Introduction to Probability

Probability Theory

<u>Some general remarks:</u>

- The mathematical theory of probability is ubiquitous.
- The systematic study of the mathematical theory of probability starts with a correspondence between Blaise Pascal and Pierre de Fermat.
- It is referred to as the ‘probability calculus’.

Probability Axioms (1)

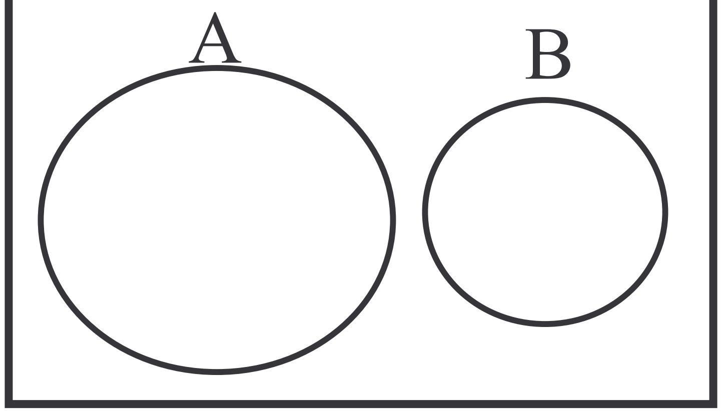
- Three basic axioms:

  (1) $0 \leq P(A) \leq 1$

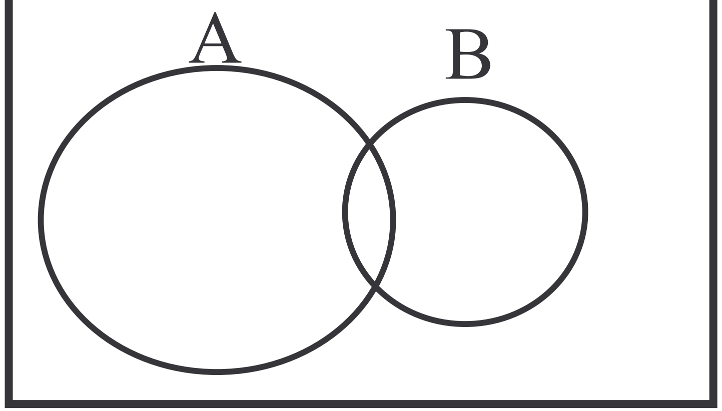
  (2) If *A* is a logical truth (i.e. tautology), then $P(A) = 1$.

  (3) Where *A* and *B* are mutually exclusive, $P(A \vee B) = P(A) + P(B)$. Alternatively: $P(A \cup B) = P(A) + P(B)$.

A and B disjoint



A and B overlap



Probability Axioms (2)

- Three basic axioms:

  (1) $0 \leq P(A) \leq 1$

  (2) If *A* is a logical truth (i.e. tautology), then $P(A) = 1$.

  (3) Where *A* and *B* are mutually exclusive, $P(A \vee B) = P(A) + P(B)$. Alternatively: $P(A \cup B) = P(A) + P(B)$.
- Other axiomatisations possible.
- A, B, etc. are propositions but with qualifications we can plug in events, theories, evidence, etc.
- Notice that $P(A) = 1$ does not necessarily mean that A is a logical truth – conditional NOT bi-conditional.

Example

- Suppose we are throwing a fair die and that only one side can come up at any given throw:

  P (the die will come up n) =1/6 where n takes any value between 1 and 6.

- Suppose we ask: P(even number or 5) = ?
- Answer: Given that the events are mutually exclusive, that means that P(even number or 5) = P(2v4v6v5) = P(2) + P(4) + P(6) + P(5)

  = 1/6 + 1/6 + 1/6 + 1/6 = 4/6.

Useful Theorems

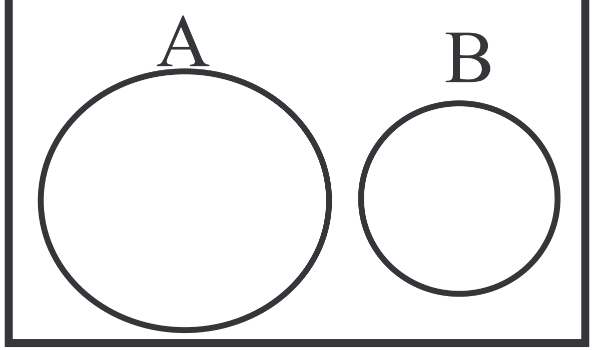
- From the basic axioms one can deduce several useful theorems. Here's a short list:

  (a) P(~A) = 1 - P(A).

