Machine Learning

Sample Spaces and Probability: Basic Concepts

Basic Concepts

- A **probability experiment** is a chance process that leads to well-defined results called outcomes.
- An **outcome** is the result of a single trial of a probability experiment.
- A **sample space** is the set of all possible outcomes of a probability experiment.

➢ As example –

Experiment	Sample space
Toss one coin	Head, tail
Roll a die	1, 2, 3, 4, 5, 6
Answer a true/false question	True, false
Toss two coins	Head-head, tail-tail, head-tail, tail-head

Basic Concepts

Example:
Find out a sample space for rolling two dice.

Die 1	Die 2					
	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Basic Concepts

- An **event** consists of a set of outcomes of a probability experiment.
- An event can be one outcome or more than one outcome.

- For example, if a die is rolled and a 6 shows, this result is called an *outcome,* since it is a result of a single trial. An event with one outcome is called a **simple event.**

- The event of getting an odd number when a die is rolled is called a **compound event,** since it consists of three outcomes or three simple events.

- In general, a compound event consists of two or more outcomes or simple events.

Basic Concepts

- There are three basic interpretations of probability:
  1. Classical probability
  2. Empirical or relative frequency probability
  3. Subjective probability

Machine Learning

Classical Probability

Classical Probability

- **Classical probability** assumes that all outcomes in the sample space are equally likely to occur.
  - For example, when a single die is rolled, each outcome has the same probability of occurring. Since there are six outcomes, each outcome has a probability of 1/6.
- **Formula for Classical Probability:**
  - The probability of any event *E* is

$$\frac{\textbf{Number of outcomes in E}}{\textbf{Total number of outcomes in the sample space}}$$

This probability is denoted by

$$\boldsymbol{P(E) = \frac{n(E)}{n(S)}}$$

This probability is called *classical probability,* and it uses the sample space *S*

Machine Learning

Empirical Probability

Empirical Probability

- The difference between classical and **empirical probability** is that classical probability assumes that certain outcomes are equally likely (such as the outcomes when a die is rolled), while empirical probability relies on actual experience to determine the likelihood of outcomes.

- In empirical probability, one might actually roll a given die 6000 times, observe the various frequencies, and use these frequencies to determine the probability of an outcome.

Method	Frequency
Drive	**41**
Fly	**6**
Train or bus	**3**
	50

For example, the probability of selecting a person who is driving is 41/50.

Empirical Probability

- **Formula for Empirical Probability**
  - Given a frequency distribution, the probability of an event being in a given class is

$$P(\mathrm{E}) = \frac{\text{frequency for the class}}{\text{total frequencies in the distribution}} = \frac{f}{n}$$

  - This probability is called *empirical probability* and is based on observation.

Machine Learning

Subjective Probability

Subjective Probability

- **Subjective probability** uses a probability value based on an educated guess or estimate, employing opinions and inexact information.

- A physician might say that, on the basis of her diagnosis, there is a 30% chance the patient will need an operation.

- A seismologist might say there is an 80% probability that an earthquake will occur in a certain area.

Machine Learning

The Addition Rules for Probability

The Addition Rules for Probability

- Two events are **mutually exclusive events** if they cannot occur at the same time (i.e., they have no outcomes in common).

➢ **Example:**
  - When a single die is rolled. Getting an odd number and getting an even number.
  - When two events A and B are mutually exclusive, the probability that A or B will occur is

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

The Addition Rules for Probability

- **Example:**
  - A box contains 3 red balls, 4 blue balls, and 5 green balls. If a person selects a ball at random, find the probability that it is either a red ball or a green ball.
- **Solution:**
  - Since the box contains 3 red balls, 4 blue balls, 5 green balls so a total of 12 balls,

    P(red or green balls) = *P*(red balls) + *P*(green balls) = (3/12) + (5/12) = (8/12) = (2/3) . Here the events are mutually exclusive.

The Addition Rules for Probability

- If *A* and *B* are *not* mutually exclusive, then

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

Example:

Staff	Females	Males	Total
Nurses	7	1	8
Physicians	3	2	5
Total	10	3	13

- The probability is

P(nurse or male) = P(nurse) + P(male) – P(male nurse)

$$= \frac{8}{13} + \frac{3}{13} - \frac{1}{13} = \frac{10}{13}$$

Machine Learning

The Multiplication Rules for Probability

The Multiplication Rules for Probability

- Two events *A* and *B* are **independent events** if the fact that *A* occurs does not affect the probability of *B* occurring.
- When two events are independent, the probability of both occurring is

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

➢ **Example of Tossing a Coin:**
  - A coin is flipped and a die is rolled. Find the probability of getting a head on the coin and a 4 on the die.

➢ **Solution:**

$P(\text{head and } 4) = P(\text{head}) * P(4) = (1/2) * (1/6) = (1/12)$

Note that the sample space for the coin is H, T; and for the die it is 1, 2, 3, 4, 5, 6.

Machine Learning

Conditional Probability

Conditional Probability

- When the outcome or occurrence of the first event affects the outcome or occurrence of the second event in such a way that the probability is changed, the events are said to be **dependent events.**

- The **conditional probability** of an event *B* in relationship to an event *A* is the probability that event *B* occurs after event *A* has already occurred. The notation for conditional probability is $P(B \mid A)$.

- When two events are dependent, the probability of both occurring is

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

Conditional Probability

- **Example:**
  - There were 5 burglaries reported in 2003, 16 in 2004, and 32 in 2005. If a researcher wishes to select at random two burglaries to further investigate, find the probability that both will have occurred in 2004.
- **Solution**
  - In this case, the events are dependent since the researcher wishes to investigate two distinct cases. Hence the first case is selected and not replaced.
    $P(C1 \text{ and } C2) = P(C1) * P(C2 \mid C1)$
    $= (16/53) * (15/52)$
    $= (60/689)$

Conditional Probability

➢ **Formula for Conditional Probability:**

- The probability that the second event *B* occurs given that the first event *A* has occurred can be found by dividing the probability that both events occurred by the probability that the first event has occurred. The formula is

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

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

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

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

Conditional Probability

- **Example:**
  - Responses obtained on a survey question as below -

Gender	Yes	No	Total
Male	32	18	50
Female	8	42	50
Total	40	60	100

- Find the probability that the respondent was a male, given that the respondent answered no.
  - The problem is to find P(M|N).

$$P(M|N) = \frac{\mathrm{P(N\ and\ M)}}{\mathrm{P(N)}} = \frac{18/100}{60/100}$$

$$= \frac{18}{100} \div \frac{60}{100} = \frac{\overset{3}{\cancel{18}}}{\underset{1}{\cancel{100}}} \cdot \frac{\overset{1}{\cancel{100}}}{\underset{10}{\cancel{60}}} = \frac{3}{10}$$