Introduction to Probability
Chapter 1: Introduction

Dr. Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

1. Motivation
2. Introduction to the concepts of probability
3. Assigning probabilities to events
4. Continuity theorem in probability

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.
4. https://medium.com/@jrodthoughts/statistical-learning-in-artificial-intelligence-systems-e68927792175 by Jesus Rodriguez (2017)

Motivation

In artificial intelligence (AI) environment, uncertainty is a key element. Due to uncertainty the AI agent does not know the precise outcome of the given situation. Uncertainty is the typical result of random/probabilistic or partially observable environment. Statistical learning is helpful in these AI situations.

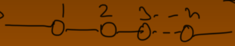
For example Bayes' theorem helps in dealing with uncertainty in the real world:

$$P(cause|effect) = P(effect)xP(effect|cause)/P(cause),$$

where $P(A|B)$ is the probability of occurrence of A given B. Replacing cause and effect with the probabilities of any state-action combination in an AI environment we arrive to the fundamentals of Bayesian learning. Many AI algorithms are based on Bayesian learning or statistical learning.

Motivation

In reliability computation of r-out-of n system the probability concepts are used where the components are assumed to have random life. r-out-of n system is a system which functions if atleast r out of its n components functions. Series and parallel system are n-out of n system and 1-out of n system, respectively.



Consider a situation where a redundant component or spare is provided to the system to increase its reliability. Then using probability concepts we can find increase in the reliability of the system.

Introduction

- Random Experiment (E): is an experiment whose outcome may not be perdicted in advance.
- Sample Space (Ω): Collection of all possible outcomes of random experiment.

Example

If E_1: Toss a coin, then $\Omega_1 = \{H, T\}$.
If E_2: Toss a coin till we get a head, then $\Omega_2 = \{H, TH, TTH, \ldots\}$.
If E_3: Lifetime of a bulb, then $\Omega_3 = [0, \infty)$.
If E_4: Radioactive particles emitted by a radioactive substance, then $\Omega_4 = \{0, 1, 2, \ldots\}$.
If E_5: Roll a pair of dice and see up face, then $\{(1,1),(1,2),\ldots,(6,6)\}$
$\Omega_5 = \{(i, j),\ i = 1, 2, 3, 4, 5, 6;\ j = 1, 2, 3, 4, 5, 6\}$.

Introduction

- Event: is subset of sample space. Event is denoted by capital letter.
- The set of all subsets is power set for a finite sample space .

Example

In E_1 the event is the toss yield a head $A_1 = \{H\}$
In E_2 we are getting head in third toss then event is $A_2 = \{TTH\}$.
In E_3 an event is $A_3 = (0, 2)$.
In E_4 if the radioactive particles emitted is 2, then $A_4 = \{2\}$.
In E_5 if sum of number on up faces is 4, then $A_5 = \{(1,3),(2,2),(3,1)\}$.
If $E_1, E_2, \ldots$ are events in Ω and events $E_1, E_2, \ldots$ are mutually exclusive, then $\bigcup_{i=1}^{\infty} E_i \subseteq \Omega$.

Sigma Field

Suppose E is an experiment with sample space as Ω. Let f be a collection of subsets of Ω. Then f is said to be a sigma field if

1. $\Omega \in f$.
2. If $A \in f$, then $\bar{A} \in f$.
3. If $A_1, A_2 \in f$, then $A_1 \cup A_2 \in f$.

Example-Sigma Field

Example (1)

E: Toss a coin, then $\Omega = \{H, T\}$. Then
$f_1 = \{\phi, \{H\}, \{T\}, \Omega\}$ is the power set and is a sigma field.
$f_2 = \{\phi, \Omega\}$ is trivial sigma field.

Example (2)

E: Toss a two coin, then $\Omega = \{HH, HT, TH, TT\}$. Then
$f_1 = \{\phi, \Omega\}$ is trivial sigma field.
f_2=Power set of Ω, is a sigma field. $= \{\phi, \{HH\}, --, \{TT\}, \{HH, HT\}, -- \{HH, HT, TH\}, -- -, \Omega\}$

$f_3 = \{\phi, \Omega, \{HH\}, \{HT, TH, TT\}\}$

Probability

Definition

E, Ω, f ✓ $A \in f$

Consider a random experiment E having sample space Ω. Let f be a sigma field of subsets of Ω. Consider an event A defined on f, then $P(A)$ is a real number called a probability of event A if $P(\cdot)$ satisfies following axioms:

1. $P(\Omega) = 1$, $\Omega \in f$.
2. $P(A) \geq 0$, $A \in f$.
3. If $A_1, A_2, \ldots$ are mutually exclusive events in f, then

$A_i \cap A_j = \phi, i \neq j ; A_i, A_j \in f$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

How to assign the probabilities

The assignment of probability is done on the basis of

1. prior experience or prior observations;
2. analysis of the experimental conditions;
3. assumptions.

Relative frequency approximation

$$f_A = \frac{\text{\# of favorable cases to the event } A}{\text{Total \# of times experiment is done}} = \frac{m_A}{m}$$

$$P(A) = \lim_{m\to\infty} \frac{m_A}{m} = \frac{\text{number of times event A occurs}}{\text{number of times experiment was done}}.$$

For large number of trails, the approximate probability obtained is quite near to the exact probability. The disadvantage of this approach is that the experiment should be repeated and is not a one off situation.

