Quantitative Methods – I (Statistics)

A. Y. 2024-25

Prof. Marco Stefanucci

Chapter 4
Probability

Probability: outline

1. Definitions and instruments: experiment, events, operations with events, Venn diagrams
2. Probability measurement
3. Joint, marginal, and conditional probabilities
4. Independency
5. Bayes' Theorem

Basic Elements of Probability Theory

Probability theory deals with random experiments, or random trials.

A random experiment is such that:

- the possible outcomes are known prior to its conduct
- the outcome is know at the end of the experiment and cannot be predicted with certainty in advance
- (can be repeated any number of times under the same conditions)

The last item is relevant only for one approach to assigning probabilities (the frequentist approach), as we shall see.

Basic Elements of Probability Theory

The main ingredients of a random experiment are:

- The sample space: a set Ω, whose elements ω correspond to the possible outcomes of the experiment.
- The family of events: a collection $\mathcal{F}$ of subsets $A \in \Omega$. We say that A occurs if the outcome ω of the experiment is an element of A.
- The probability measure: a function P with range $[0,1]$, defined on $\mathcal{F}$, and satisfying certain properties.

The triple $(\Omega, \mathcal{F}, P)$ is called a probability space.

Definitions

Experiment: process that, when performed, results in one of many uncertain outcomes

Ex. Toss of a coin, roll of a die, select a worker,…

Outcomes: results of an experiment

Ex.	Toss of a coin:	Head, Tail
	Roll of a die:	1, 2, 3, 4, 5, 6
	Select a worker:	Male, Female

Definitions

Sample space: **<u>all</u>** outcomes of an experiment, denoted by Ω

Experiment	Outcomes	Sample Space
Toss a coin once	Head, Tail	$S = \{\text{Head, Tail}\}$
Roll a die once	1, 2, 3, 4, 5, 6	$S = \{1, 2, 3, 4, 5, 6\}$
Toss a coin twice	HH, HT, TH, TT	$S = \{HH, HT, TH, TT\}$
Play lottery	Win, Lose	$S = \{\text{Win, Lose}\}$
Take a test	Pass, Fail	$S = \{\text{Pass, Fail}\}$
Select a worker	Male, Female	$S = \{\text{Male, Female}\}$

Sample Space

The set Ω can be

- finite (it contains a finite number of elements ω), or
- countable, i.e., it can contain a countably infinite set of elements (e.g., number of clients), or
- uncountable (e.g., a subset of real numbers, like the waiting time at the bus station)

Ex. Toss a coin until Head occurs for the first time. The sample space is $\Omega = \{H, TH, TTH, TTTH, \ldots\}$ and contains a countably infinite number of elementary events.

Experiment and Sample Space: examples

Experiment	Sample space
Toss of a coin	$\{H, T\}$
Two successive tosses of a coin	$\{(HH),(HT),(TT),(TH)\}$
Rolling a die	$\{1,2,3,4,5,6\}$
Rolling two dice	$\{(1,1),(1,2),\ldots,(1,6),$ $(2,1),\ldots,(2,6),\ldots,$ $(6,1),\ldots,(6,2),\ldots,(6,6)\}$
University exam	$\{\text{Fail }(<18),18,19,\ldots,30,30\text{ c.l.}\}$
Number of clients in a day	$\{0,1,2,\ldots,\}$
Duration of a bulb	$R^+ = [0,\infty)$
Daily return of a stock	$R = (-\infty,\infty)$

Events

An event is a subset of Ω.

- **Elementary event**: one of the possible outcomes of the experiment.
  It will be denoted by w.

$$\Omega = \{w_1, w_2, w_3, \ldots\}$$

Elementary events $w \in \Omega$ are disjoint (mutually exclusive).

For instance, if you toss a coin, either Head or Tail can occur, but not both.

Compound (or composite) event

A compound or composite event is obtained by combining elementary events. An event A is a subset of a sample space Ω, $A \subset \Omega$, obtained by combining its elementary events.

We say that the event A occurs if the outcome of the experiment, $w \in \Omega$, is an element of the set A.

Example: rolling of a die.

$$\Omega = \{w_i = i, i = 1, \dots, 6\}.$$

The event "an even number on the upper side of a die" is the composite event: $A = \{2, 4, 6\}$

Simple and Compound event

Event (or simple event): 1 of final outcomes for an experiment, denoted by E_i

Compound event: collection of 2 or more outcomes for an experiment, denoted by capital letters A, B, C, …

Ex: in rolling a die once, we have 6 *simple* events:

$\boldsymbol{E_1} = \{1\}$, $\boldsymbol{E_2} = \{2\}$, $\boldsymbol{E_3} = \{3\}$, $\boldsymbol{E_4} = \{4\}$, $\boldsymbol{E_5} = \{5\}$, $\boldsymbol{E_6} = \{6\}$

$\boldsymbol{A}$ = "An even number is obtained",

$\boldsymbol{A} = \{2, 4, 6\}$, is a *compound* event

Union on sets

The symbol ∪ (or OR) is employed to denote the union of two sets. Thus, the set A ∪ B - read "A union B" or "the union of A and B" - is defined as the set that consists of all elements belonging to either set A **OR** set B (**or both**).

Ex. suppose that ***Committee A***, consisting of the 5 members *Jones, Blanshard, Nelson, Smith*, and *Hixon*, meets with ***Committee B***, consisting of the 5 members *Blanshard, Morton, Hixon, Young*, and *Peters*. The union of Committees A and B must then consist of **8 members** rather than 10 (Jones, Blanshard, Nelson, Smith, Morton, Hixon, Young, and Peters).

Intersection on sets

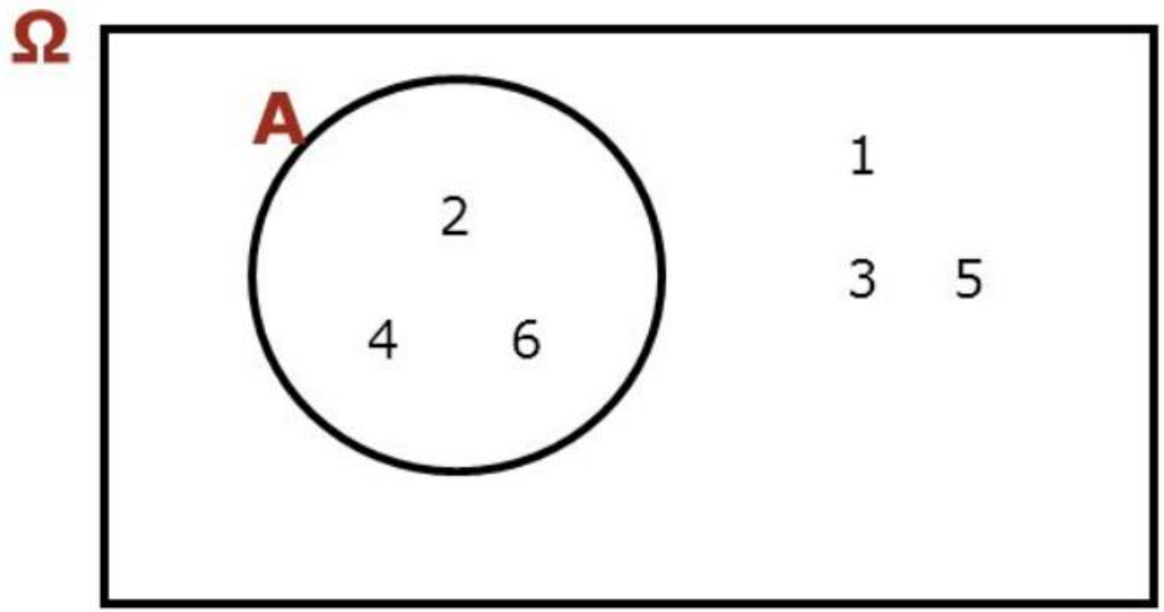
The intersection operation is denoted by the symbol ∩ (or AND). The set A ∩ B - read "A intersection B" or "the intersection of A and B" - is defined as the set composed of all elements that belong to both A **AND** B.

Ex. suppose that ***Committee A***, consisting of the 5 members *Jones, Blanshard, Nelson, Smith*, and *Hixon*, meets with ***Committee B***, consisting of the 5 members *Blanshard, Morton, Hixon, Young*, and *Peters*. The intersection of the two committees is the set consisting of **2 members** (Blanshard and Hixon).

Venn Diagrams

Picture that depicts all the possible outcomes for an experiment.

This Venn Diagram shows Ω and the event A="*An even number is obtained*" for rolling a die once

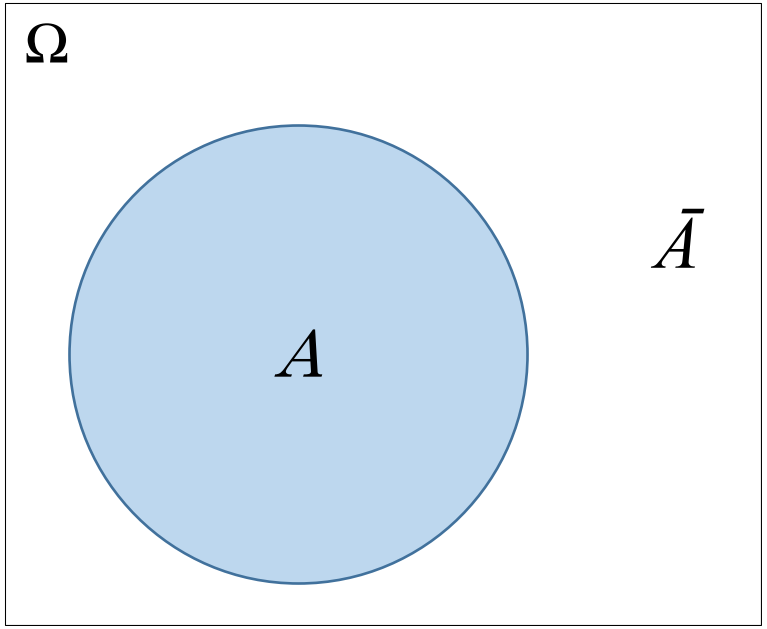


Algebra of Events

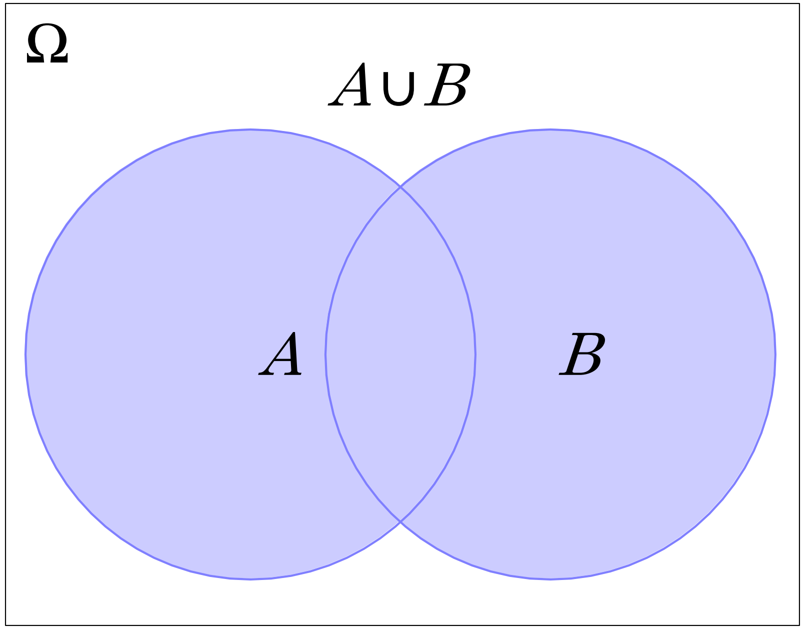
We introduce a set of operations and rules that are used to generate other events.