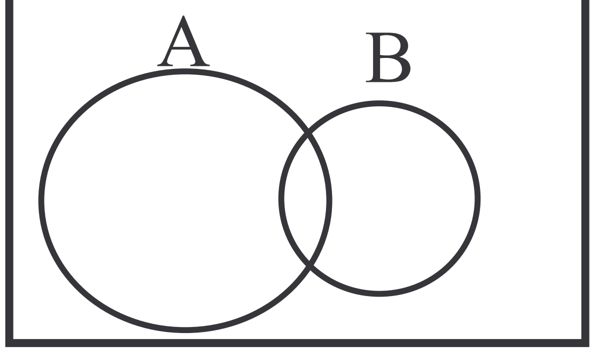
  (b) If *A* and *B* are logically equivalent, then P(A)=P(B).

  (c) P (A v B)= P(A) + P(B) – P (A&B).

A and B disjoint



A and B overlap



A Simple Proof

- Proving theorem (a): $P(\sim A) = 1 - P(A)$.

(i) $P(A \vee \sim A) = P(A) + P(\sim A)$	Axiom 3
(ii) $P(A \vee \sim A) = 1$	Axiom 2
(iii) $P(A) + P(\sim A) = 1$	Steps i, ii
$\therefore P(\sim A) = 1 - P(A)$	

Conditional Probabilities

- Definition:

  $P(A/B) = P(A\&B)/P(B)$ where $P(B) > 0$.

- Some take conditional probability as primitive and axiomatise directly on this notion.
- Conditional probabilities are important to us since we are interested in the relationships between theory and evidence but also between old and new theories.

  <u>So, we can ask</u>:

  - $P(T_1/E_1) = ?$
  - $P(T_1/T_2) = ?$

Central Questions

- A central foundational concern is expressed by the following question: What does it mean to say "*A* has a probability of value *p*"?

- Another related concern is expressed by this question: How are probability values to be measured?

Interpretations of Probability

- What does the term ‘probability’ mean?
- Five interpretations:

  (1) Classical

  (2) Subjective/Personal

  (3) Logical

  (4) Frequency

  (5) Propensity
- Any interpretation given must satisfy the basic axioms given earlier.

Classical

- Prominent advocate: Pierre Simon de Laplace.
- Assigns probabilities in the absence of any evidence OR in the presence of symmetrical evidence.
- <u>Principle of indifference</u>: Whenever there is no evidence favouring one possibility over another, we assign the same probability to them.
- $P(A) = m/n$ where *m* is the number of favourable cases and *n* the number of all *equally possible* cases.
- Example: Fair coin
  On the basis of symmetry we say:
  $P(\text{turning up heads}) = 1/2$

Classical (2)

- Some major problems:

  (1) Applicable only to those cases that are equipossible.

  <u>Reply:</u>

  Bite the bullet, i.e. not applicable in other genuine cases.

  (2) If we have symmetrical evidence the classical interpretation is fine, BUT if we are in a state of ignorance, we should not assign any probabilities whatsoever.

  <u>Reply:</u>

  Sometimes we NEED to make a choice. It’s best to choose in accordance with the principle of indifference.

Classical (3)

(3) The principle of indifference can be used in incompatible ways. This leads to paradoxes:

Suppose: Factory produces cubes with side-length between 0 and 1 foot.

Question: What is the probability that a random cube has a side-length between 0 and ½ foot?

Answer: 0.5 (assumption: uniform distribution over side-length).

Question can be restated 'equivalently' as: What is the probability that a random cube has a face-area between 0 and ¼ square-feet?

NB: ¼ square-feet of face-area corresponds to ½ foot of side-length.

Answer: 0.25 (assumption: uniform distribution over face-area).

Problem: We cannot allow the same event to have different probabilities. NB: The principle of indifference says nothing about which assumption is the right one.

Food for Thought

- Is probability an objective or a subjective matter?

Reading

- Curd, M. and Cover, J.A. (1998) ‘Commentary’, in Curd and Cover, pp. 628-632.
- Earman, J. and Salmon, W.C. (1999) ‘The Confirmation of Scientific Hypotheses’, in M. Salmon et al. (eds.) Introduction to the Philosophy of Science, Indianapolis: Hackett Publishing Company, ch. 2, pp. 66-77.