Elements of Probability

Fall 2024
Sunita Sarawagi

Thanks to Prof Ajit Rajawade for the initial version of the slides

Data interpretation from samples is uncertain

- Need a formal representation of uncertainty
- Probability provides a formal framework of expressing uncertainty when drawing conclusions from finite samples of a much larger population.

Probability in Computer Science

- Algorithm design
  - Randomized algorithms: steers around unlikely situations
  - Several hard problems that can only be solved efficiently with high probability
- Performance analysis
  - What is the probability that you will find the next accessed page in cache?
  - What is the probability that the length of the queue will be greater than 5 when a job arrives at a server?
- Network protocol design
- Machine Learning/AI: Is all about probability and statistics

Topic Overview

- Terminology: sample space, event, probability
- Composition of events; mutual exclusion and independence
- Axioms of probability
- Principles of counting
- Conditional probability and Bayes' theorem
- Some paradoxes!

Sample space

- Consider an experiment whose outcome is not known in advance.

- Example 1: A coin toss

- But we do know the complete set of possible outcomes: Heads or tails

- The set of all possible outcomes of an experiment is called the **sample space**.

Sample space

- Example 2: Measurement of your body temperature (assume it’s an integer) with a thermometer. What’s the sample space?
  - Say between 30 to 40 degrees Celsius, so the sample space = {30,31,…,39,40}

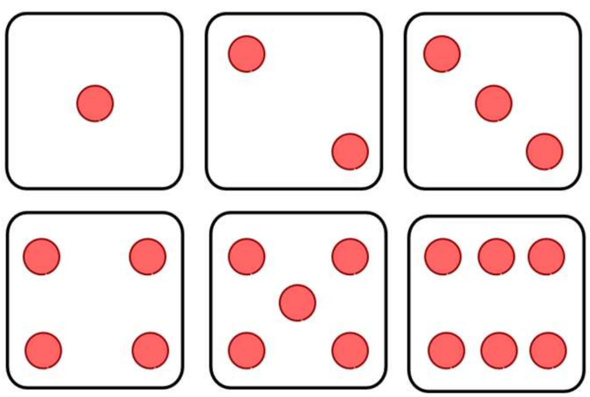
- Example 3: An experiment to randomly choose a student from the CSE 2024 batch at IITB and declare him/her the branch topper
  - Sample space = set of all students in that batch

Sample space

- Example 4: Consider a four-country ODI series between India, Pakistan, Bangladesh and Australia. What is the set of rankings?
  - Sample space = set of all 4! permutations of the string IPBA

- Example 5: An experiment to roll a die

  -Sample space = {1,2,3,4,5,6}



Event

- Any subset of the sample space is called an **event**.

- If the outcome of an experiment is contained in Event E, then we say E has *occurred*.

- In example 1, if $E = \{H\}$, then E is the event that the coin produced a heads.

- In example 2, if $E = \{\text{set of temperatures from 33 to 37}\}$, then E is the event that the temperature was “normal” (i.e. not exceeding 37 and not less than 33)

Composition of Events

- Given event E, event E^c is the event that E did not occur. E^c is called the **complement** of E.

- Given events E and F, the event G that either E or F (or both) occur is called as the **union** of E and F, and denoted as $G = E \cup F$.

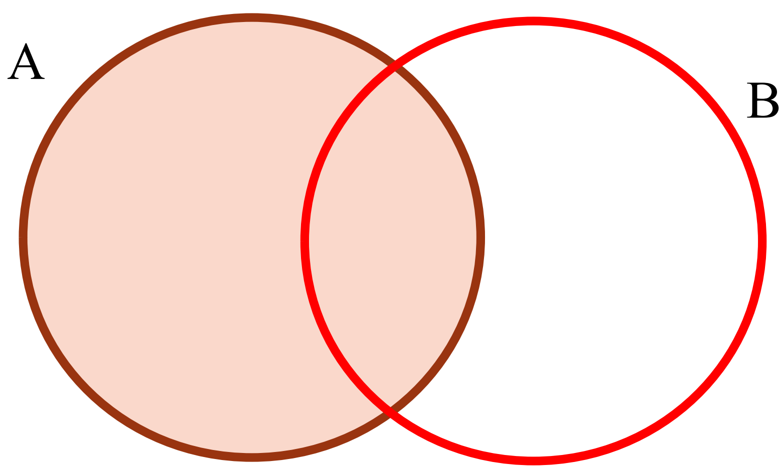
- Given events E and F, the event G that both E and F occur is called as the **intersection** of E and F, and denoted as $G = E \cap F$ or $G = EF$

Composition of Events

- Union and intersection can be extended to handle any arbitrary number of events.

- If two events cannot occur together (for example?), then their intersection is a null set. Such events are called mutually exclusive.

Occurrence of one excludes the possibility of occurrence of the other.



This is called as a Venn diagram in set theory.

Composition of Events

- An event and its complement – are always mutually exclusive events.
- Example:
  - Let F be the event that a patient tests negative for a certain disease in a medical test.
  - Let G be the event that (s)he tests positive for the same disease in the same test.
  - Then F and G are mutually exclusive.

- Example:
  - Let E be the event that the sum of three consecutive dice throws was greater than or equal to 3.
  - Let F be the event that the sum of three consecutive dice throws was greater than or equal to 4.
  - Then E and F are NOT mutually exclusive. In fact F is a subset of E.