The operations between events can be visualized using Venn diagrams.

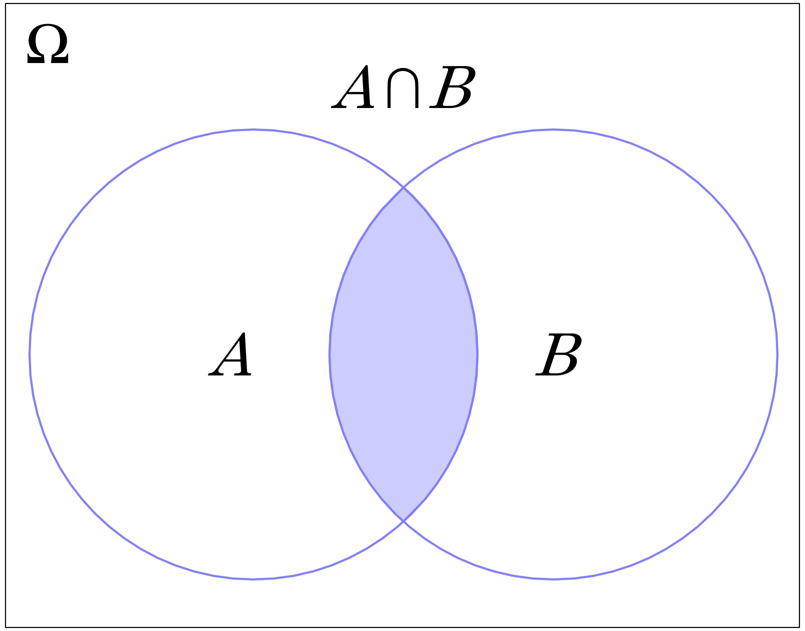
Complementary event $\bar{A}$



Union



Intersection



Algebra of Events: Complementary event

Let A and B be two arbitrary events defined in Ω

Complementary (contrary) event. We denote by $\bar{A}$ the event occurring when A does not occur. It contains all the elementary events $\omega \in \Omega$ that do not belong to A:

$$\bar{A} = \{\omega \in \Omega : \omega \notin A\}$$

Algebra of Events: Intersection and Union

Let A and B be two arbitrary events defined in Ω

Intersection. The intersection of A and B, denoted $A \cap B$, is the event occurring if both the events A and B occur. Hence,

$A \cap B = \{\omega \in \Omega : \omega \in A \text{ and } \omega \in B\}$ as a result, $\omega \in A \cap B$ if $\omega \in A$ and $\omega \in B$

Union. The union of A and B, denoted $A \cup B$, is the event occurring if either A or B occurs. It consists of all the elementary events that belong to A or B or to both. Hence,

$A \cup B = \{\omega \in \Omega : \omega \in A \text{ or } \omega \in B\}$ as a result, $\omega \in A \cup B$ if $\omega \in A$ or $\omega \in B$

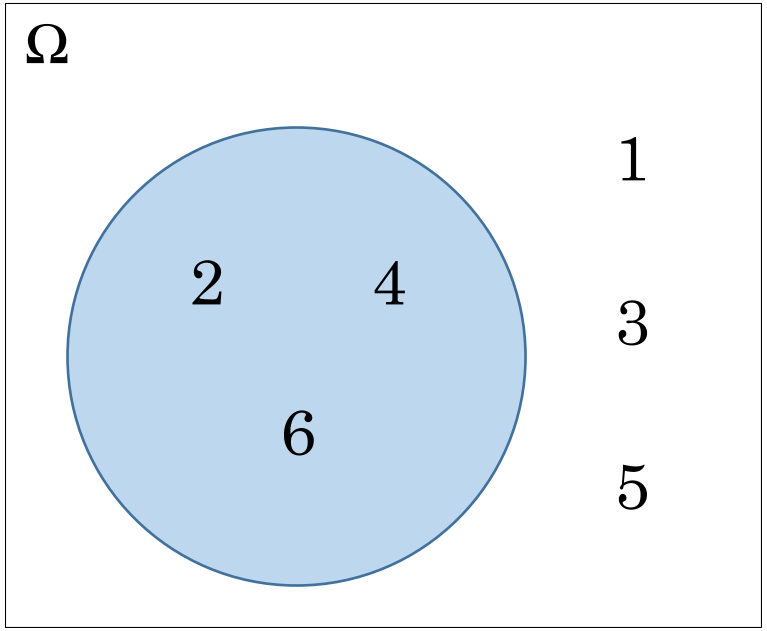
Complementary events

Complementary event: collection of all outcomes not in A, denoted by $\bar{A}$

Ex: roll a die once and define:

A= an even number is observed= $\{2, 4, 6\}$

$\bar{A}$= an odd number is observed= $\{1, 3, 5\}$



Union of events

To calculate the union of two or more sets, we combine the elements within each set. Duplicated values must only be counted once.

The symbol ∪ represents the union of sets.

Ex. The union of the set of even numbers

E={2, 4, 6, 8, 10} and odd numbers O={1, 3, 5, 7, 9} is the set

$$E \cup O = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

Intersection of events

To calculate the intersection of two or more sets, we calculate the number of values that are contained within both / all of the sets only.

The symbol ∩ represents the intersection of sets.

For the two sets A and B, $A \cap B$ is pronounced A intersection B.

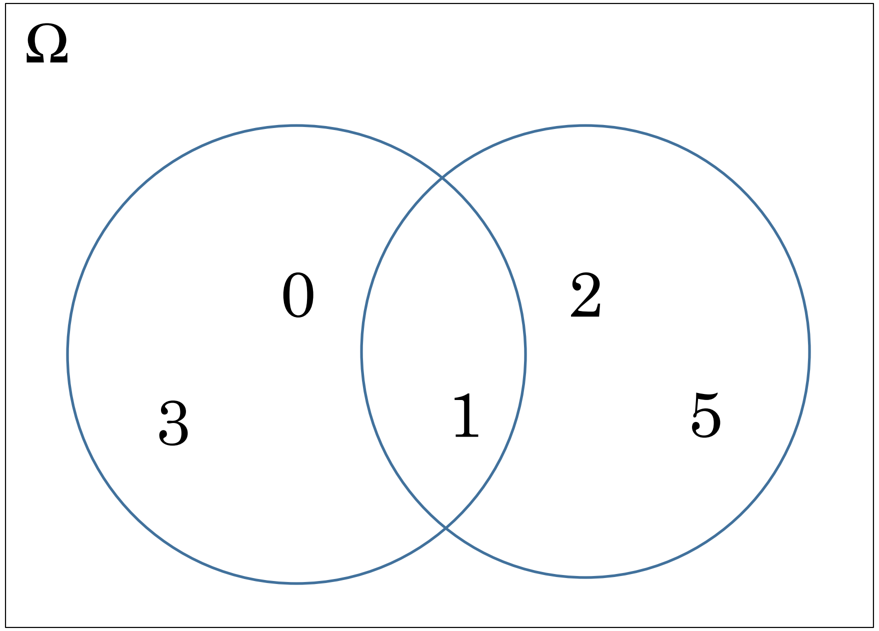
Ex. The intersection of the set of odd numbers $O = \{1, 3, 5, 7, 9\}$ and the set of prime numbers $P = \{2, 3, 5, 7, 11\}$ is the set

$$O \cap P = \{3, 5, 7\}.$$

Union and Intersection: example

Given

$$A = \{0,\ 1,\ 3\}$$
$$B = \{1,\ 2,\ 5\}$$



Then,

- *Union:* $A \cup B = \{0, 1, 2, 3, 5\}$
- *Intersection:* $A \cap B = \{1\}$

Properties of the operations between events

- A composite event can be written as the union of elementary events.
- Unions and intersections are related by the de Morgan's laws

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

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

- The complement of Ω is the impossible event, denoted Ø (empty space).
- The certain event is the event which always occurs. It is complementary to Ø and contains all the possible outcomes of the experiment. Moreover, $A \cup \bar{A} = \Omega$.