Classical Method

If a random expt. can result in n mutually exclusive and equally likely outcomes and if $n(A)$ of them have an attribute A

Here we assume that the possible outcomes of random experiment are equally likely and their total number is finite. Then

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{\text{number of favourable cases to event } A}{\text{number of cases in } \Omega}.$$

e.g. $\Omega = \{HH, HT, TH, TT\}$

A = getting atleast one head $= \{HH, HT, TH\}$

$P(A) = \frac{3}{4}$

Simple consequences of axioms

Consider the experiment E on (Ω, f).

1. $P(\phi) = 0, \ \phi = \{\} \in f$.
2. A and B are two events in f, then

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

3. If $A, B \in f$ and $A \subseteq B$, then $P(A) \leq P(B)$.
4. Let $A_i \in f, \ i = 1, 2, \ldots, n$, then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i,j=1 \\ i<j}}^{n} P(A_i \cap A_j) + \sum_{\substack{i,j,k=1 \\ i<j<k}}^{n} P(A_i \cap A_j \cap A_k) + \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n).$$

Example

Example

Suppose we are rolling two fair dices independently. We want to find the probability that

1. the sum of faces up is 7.
2. total sum of numbers on faces up is greater than 9.

Solution: 1.Let event A_1 denote sum of faces up is 7. Favourable cases for $A_1 = \{(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)\}$. Also total number of cases are 36. Hence

$$P(A_1) = \frac{\text{Number of favourable cases}}{\text{Total number of cases}} = \frac{6}{36} = \frac{1}{6}.$$

2. Event A_2 denote sum of numbers on up-faces is greater than 9. Then $A_2 = \{(4,6),(6,4),(5,5),(5,6),(6,5),(6,6)\}$. Therefore

$$P(A_2) = \frac{6}{36} = \frac{1}{6}.$$

Example

Example

An urn contains 5 red, 2 black and 4 yellow balls. Two balls are drawn at random from the urn. Find the probability that both balls are of same colour.

Total number of balls are 11. Two balls are drawn out of 11 balls in $\binom{11}{2}$ ways. Let event E_1 denote that both balls are of same colour. If balls are red the number of ways of choosing them are $\binom{5}{2}$. If balls are black the number of ways of choosing them are $\binom{2}{2}$. If balls are yellow the number of ways of choosing them are $\binom{4}{2}$. Therefore number of favourable cases to E_1 is $\binom{5}{2}+\binom{2}{2}+\binom{4}{2}$. Hence required probability is

$$P(E_1)=\frac{\binom{5}{2}+\binom{2}{2}+\binom{4}{2}}{\binom{11}{2}}=\frac{17}{55}.$$

Example

Example

Four persons A, B, C, and D take turns (in the sequence $A, B, C, D, A, B, C, D, A, \ldots)$ in tossing a biased coin.The biased coin has probability $3/4$ of head up. The first person to get a tail wins. We want to determine the probability that B wins. The probability of getting a tail in tossing the coin is $p = 1/4$ and $q = 1 - p$. Then required probability is

$$\begin{aligned} P(B \text{ wins}) &= qp + q^5 p + q^9 p + \cdots \\ &= pq(1 + q^4 + q^8 + \cdots) \\ &= \frac{pq}{1 - q^4} \\ &= 0.274. \end{aligned}$$

Definition

(a) A sequence of events $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$ are said to be monotonically increasing if, for all n, $A_n \subseteq A_{n+1}$

(b) A sequence of events $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$ are said to be monotonically decreasing if, for all n, $A_n \supseteq A_{n+1}$

Definition (Limit of sequence)

(a) For monotonically increasing sequence of events $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$ the limit of sequence of events is defined as

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$$

(b) For monotonically decreasing sequence of events $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$ the limit of sequence of events is defined as

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$$

Example

Example

Let $A_{n-1} = \{\omega : 0 < \omega < \frac{n-2}{n-1}\},\ n = 2, 3, \ldots$. Then $A_1 = \phi,\ A_2 = (0, \frac{1}{2}),\ A_3 = (0, \frac{2}{3}), \ldots$. Hence sequence of events $\{A_n\}_{n=1}^{\infty}$ is monotonically increasing. The limit of sequence is

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$$
$$= \{\omega : 0 < \omega < 1\}.$$

Continuity theorem in probability

Let (Ω, f, P) be a probability model.

(a) If $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$, be monotonically increasing sequence of events, then

$$P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P\left(A_n\right)$$

(b) If $\{A_n\}_{n=1}^{\infty}$, $A_n \in f$, be monotonically decreasing sequence of events, then

$$P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P\left(A_n\right)$$

Summary

To analyze the the algoritmns and computer systems, computer scientists need powerful tools. Many of the tools require the foundation in the probability theory. Hence we require to study the concepts of probability. The Russian mathemematican Kolmogorov (1903-1987) provided foundational work of the modern probability theory. In this chapter we introduced the concepts of random experiment, the sample space and the mathematical definition of probability. Followed by the assignment of the probabilities to the events. Some simple consequences of the axioms of the definition of the probability were also discussed. Finally we discussed the continuity theorem of probability. Various examples are provided to understand the concepts.