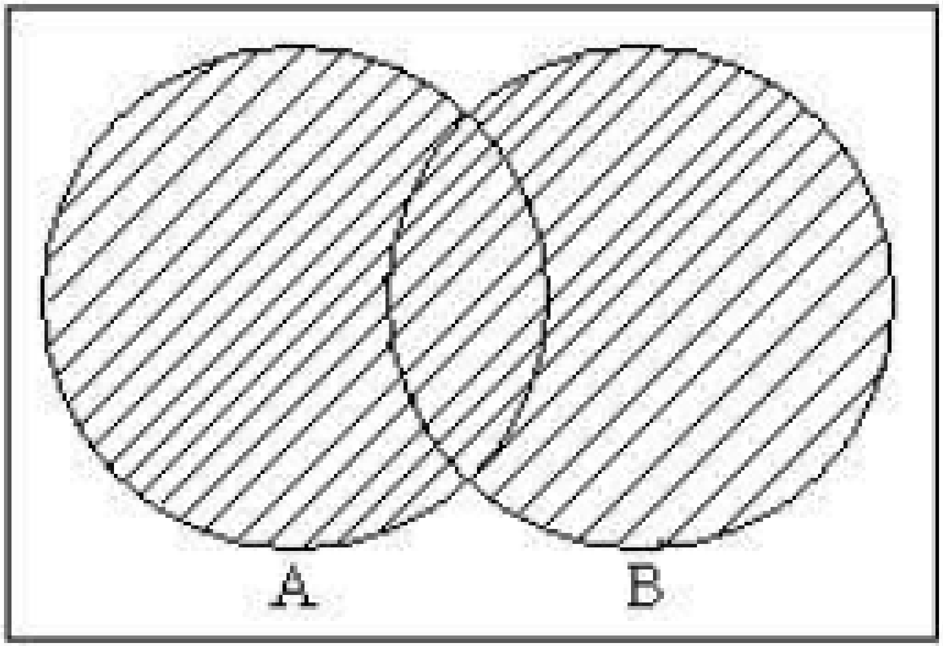
∪: union, or

∩: intersection, and

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

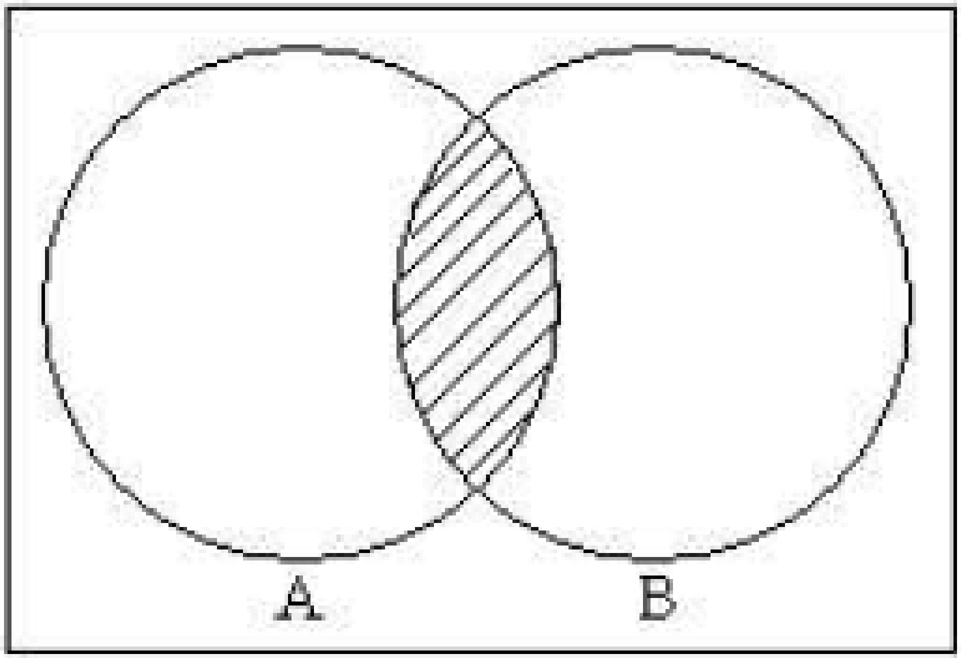
Union of sets

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



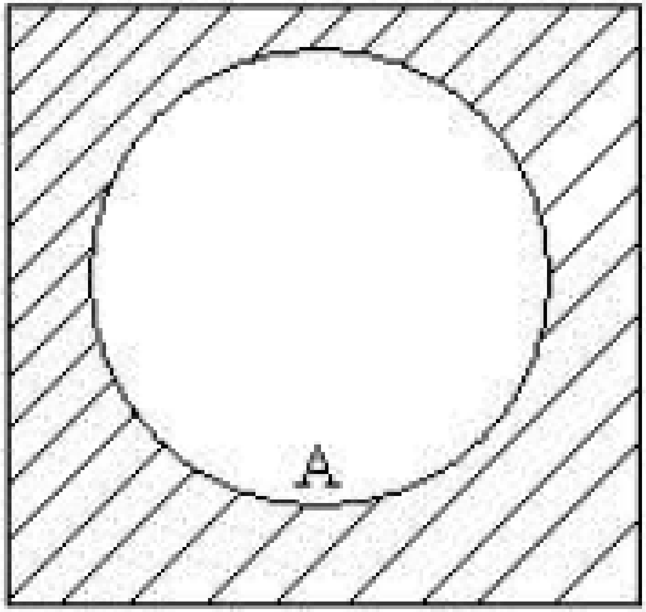
Intersection of Sets

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



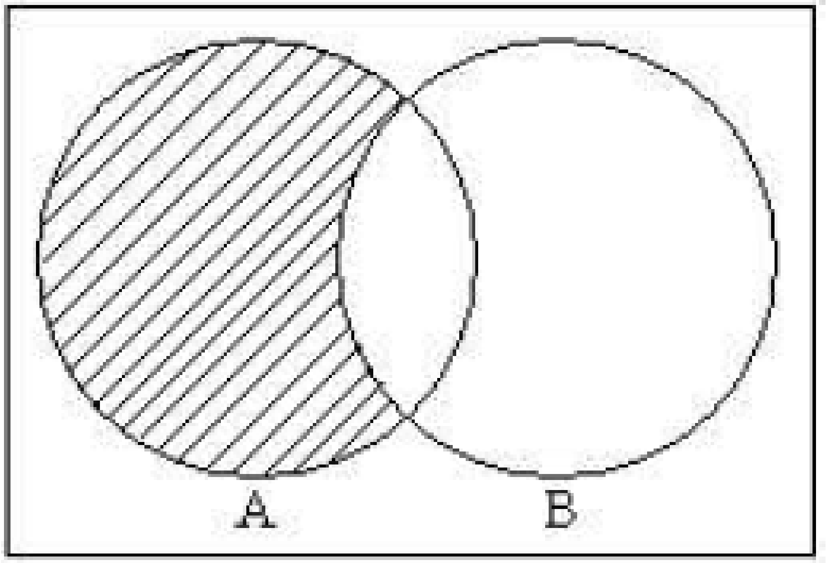
Complement

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



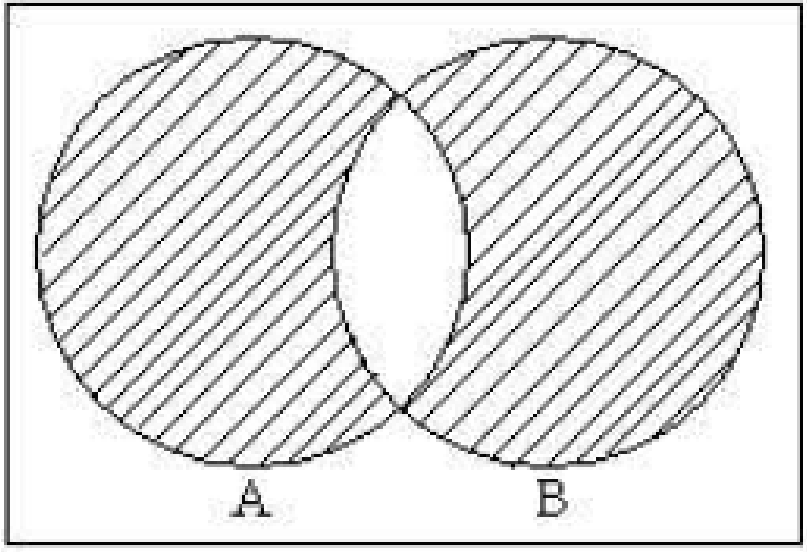
Difference

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



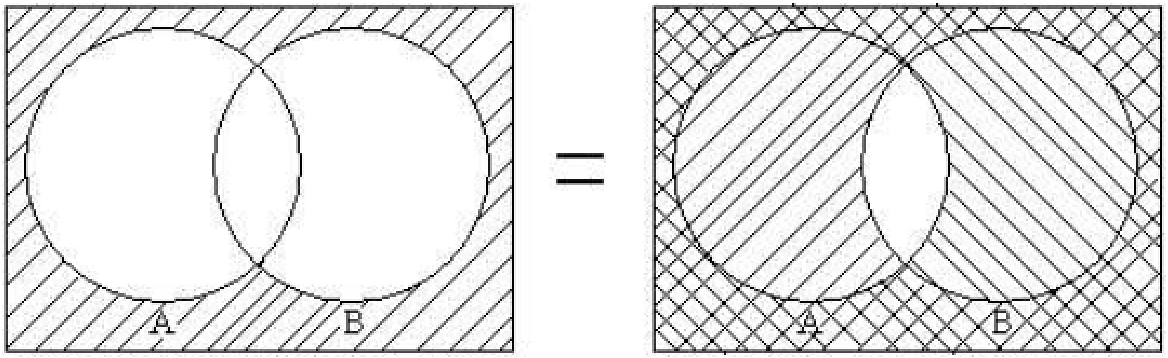
Symmetric Difference

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



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 $s \in B^c$
  - $s \in (A^c \cup B^c)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

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,
  $P(A \cup B) = P(A) + P(B)$

Two Schools

- Frequentists: fraction of times a event occurs if it is repeated “N” times
- Bayesians: a probability is a degree of belief

Conditional Probability

- We define conditional probability of A given B, as

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

where P(A|B) is the probability of event A, given that B has already happened

Assuming $P(B) > 0$

Conditional Probability: properties (I)

- Since

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

- Then,

$$\boxed{\begin{array}{l} \bullet\ P(A \cap B) = \begin{cases} P(A/B) \cdot P(B) \text{ if } P(B) > 0 \\ P(B/A) \cdot P(A) \text{ if } P(A) > 0 \end{cases} \\ \bullet\ \text{If } A \cap B = \varnothing \text{ then } P(A/B) = P(B/A) = 0 \end{array}}$$