Properties of the operations between events

	Union	Intersection
Idempotency	$A \cup A = A$	$A \cap A = A$
Neutral event	$A \cup \emptyset = A$	$A \cap \Omega = A$
Commutativity	$A \cup B = B \cup A$	$A \cap B = B \cap A$
Associativity	$(A \cup B) \cup C = A \cup (B \cup C)$	$(A \cap B) \cap C = A \cap (B \cap C)$

Mutually exclusive (disjoint) events

- Two events are said to be **mutually exclusive**, or **disjoint**, if they cannot occur simultaneously, i.e. their intersection is the empty set: $A \cap B = \emptyset$.
- An event A is said to be **included** in B, or a subset of B, if $A \cap B = A$, in which case we write $A \subset B$.
- A collection of events A_i, $i = 1, 2, \ldots, m$, are called a **partition** of the sample space Ω if they are all disjoint,

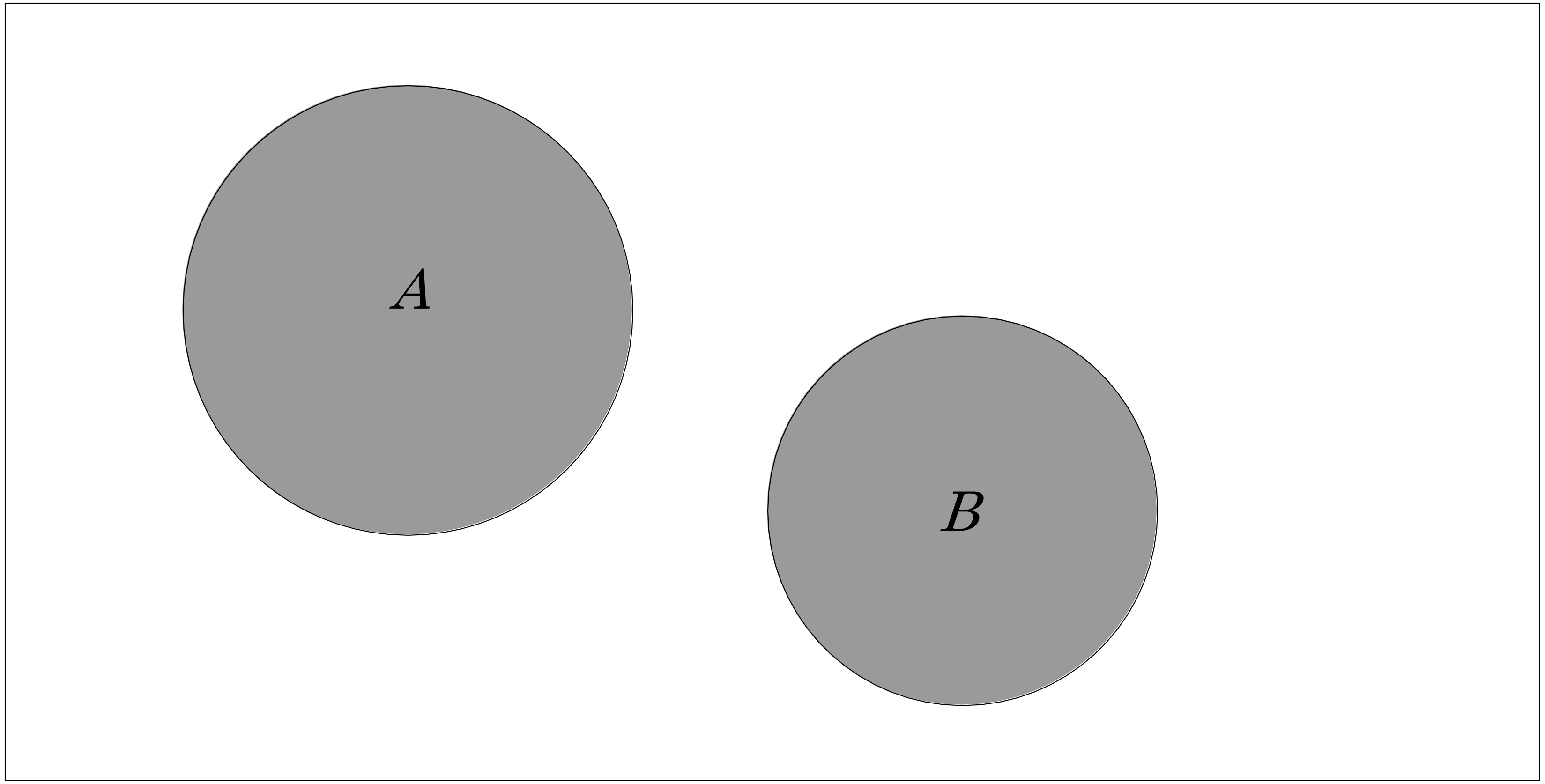
$$A_i \cap A_j = \emptyset, \ \forall \ i \neq j$$

and their union is the sample space,

$$A_1 \cup A_2 \cup \cdots \cup A_m = \Omega$$

Mutually exclusive (disjoint) events

Disjoint events: $A \cap B = \emptyset$



Mutually exclusive (disjoint) events

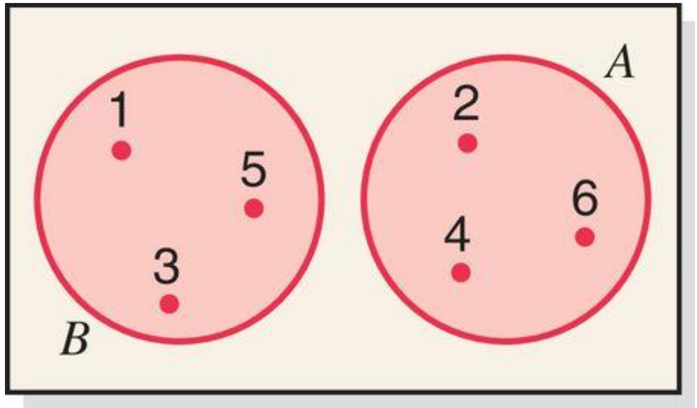
Events that cannot occur together

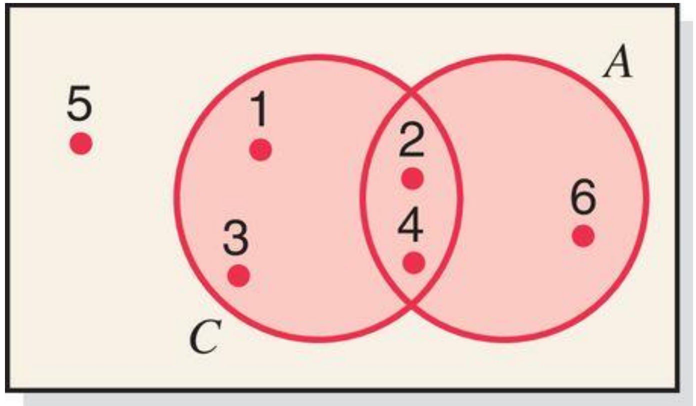
Ex: roll a die once and define:

A= an even number is observed= {2, 4, 6}

B= an odd number is observed= {1, 3, 5}

C= a number less than 5 is observed= {1, 2, 3, 4}





A and B mutually exclusive

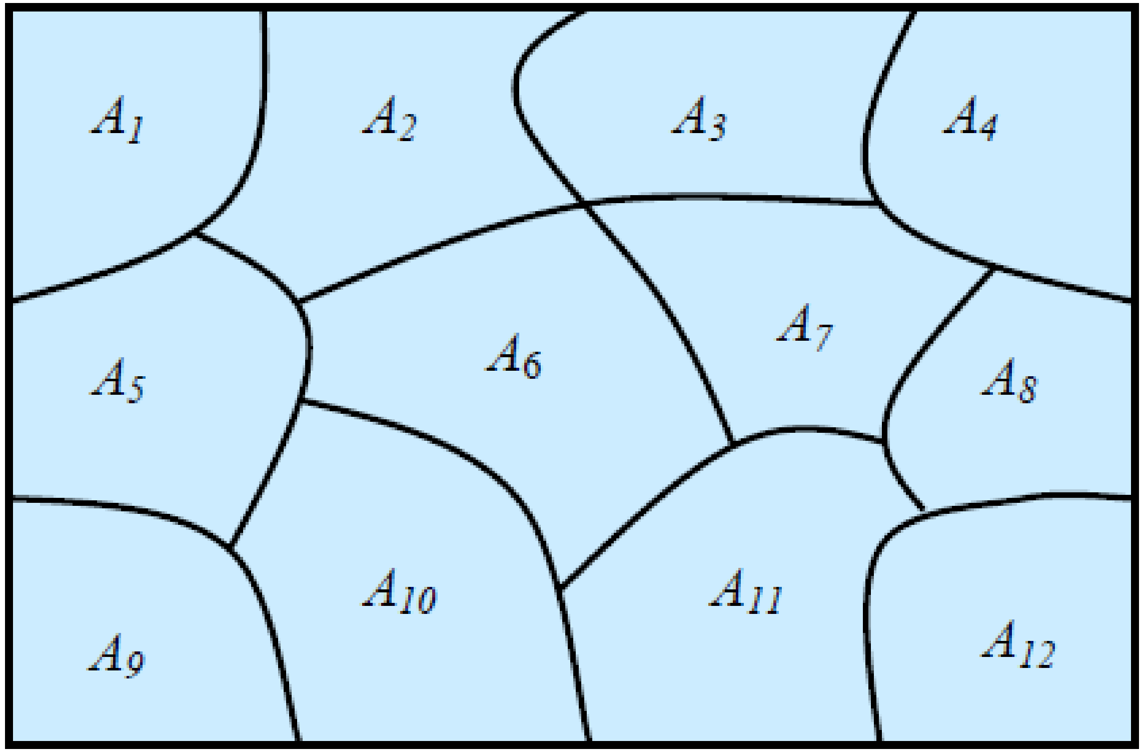
A and C not mutually exclusive

Partition of events

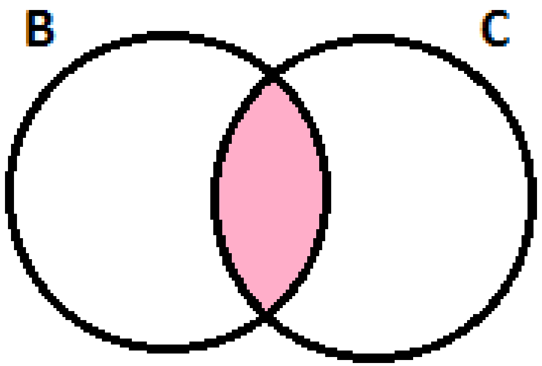
Collection of events that are mutually disjoint and such that their union is the entire sample space:

1) $A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n = \emptyset$

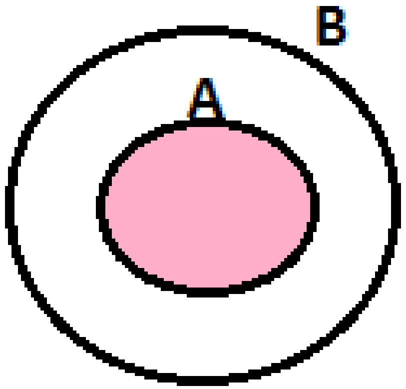
2) $A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n = \Omega$