Probability of an event

- We conduct an experiment, whose outcomes are uncertain but come from a sample space S
- We are interested in a subset S of the sample space, which we call an event E.
- If we repeat the experiment under identical conditions very large number of times,
  - Probability of E, P(E) is the fraction of times that outcome is in event E

- Example: rolling of dice.

Axioms of probability

- For an event E from sample space S, we have:

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

Axiom 2 : $P(S) = 1$

Axiom 3 : For mutually exclusive events $E_1, E_2, \ldots, E_n$, we have

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i), \qquad n = 1, 2, \ldots, \infty$$

The notion of relative frequency of event E obeys the above axioms

Properties derivable from axioms

- Properties (can be proved by Venn diagrams):

$$P(A \cup B) = P(A) + P(B) - P(AB) \leq P(A) + P(B)$$

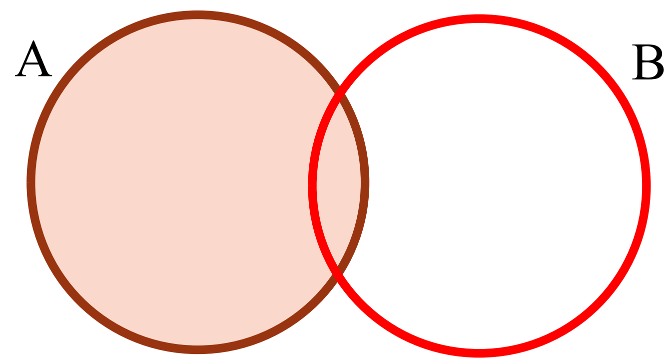
This implies

$$(1) P(A^c) = 1 - P(A)$$

$$(2) A \subseteq B \rightarrow P(A) \leq P(B)$$

Is the converse of (2) also true?

No.



Equally likely outcomes

- We will assume that each of the singleton outcomes in the sample space is equally likely.
- So, if the experiment is to roll a die, then all six faces will show up with equal probability.
- We will assume finite sample spaces for now.
- In such a case, the probability of an event E is given as:

$$P(E) = \frac{\text{Number of points in E}}{\text{Number of points in sample space}}$$

Principles of counting: motivating example

- Useful when solving problems on discrete probability.

- For example: Suppose a box contains 6 white and 5 black balls. If you draw two balls at random, what is the probability that one is white and the other is black?

Principles of counting: Product rule

Suppose a procedure can be broken down into a sequence of k tasks, and there are

- n_1 ways to do task 1,
- n_2 ways to do task 2, …
- n_k ways to do task k.

Then there are $n_1 n_2 .. n_k$ ways to do the entire procedure.

```
c = 0
for (i_1 = 1 to n_1)
{
  for (i_2 = 1 to n_2)
  {
      .
      .
      for (i_k = 1 to n_k)
      {
          c = c + 1
      }
      .
      .
  }
}
```

Principles of counting: example

- For example: Suppose a box contains 6 white and 5 black balls. If you draw two balls at random, what is the probability that one is white and the other is black?

- There are two scenarios: (1) the first ball is white and second is black, or (2) vice versa.
- For (1), the probability that the white ball is picked is 6/11, and the probability that the black ball is picked is 5/10 (10 balls remain after the first white ball is picked). The overall probability is 30/110 (product rule).
- For (2), the probability that the black ball is picked is 5/11, followed by a 6/10 probability of picking a white ball, leading to an overall probability of 30/110 (product rule).
- The total probability is (30+30)/110 = 6/11 (sum rule).

Conditional Probability

- An important concept.
- Helps one quantify uncertainty of outcomes under partial knowledge or constraints.
- For example
  - What is the probability that the outcome of a dice roll is 2 given that it is even?

Let A, B be two events. Conditional probability of A, given that B has already occurred

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

Bayes Formula

Example 3.7.d. In answering a question on a multiple-choice test, a student either knows the answer or she guesses. Let p be the probability that she knows the answer and $1 - p$ the probability that she guesses. Assume that a student who guesses at the answer will be correct with probability $1/m$, where m is the number of multiple-choice alternatives. What is the conditional probability that a student knew the answer to a question given that she answered it correctly?

- S = {KC, GC, ~~KI~~, GI}
- B = Correct answer = {KC, GC}
- A = {KC, ~~KI~~}
- P(A,B) = P(KC) = p
- P(B) = p + (1-p)/~~m~~ $\times \frac{1}{m}$

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{p}{p + \frac{(1-p)}{m}}$$

$m = 5$

$p = 0.6$

$$P(K|C) = \frac{0.6}{0.6 + \frac{0.4}{5}} = \frac{1}{1 + \frac{2}{3} * 5} = 0.88$$

Example 3.7.e. A laboratory blood test is 99 percent effective in detecting a certain disease when it is, in fact, present. However, the test also yields a "false positive" result for 0.1 percent of the healthy persons tested. (That is, if a healthy person is tested, then, with probability 0.01, the test result will imply he or she has the disease.) If .5 percent of the population actually has the disease, what is the probability a person has the disease given that his test result is positive?

$$P(T_p \mid D) = 0.99$$

$$P(T_p \mid D^c) = 0.001$$

$$P(D) = 0.005$$

$$P(D \mid T_p) = \frac{P(T_p \mid D)P(D)}{\underbrace{P(T_p \mid D)P(D) + P(T_p \mid D^c)P(D^c)}_{P(T_p)}}$$

$$= \frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.001 \times (1 - 0.005)}$$