Conditional Probability: properties (II)

- Since

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

- Then,

- If $A \subset B$ then $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} \geq P(A)$
- If $A \supset B$ then $P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$

Conditional probabilities

- Example: 2 fair coins are tossed –*at least one is a head* -- find the probability that both are heads:
  - Methodology, define Ω and the probability function P.
    - Sample Space Ω={2H, Both,2T}
    - Define the probability of the elementary events:
    - P(2H)= P(2T)= 1/4 , P(Both)=2/4
  - We are given some inside info B= {2H, Both}; P(B)=1/4+2/4=3/4
  - Determine A ={2H}; P(A)=1/4
- But $P(A|B)= P(A \cap B)/P(B)= P(A)/P(B)=(1/4)/(3/4)=1/3$

Independence

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

  $P(A \cap B) = P(A)\ P(B)$, if *A* and *B* are independent.
- Example (cont.): 2 fair coins are tossed –*at least one is a head* -- find the probability that both are heads
  - Are A =\{2H\} and B= \{2H, Both\} independent?
  - $P(A \cap B) = P(A)\ P(B)$?
  - $P(A) = \{2H\} = 1/4$; $P(B) = 3/4$ thus $P(A)P(B) = 3/16$
  - $P(A \cap B) = P(A) = 1/4$
- NO, thus dependent events

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.

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 group
- r <n, order matters
- Vn,r are the number of r ordered groups of different elements that can be formed with n elements.
  - $V_{n,r} = n (n-1) (n-2) \ldots (n-r+1)$
  - Example: five-letter 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 don't matters