Algebra of events: examples

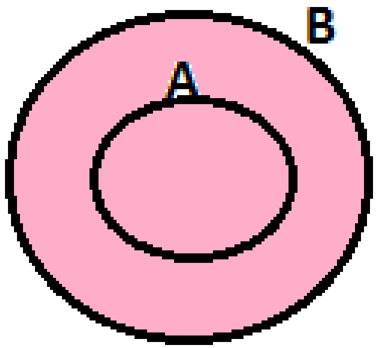


B∩C



B∩A

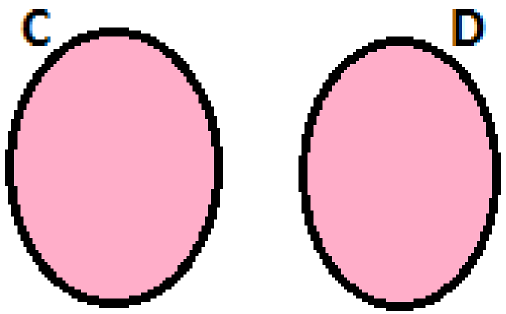
A ⊂ B



A U B

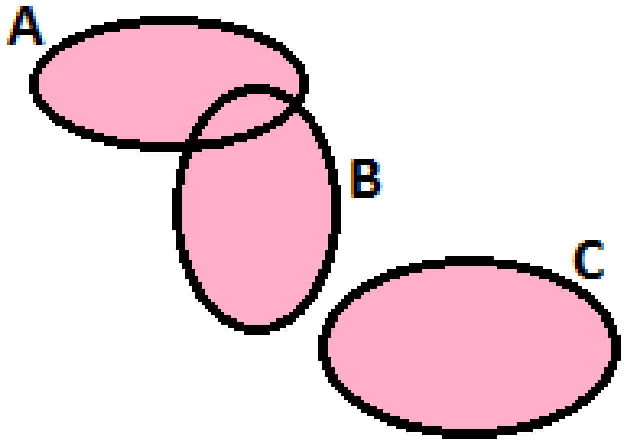
A ⊂ B

Algebra of events: examples

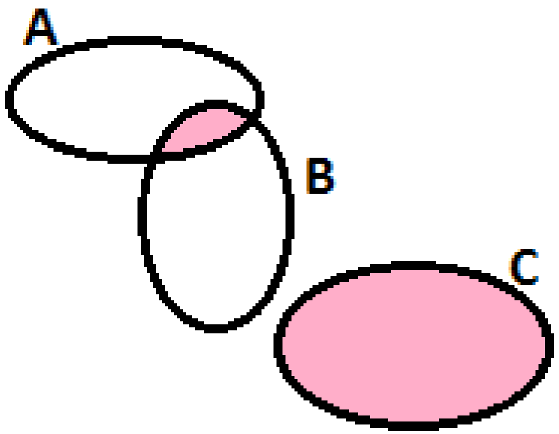


C U D

Disjoint events

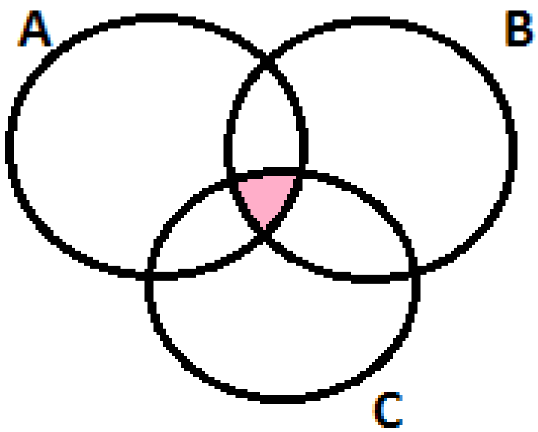


A U B U C

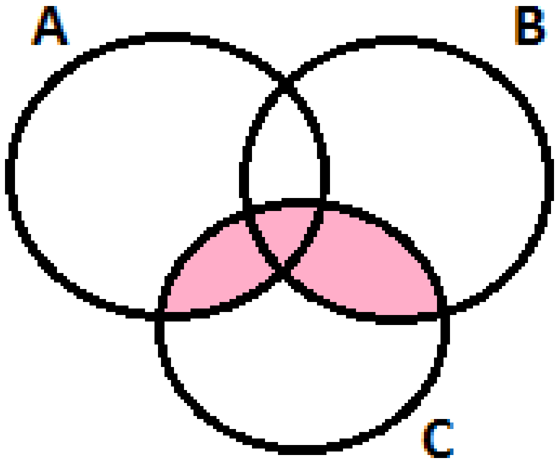


(A∩B) UC

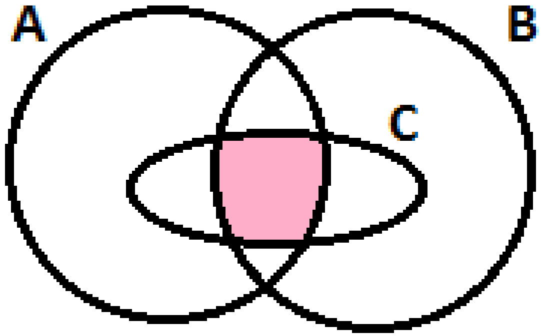
Algebra of events: examples



(A∩B)∩C



(A∩C) U (B∩C)



(A∩B)∩C

Counting rule

When the experiment has a large number of outcomes, it may not be easy to list them all ➔ to count them use the counting rule

If an experiment consists of k steps, each with n_i outcomes, then the total number of final outcomes is

$$n_1 \times n_2 \times n_3 \times \cdots \times n_k$$

Ex: the experiment of tossing a coin (2 outcomes) 5 times has

$$2 \times 2 \times 2 \times 2 \times 2 = 2^5 = 32$$

Counting rule: Example

Ex: A car buyer can choose between a fixed and a variable interest rate and can also choose a payment period of 36 months, 48 months, or 60 months. How many total outcomes are possible?

How many steps?
2 (choosing the interest rate and choosing the period)

How many outcomes each?
2 for step 1 (interest rate), 3 for step 2 (periods)

$$2 \times 3 = 6$$

Factorial

The factorial of a number is obtained by multiplying all the integers from that number to 1

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 1$$

Ex: the factorial of 5 is:

$$5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$

NB: conventional rule $0! = 1$

Permutation

Permutation is basically called as a arrangement where order **does matters**.

Here we need to arrange the digits, numbers, alphabets, colors and letters taking some or all at a time.
It is represented as,

$$P_{n,x} = \frac{n!}{(n-x)!}$$

Combination

Combination is basically called as a selection where order **does not matters**.

Here we need to arrange the digits, numbers, alphabets, colors and letters taking some or all at a time.
It is represented as,

$$C_{n,x} = \frac{n!}{x!\,(n-x)!} \quad or \quad C_{n,x} = \binom{n}{x}$$

1. $n \geq x$
2. $C_{n,n} = 1$
3. $C_{n,0} = 1$

Combination and Permutation

Permutation is an arrangement of objects **in a definite order**.

Number of all permutations of n things, taken x at a time, is given by

$$P_{n,x} = \frac{n!}{(n-x)!}$$

Combination is selection of objects where **order does not matter**.

Number of all combinations of n things, taken x at a time, is given by

$$C_{n,x} = \frac{n!}{x!\,(n-x)!}$$

Combinations: examples

Ex: 3 members of a faculty committee need to be randomly chosen from a set of 5. How many combinations are possible?

$$C_{n,x} = \frac{n!}{x!\,(n-x)!}$$

$$C_{5,3} = \frac{5!}{3!\,(5-3)!} = 10$$

Check: denote the 5 candidates as: A B C D E

The possible combinations are:

ABC , ABD , ABE , ACD , ACE , ADE , BCD , BCE , CDE , BDE

NB: order not important ➔ ABC is equal to ACB and BCA.

Permutations: example

Ex: 3 members needs to be randomly chosen from a set of 5. How many permutations are possible?

$$P_{n,x} = \frac{n!}{(n-x)!}$$

$$P_{5,3} = \frac{5!}{(5-3)!} = \frac{120}{2} = 60$$

Permutation with repetition

A **permutation with repetition** (it is possible to select many times the same unit) is equal to

$$P_{n,x} = n^x$$

Ex: 3 members needs to be randomly chosen from a set of 5 with repetition. How many permutations are possible?

$$P_{n,x} = n^x$$

$$P_{5,3} = 5^3 = 5 \times 5 \times 5 = 125$$

Combination and Permutation

	Repetition	Non-Repetition
Ordered **Permutation**	tuples: n^k	k-permutation: $\frac{n!}{(n-k)!}$
Unordered **Combination**	combination with rep: $\binom{k+n-1}{n-1}$	combination: $\binom{n}{k}$

Probability

Measures the chances that an event will occur

Basis for inferential statistics: inference is used to take decisions under uncertainty ➔ Probability evaluates the uncertainty involved in those decisions.

Probability definition and its measure

Given a collection of events defined on the sample space Ω, probability P(.) is a function, which assigns a number to every event in the collection according to the following rules (axioms of probability):

1. $P(A) \geq 0$
2. $P(\Omega) = 1$
3. $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$ (additivity)

Probability: properties

$P(A)$ = probability of event A, i.e. the likelihood of occurrence of A

Main properties

1. $\mathrm{P}(\Omega) = 1$ ➔ certain event
2. $0 \leq \mathrm{P}(\mathrm{A}) \leq 1$
3. $\mathrm{P}(\overline{\mathrm{A}}) = 1 - \mathrm{P}(\mathrm{A})$
4. $\mathrm{P}(\emptyset) = P(\overline{\Omega}) = 1 - 1 = 0$ ➔ impossible event

Probability: theorems and corollaries

These are simple consequences of the axioms of probability:

- $0 \leq P(A) \leq 1$
- $P(\emptyset) = 0$
- $P(\bar{A}) = 1 - P(A)$
- If $B \subseteq A$, then $P(B) \leq P(A)$
- If $P(B) = 1$, then $P(B \cap A) = P(A)$
- If $P(B) = 0$, then $P(B \cup A) = P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Probability: Classical approach

How to compute $P(A)$ depends on the approach:

1. Classical ➔ $P(A) = \frac{Number\ of\ cases\ favourable\ to\ A}{Number\ of\ cases\ possible}$

Ex: roll a die once, and find the probability of A="even number", B="number less than 5", C="number different from 2".