- Cn,r is the number of R groups of different elements can be formed with n elements.

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

- 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
- Pn is the number of different ways to sort n elements.
  - Pn = 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
- Vn,r are the number of r ordered groups of different elements that can be formed with n elements, we can repeat the elements.
  - $V_{n,r} = n^r$
  - Example: different license plates that have only the digits 5,7,9.

Combinations with repetition

- n: number of elements
- r: number of elements in the groups
- r <n, order don't matters
- Cn,r is the number of R groups of different elements can be formed with n elements that can be repeated.

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

- Example: results when throwing 4 equal dices.

Permutations with repetition

- n: number of elements
- r: number of elements in the groups
- r=n, order matters
- Pn 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.

Example: 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? 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(same)=1-0.9917…= 0.0082…
- 4 w/ no sharing the date :
  - (365*364*363*362)/(365)^4 = 0.01635;
- Pattern p(no sharing the date):

$$\frac{365!/(365-n)!}{365^n}$$

Birthday problem

# people in rm	P(2 sharing birthday)
5	0.027
10	0.116
20	0.252
25	0.411
50	0.970
70	0.994
80	0.99991
90	0.999993

Birthday problem

- Thus, at 23 people, the P(2 people share birthday) = 0.5072
- We are 9, p(not sharing)=0,9, true?