$P(A) = 3/6 = 0.5$ $\quad$ $P(B) = 4/6 = 0.67$ $\quad$ $P(C) = \frac{5}{6} = 0.83$

Probability: Frequentist approach

2. Frequentist: probability is the proportion of times that the event occurs in a long run of observations.

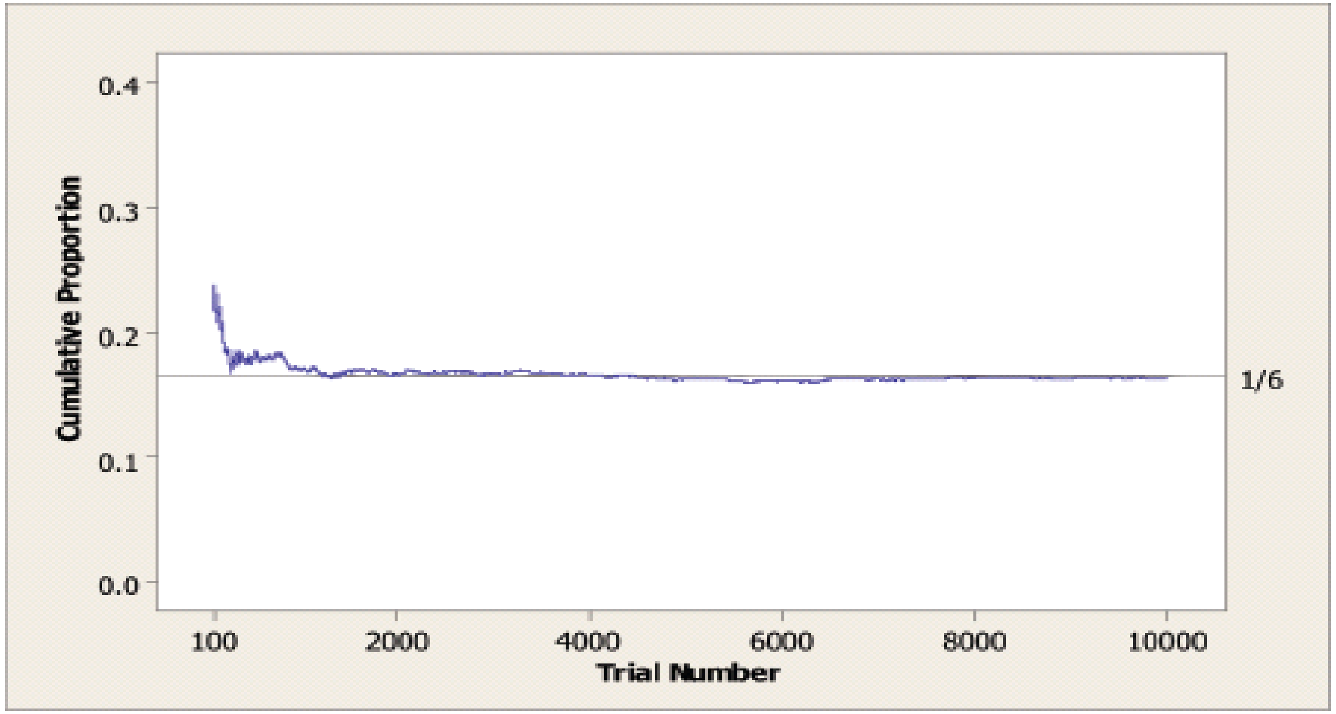
If in the short-run, the proportion of times that something happens is highly random…

Probability: Frequentist approach

2. Frequentist: probability is the proportion of times that the event occurs in a long run of observations

…in the long-run, this proportion becomes very predictable (Law of Large Numbers).

E.g.: keeping tossing a die, the share of 6 obtained will progressively converge to… 1/6



Probability: Subjective approach

3. Subjective → based on individual belief, experience, information

Ex: the probability that each one of you attaches to the event "*I will get 30 as final mark in Statistics*"

Bayesian statistics is a branch of statistics that uses subjective probability as its foundation

Joint probabilities

Ex: roll a die once and consider

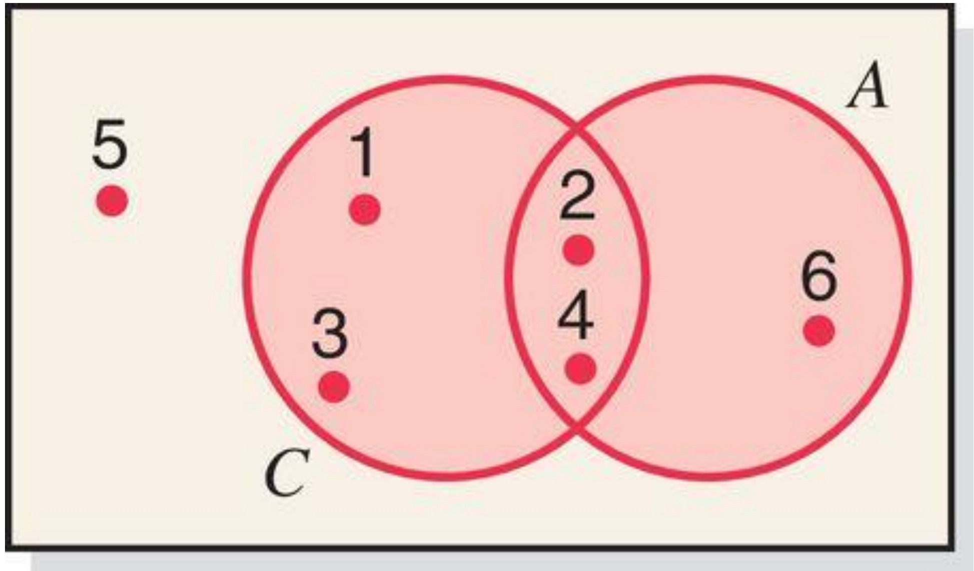
A= an even number is observed= {2, 4, 6}

B= an odd number is observed= {1, 3, 5}

C= a number less than 5 is observed= {1, 2, 3, 4}

The ***joint probability*** of A and C is

$$P(A \cap C) = \frac{2}{6} = 0.333$$



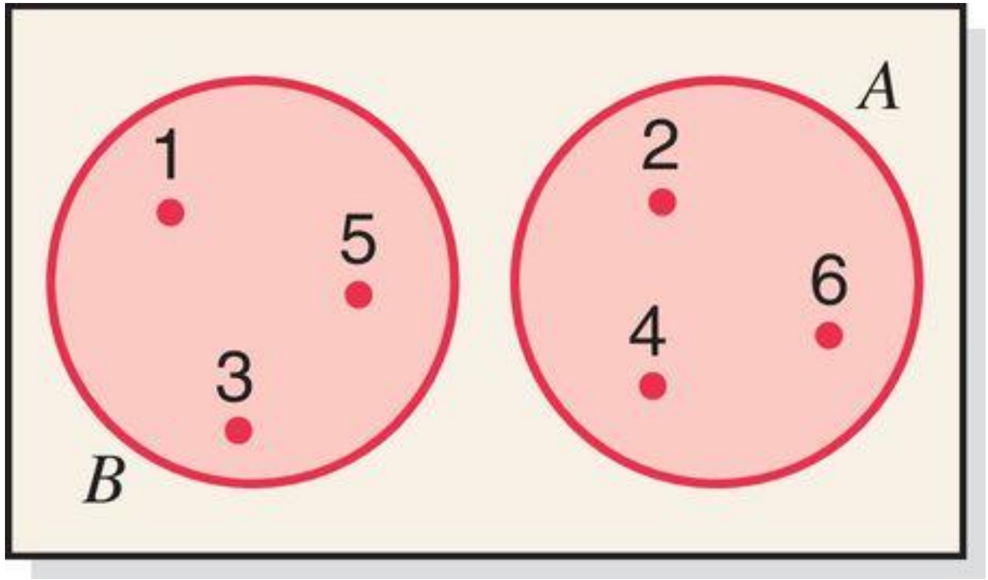
Joint probabilities: mutually exclusive events

Ex: roll a die once and consider

A= an even number is observed= {2, 4, 6}

B= an odd number is observed= {1, 3, 5}

C= a number less than 5 is observed= {1, 2, 3, 4}



The ***joint probability*** of A and B is

$$P(A \cap B) = \mathrm{P}(\emptyset) = 0$$

Conditional probabilities

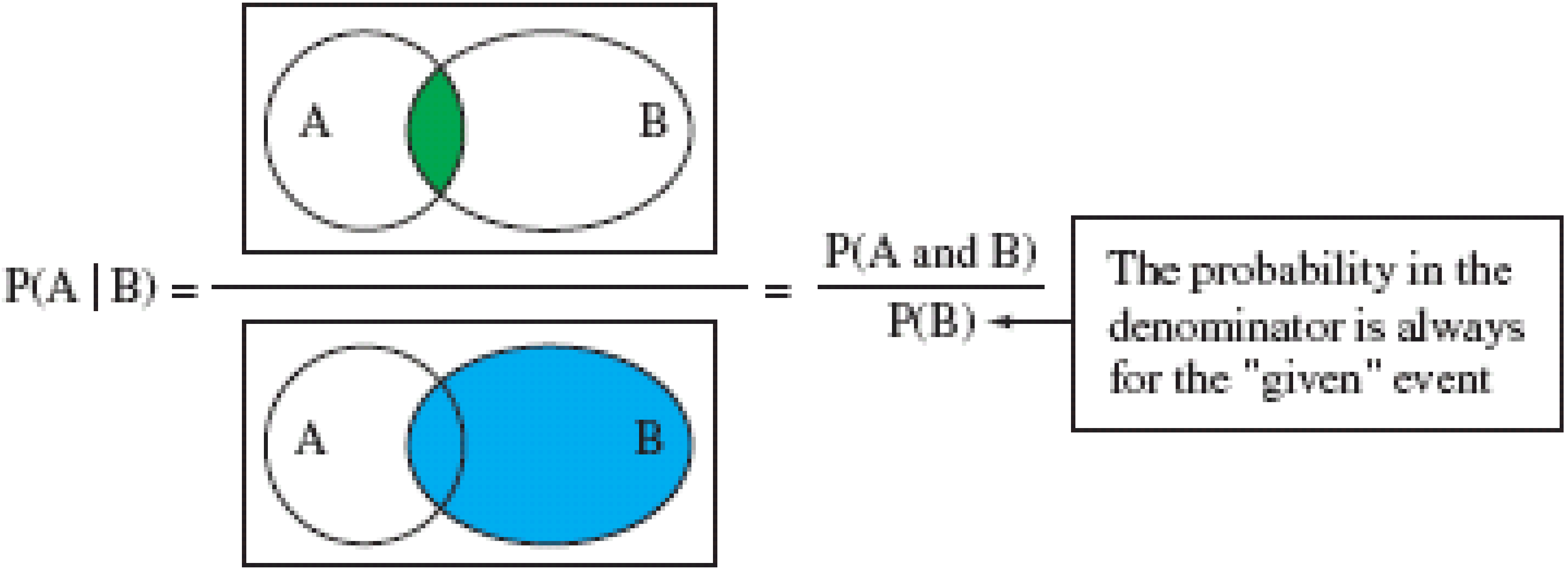
We aim at defining the probability of an event, given prior information on the occurrence of another event.

Suppose that we know that the event B, with $P(B) > 0$, has already occurred. We define the conditional probability of the event A, given that event B has already occurred, as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) > 0$$

Conditional probabilities

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) > 0$$



Conditional probabilities: example

$p(A/B)$

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

$A = \{1, 3, 5\}$

$B = \{3, 4, 5\}$

$p(A/B)$

Conditional probabilities: example

$$p(A|B) \qquad S = \{1, 2, 3, 4, 5, 6\}$$

$$A = \{1, 3, 5\} \qquad B = \{3, 4, 5\}$$

$$p(A|B) = \frac{2}{3} \qquad A \cap B = \{3, 5\} \leftarrow 2$$

Conditional probabilities: example

$$p(A/B) \qquad S = \{1, 2, 3, 4, 5, 6\}$$

$$A = \{1, 3, 5\} \qquad B = \{3, 4, 5\}$$

$$p(A/B) = \frac{2}{3} \qquad A \cap B = \{3, 5\} \leftarrow 2$$

$$p(A/B) = \frac{p(A \text{ and } B)}{p(B)}$$

Conditional probabilities: example

$p(A|B)$

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

$A = \{1, 3, 5\}$ $\quad B = \{3, 4, 5\}$

$p(A|B) = \frac{2}{3}$ $\quad A \cap B = \{3, 5\} \leftarrow 2$

$$p(A|B) = \frac{p(A \text{ and } B)}{p(B)} = \frac{\frac{2}{\cancel{6}} \cdot \cancel{6}}{\frac{3}{\cancel{6}} \cdot \cancel{6}} = \frac{2}{3}$$

Conditional probabilities: example

We toss a fair coin three successive times. We wish to find the conditional probability $\mathbf{P}(A \mid B)$ when A and B are the events

$$A = \{\text{more heads than tails come up}\}, \qquad B = \{\text{1st toss is a head}\}.$$

Conditional probabilities: example

We toss a fair coin three successive times. We wish to find the conditional probability $\mathbf{P}(A \mid B)$ when A and B are the events

$$A = \{\text{more heads than tails come up}\}, \qquad B = \{\text{1st toss is a head}\}.$$

The sample space consists of eight sequences,

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\},$$

which we assume to be equally likely. The event B consists of the four elements HHH, HHT, HTH, HTT, so its probability is

$$\mathbf{P}(B) = \frac{4}{8}.$$

Conditional probabilities: example

$$\mathbf{P}(B) = \frac{4}{8}.$$

The event $A \cap B$ consists of the three elements outcomes HHH, HHT, HTH, so its probability is

$$\mathbf{P}(A \cap B) = \frac{3}{8}.$$

Thus, the conditional probability $\mathbf{P}(A \mid B)$ is

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \frac{3/8}{4/8} = \frac{3}{4}.$$

Conditional probabilities: example

If you roll a fair die twice and observe the numbers that face up, find the probability that the sum of the numbers is 8, given that the first number is 3.

Conditional probabilities: example

Solution We begin by recalling that the sample space when we roll a fair die twice is the set $S = \{(1, 1), (1, 2), \ldots, (6, 6)\}$ containing the 36 different equally likely outcomes.

The two events under consideration are

A: The sum of the numbers is 8.

B: The first number is 3.

We also need

$A \cap B$: The sum of the numbers is 8 and the first number is 3.

But this can only happen in one way: $A \cap B = \{(3, 5)\}$. From the formula, then,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/36}{6/36} = \frac{1}{6}.$$

Example

A box contains 3 blue pens and 5 red pens. What is the probability of getting 2 red pens if they are extracted at once?

Let A =«2 red pens obtained», then

$$P(A) = \frac{Number\ of\ cases\ favourable\ to\ A}{Number\ of\ possible\ \ cases} = \frac{C_{5,2}}{C_{8,2}} = \frac{10}{28} = \frac{5}{14}$$

Same example with conditional probabilities

A box contains 3 blue pens and 5 red pens. What is the probability of getting 2 red pens if they are extracted at once?

A =«2 red pens obtained», but let's define also two single events

R_1=«the 1° pen is red», R_2=«the 2° pen is red»,

$$P(A) = P(R_1)P(R_2|R_1) = \frac{5}{8}\frac{4}{7} = \frac{5}{14}$$

Independent events

A and B are independent if:

$$P(A|B) = P(A) \qquad \text{or} \qquad P(B|A) = P(B)$$

Ex: in our example, since

$P(A|B) = 0.79 \neq P(A) = 0.6$ ➔ A and B are dependent

Clearly $P(B|A) = 0.25 \neq P(B) = 0.19$

Notice:

If A and B are dependent ➔$P(A \cap B) = P(A|B)P(B)$

If A and B are independent ➔$P(A \cap B) = P(A)P(B)$

Independent events: example

A box contains a total of 100 DVDs that were manufactured on two machines, A and B. Of the total DVDs, 15 are defective. Of the 60 DVDs that were manufactured on Machine A, 9 are defective.

Let D be the event that a randomly selected DVD is defective, and let A be the event that a randomly selected DVD was manufactured on Machine A. Are D and A independent?

$$P(\mathrm{D}) = \frac{15}{100} = 0.15 \qquad \text{and} \qquad P(D|A) = \frac{9}{60} = 0.15$$

Since $P(D|A) = 0.15 = P(D) = 0.15$ ➔ D and A are independent

Contingency table and joint probabilities

Shows the ***joint*** distribution of two variables:

Ex: 100 employees classified by gender (male or female) and opinion (in favour or against giving bonus to CEO)

	In Favor	**Against**
Male	15	45
Female	4	36

Consider 2 events: A="be male", B="be in favour of bonus".

The ***joint probability*** of A and B is $P(A \cap B) = \frac{15}{100} = 0.15$

Contingency table and marginal probabilities

Marginal probability: prob. of an event regardless of any other event

Ex:

	In Favor	Against	Total
Male	15	45	60
Female	4	36	40
Total	19	81	100

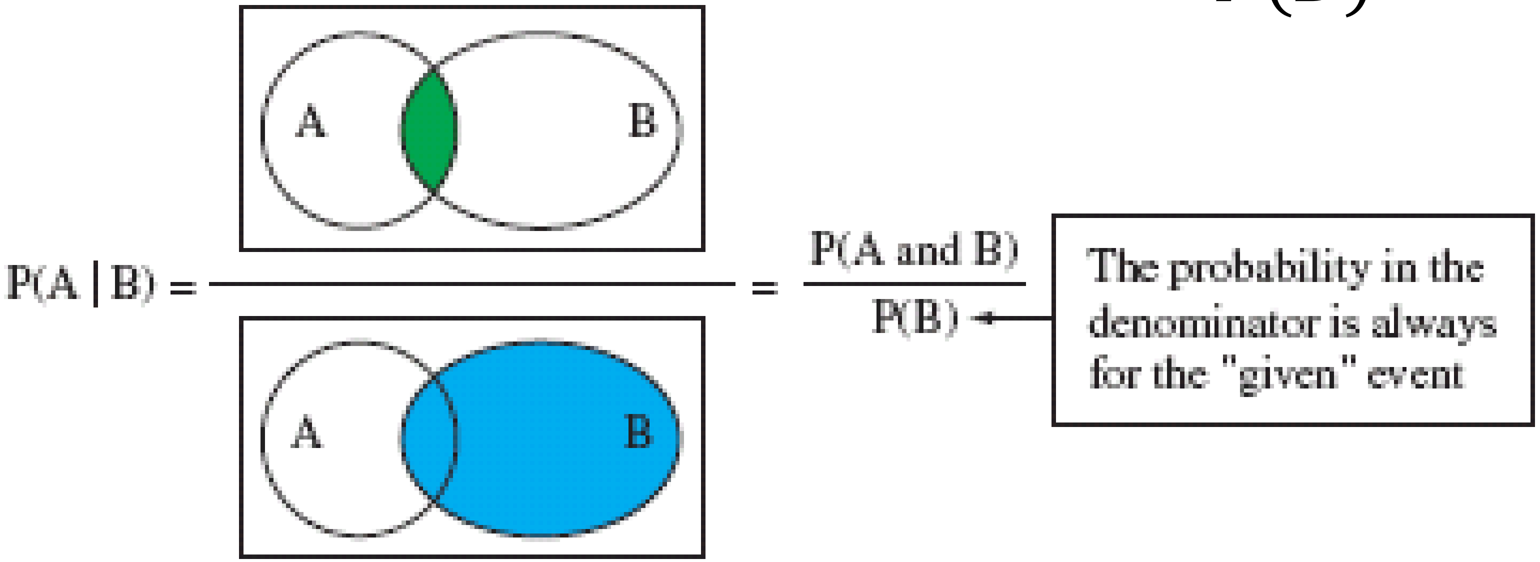
Let A="be male", B="be in favour of bonus",

then $P(A) = 60/100 = 0.6$ and $P(B) = 19/100 = 0.19$

Contingency table and conditional probabilities

The probability of A given B (i.e. likelihood of occurrence of A ***given*** that B has already occurred) is:

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



Contingency table and conditional probabilities

Recalling the example above,

	In Favor	Against	Total
Male	15	45	60
Female	4	36	40
Total	19	81	100

$$P(A|B) = \frac{P(A \cap B)}{P(\mathrm{B})} = \frac{15/100}{19/100} = 0.79$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{15/100}{60/100} = 0.25$$

Example

The following contingency table gives the distribution on 2000 randomly selected adults, by gender and having or not they have ever shopped on the Internet.

	Have Shopped	Have Never Shopped
Male	500	700
Female	300	500

Example

	Have Shopped	Have Never Shopped
Male	500	700
Female	300	500

Find the probability that 1 adult randomly selected

a) Has never shopped on the Internet

Example

	Have Shopped	Have Never Shopped
Male	500	700
Female	300	500
	800	1200

Find the probability that 1 adult randomly selected

a) Has never shopped on the Internet

$$P(\bar{S}) = \frac{1200}{2000} = 0.60$$

Example

	Have Shopped	Have Never Shopped
Male	500	700
Female	300	500
	800	1200

Find the probability that 1 adult randomly selected

b) Is a Male

Example

	Have Shopped	Have Never Shopped	
Male	500	700	1200
Female	300	500	800
	800	1200	

Find the probability that 1 adult randomly selected

b) Is a Male

$$P(M) = \frac{1200}{2000} = 0.60$$

Example

	Have Shopped	Have Never Shopped	
Male	500	700	1200
Female	300	500	800
	800	1200	

Find the probability that 1 adult randomly selected

c) has shopped on the Internet given that she is a female

Example

	Have Shopped	Have Never Shopped	
Male	500	700	1200
Female	300	500	800
	800	1200	

Find the probability that 1 adult randomly selected

c) has shopped on the Internet given that she is a female

$$P(S|F) = \frac{300}{800} = 0.375$$

Example

	Have Shopped	Have Never Shopped	
Male	500	700	1200
Female	300	500	800
	800	1200	

Find the probability that 1 adult randomly selected

d) Is a male given that he has never shopped on the Internet

Example

	Have Shopped	Have Never Shopped	
Male	500	700	1200
Female	300	500	800
	800	1200	

Find the probability that 1 adult randomly selected

d) Is a male given that he has never shopped on the Internet

$$P(M|\bar{S}) = \frac{700}{1200} = 0.5833$$

Example

	Have Shopped	Have Never Shopped	
Male	500	700	1200
Female	300	500	800
	800	1200	

e) Are the events “male” and “female” mutually exclusive?

f) What about the events “have shopped” and “male”?

g) Are the events “female” and “have shopped” independent?

Example

	Have Shopped	Have Never Shopped	
Male	500	700	1200
Female	300	500	800
	800	1200	

e) Are the events "male" and "female" mutually exclusive? Yes

f) What about the events "have shopped" and "male"? NO!

g) Are the events "female" and "have shopped" independent? NO, indeed

$$P(S) = 1 - P(\bar{S}) = 1 - 0.60 = 0.4 \neq P(S|F) = 0.375$$

Probabilities: example

250 students were classified according to their major and satisfaction

	Happy	Unhappy
Psychology	80	20
Communication	115	35

Find probability that 1 randomly selected student is:

A) psychology major and happy with that major

B) communication major or unhappy with the major

C) Communication major given that he/she is happy with the major

Probabilities: example

250 students were classified according to their major and satisfaction

	Happy	Unhappy
Psychology	80	20
Communication	115	35

Answers

A) $\frac{80}{250} = 0.32$

B) $\frac{115+35+20}{250} = 0.68$

C) $\frac{115}{115+80} = 0.59$

Probabilities: example

250 students were classified according to their major and satisfaction

	Happy	Unhappy
Psychology	80	20
Communication	115	35

Find probability that 1 randomly selected student is:

D) Unhappy with the major given that he/she is a psychology major

E) Happy with the major

F) Psychology major

Probabilities: example

250 students were classified according to their major and satisfaction

	Happy	Unhappy
Psychology	80	20
Communication	115	35

Answers:

D) $\frac{20}{100} = 0.20$

E) $\frac{115+80}{250} = 0.78$

F) $\frac{100}{250} = 0.40$

Probabilities: example

250 students were classified according to their major and satisfaction

	Happy	Unhappy
Psychology	80	20
Communication	115	35

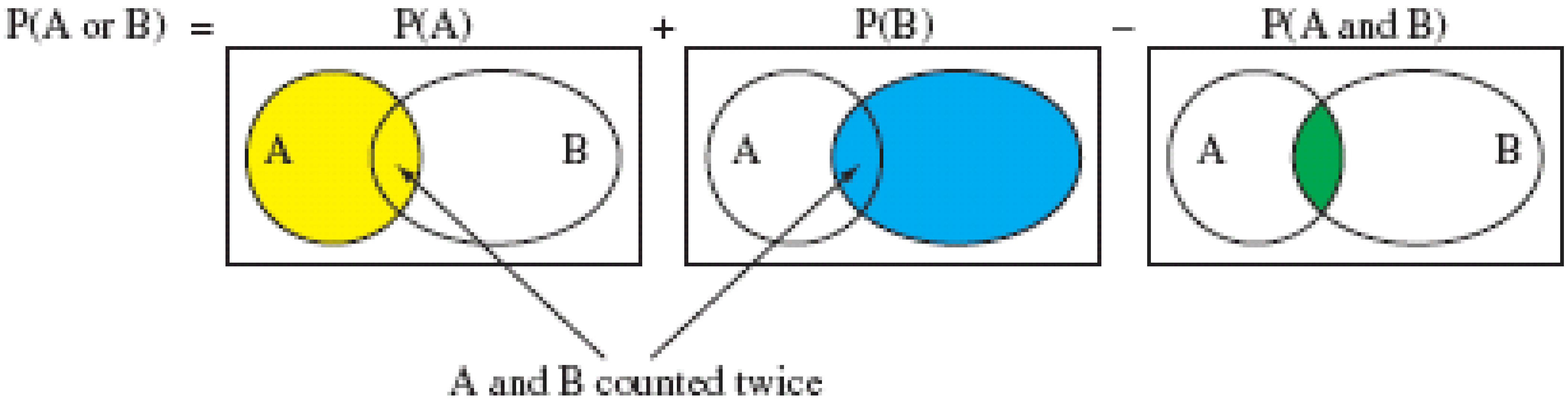
Are “psychology major” and “happy with major” independent events?

Are they mutually exclusive?

Addition rule

Consider 2 events, A and B, then the probability of their union (i.e. at least one of the two events occurs) is:

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



Addition rule: example

Amy is trying to purchase concert tickets online for two of her favorite bands, the Kings of Leon (K) and the Lumineers (L). She estimates that her probability of being able to get tickets for the Kings of Leon is 0.14, the probability of being able to get tickets for the Lumineers is 0.23, and the probability of being able to get tickets for both concerts is 0.026. What is the probability that she will be able to get tickets for at least one of the two concerts?

$$P(K \cup L) = P(K) + P(L) - P(K \cap L) = 0.14 + 0.23 - 0.026 = 0.344$$

Law of total probability

In probability theory, the law (or formula) of total probability is a fundamental rule relating marginal probabilities to conditional probabilities.

It expresses the total probability of an outcome which can be realized via several distinct events, hence the name.

$$P(A) = P(A|B) \cdot P(B) + P(A|\bar{B}) \cdot P(\bar{B})$$

Bayes Theorem

Allows to update the probability of an event whenever new info about the occurrence of a related event become available.

Consider and event A, with a certain probabilty P(A), and suppose you know that B has now occurred ➔ the probabilty of A can adjusted to take into account this additional information

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

Bayes Theorem

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \qquad P(B|A) = \frac{P(B \text{ and } A)}{P(A)}$$

$$P(A \text{ and } B) = P(B \text{ and } A)$$

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

Bayes Theorem

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

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

Bayes Theorem: example

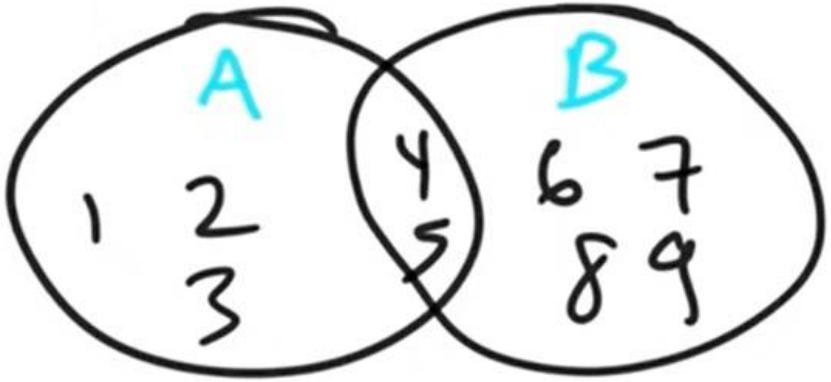
$A = \{1, 2, 3, 4, 5\}$ $B = \{4, 5, 6, 7, 8, 9\}$

$P(A|B) = ?$

Bayes Theorem: example

$A = \{1, 2, 3, \underline{4, 5}\}$ $\quad$ $B = \{\underline{4, 5}, 6, 7, 8, 9\}$

$P(A|B) = ?$



$P(A) = \frac{5}{9}$

$P(B) = \frac{6}{9}$

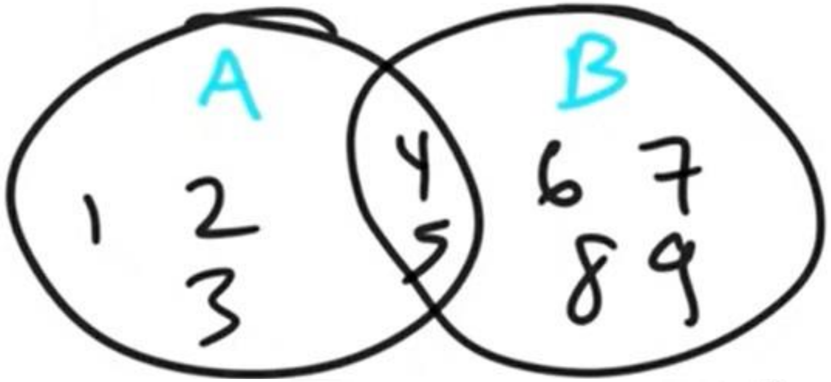
$P(B|A) = \frac{2}{5}$

$S = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$

Bayes Theorem: example

$A = \{1, 2, 3, \underline{4, 5}\}$ $\quad$ $B = \{\underline{4, 5}, 6, 7, 8, 9\}$

$P(A|B) = ?$



$P(A) = \frac{5}{9}$

$P(B) = \frac{6}{9}$

$P(B|A) = \frac{2}{5}$

$$P(A/B) = \frac{P(B/A) \cdot P(A)}{P(B)}$$

$$= \frac{\frac{2}{5} \cdot \frac{5}{9}}{6/9} = \frac{2}{6}$$

Bayes Theorem: example

$$P(A) = \frac{5}{9}$$

$$P(B) = \frac{6}{9}$$

$$P(B|A) = \frac{2}{5}$$

$$P(A/B) = \frac{P(B/A) \cdot P(A)}{P(B)}$$

$$= \frac{\frac{2}{5} \cdot \frac{5}{9}}{6/9} = \frac{2}{6}$$

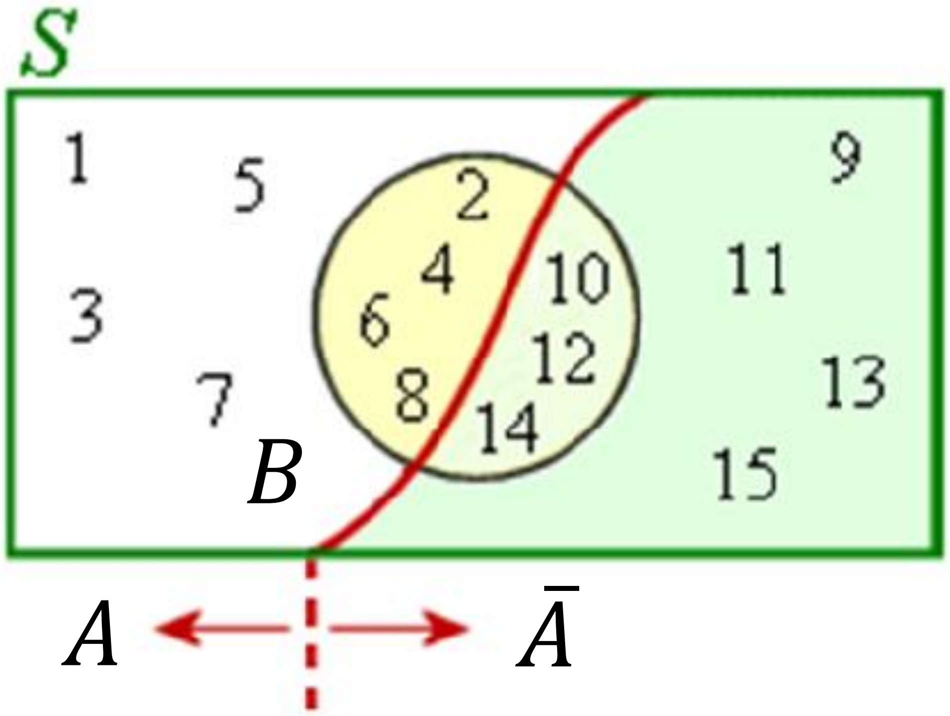
$$P(A/B) = \frac{P(A \text{ and } B)}{P(B)}$$

$$= \frac{\frac{2}{9}}{6/9} = \frac{2}{6}$$

Bayes Theorem: example

Consider the same example above and suppose you know that B has now occurred ➔

What is the probability that A occurs given that now you know that B has already occurred?



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

Bayes Theorem: how to solve

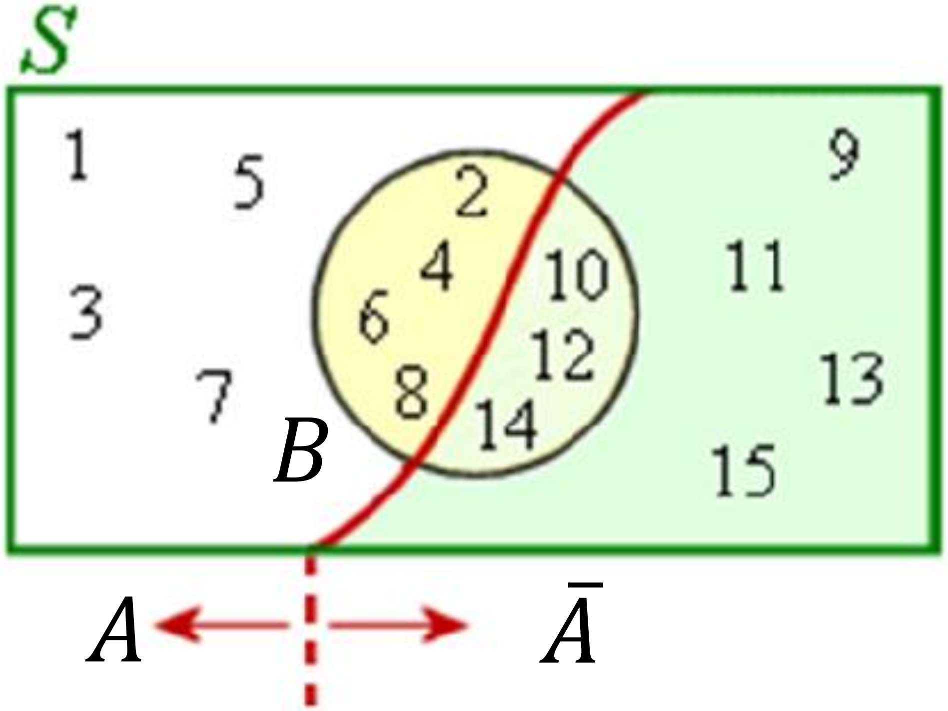
To solve it:

1) Locate the 2 events : A and B
2) Assign correct probabilities

$P(A) = \frac{8}{15}, P(B) = \frac{7}{15}, P(\bar{A}) = 7/15$

$P(B|A) = 4/8 \rightarrow P(\bar{B}|A) = 4/8$

$P(B|\bar{A}) = 3/7 \rightarrow P(\bar{B}|\bar{A}) = 4/7$



3) Find the required probability

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{\frac{4}{8} \cdot \frac{8}{15}}{\frac{7}{15}} = 4/7$$

Bayes Theorem: example

20% of the inbox messages is unwanted. 70% of the unwanted email contain deceptive words (Lottery, Win, Confirm your bank account, etc), while legitimate email messages do contain those words with a proability of 5%. Compute the probability that a message containing deceptive words is unsolicited and needs to be moved to Spam folder

1) Locate the two events:

D=«email contains deceptice words»

S=«email is moved to Spam folder»

Bayes Theorem: example

20% of the inbox messages is unwanted. 70% of the unwanted email contain deceptive words (Lottery, Win, Confirm your bank account, etc), while legitimate email messages do contain those words with a proability of 5%. Compute the probability that a message containing deceptive words is unsolicited and needs to be moved to Spam folder

2) Assign correct probabilities

$P(S) = 0.2 \rightarrow P(\bar{S}) = 0.8$

$P(D|S) = 0.7 \rightarrow P(\bar{D}|S) = 0.3$

$P(D|\bar{S}) = 0.05 \rightarrow P(\bar{D}|\bar{S}) = 0.95$

Bayes Theorem: example

20% of the inbox messages is unwanted. 70% of the unwanted email contain deceptive words (Lottery, Win, Confirm your bank account, etc), while legitimate email messages do contain those words with a proability of 5%. Compute the probability that a message containing deceptive words is unsolicited and needs to be moved to Spam folder

3) Find the solution:

$$P(S|D) = \frac{P(D|S)P(S)}{P(D|S)P(S) + P(D|\bar{S})P(\bar{S})} = 0.78$$

Bayes Theorem: example

40% of the smokers of a little town prefer brand A, while the rest prefer other brands. Of those preferring brand A, 30% are females, while the 40% of those preferring other brands are females.

What is the probability that a randomly selected smoker prefers brand A, given that she is a female?

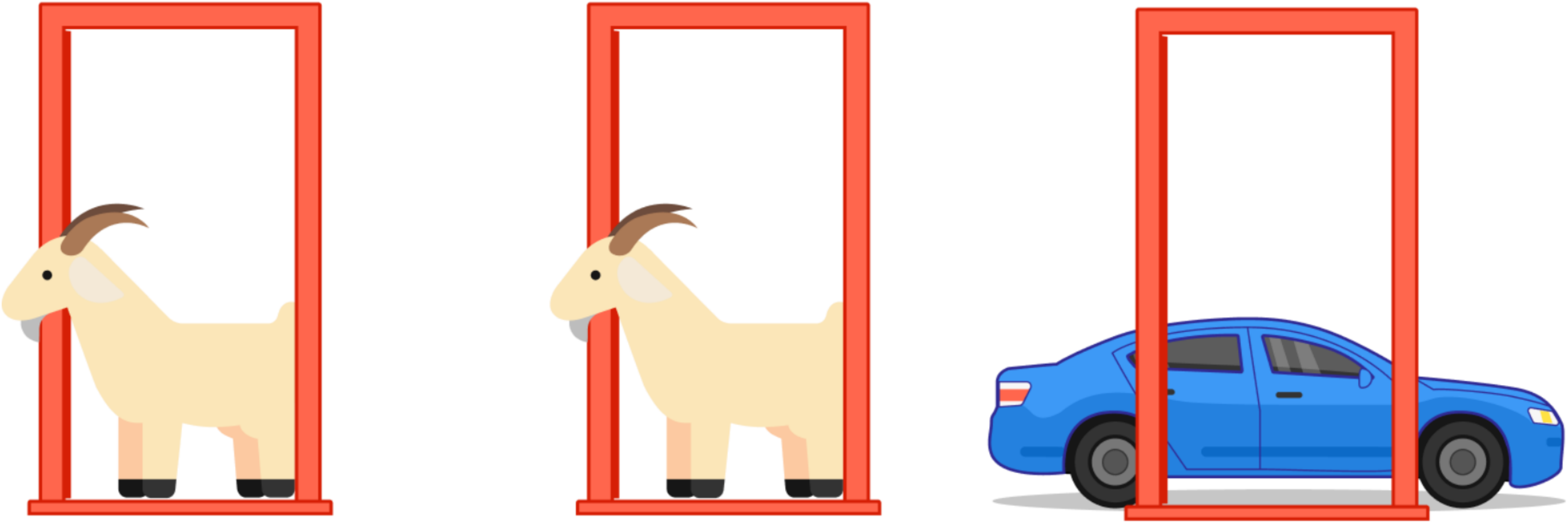
1) Locate the two events: F =«female», A=«prefer brand A»
2) Assign correct probabilities
3) Find the solution

$$P(A|F) = \frac{P(F|A)P(A)}{P(F)} = 1/3$$

Monty Hall

Monty Hall

Monty Hall



https://www.youtube.com/watch?v=mhlc7peGlGg

Or

https://www.youtube.com/watch?v=C4vRTzsv4os