Probability Theory

Introduction

Fourier Series and Fourier Transform are widely used techniques to model deterministic signals.

In a typical communication system, the output of an information source (e.g. binary output symbols from a computer) is random in nature. It is not possible to express such random signals by explicit time functions.

However, a random signal may exhibit certain statistical properties which can be modelled in terms of probabilities.

In this way, the behaviour and the performance of a communication system, in the presence of random signals, can be analysed.

Probability theory deals with concepts, not with reality. It is so formulated as to related to the real world but not in an exact manner. Results are derived/deducted from certain axioms.

A central concept in the theory of probability is the event. By associating probabilities with various events in a clearly specified experiment, we can derive useful results from the experiment. Events are combined in various ways to form other events and set theory is used to describe these operations.

Set, Probability, and Relative Frequency

Set definitions :

1. Set - A collection of objects.

2. Elements - objects of the set.

   2.1 Subset of a set A - a set where elements are also elements of the set A.

3. We shall only consider sets whose elements are taken from a largest set S, which we shall call space or universe set. Hence, all sets will be subsets of the set S.

4. An element of a set is any kind of object. It might be a collection of other objects from sets. Hence the name a class of sets or set of sets.

   Eg. $S = \{e_1, e_2\}$, there are 2^2 subsets namely the subsets {empty}, $\{e_1\}$, $\{e_2\}$, $\{e_1, e_2\}$. The element of the set S is e_1 and $\{e_1\}$ is a set with element e_1.

5. A null or empty set φ contains no element. We do not want to exclude the possibility of having an empty set because in performing certain operations on sets, we might end up with an empty set.

6. In probability, we assign probability to various subsets of the space S and not to elements. The subsets are called events. On the other hand, we define functions (eg.

random variables) only on elements.

Probability definitions:

1. Outcomes - The result of an experiment (elements of a sample space S).

2. Sample space - The set of all possible outcomes of the random experiment.

3. Random experiment - An experiment whose outcome is not known in advance.

Eg. If we throw a die and a 4 comes out, the outcome of the random experiment is 4. The outcome is the element 4 and not the event {4}.

At this trial of the experiment, the following events may occur: {4}, {4, 3}, {even}.

4. Event - (Subset of a sample space S). A collection of outcome(s) of a random experiment. Subsets of a sample space S form a Borel field where $A_1 + A_2 + ...$ and $A_1A_2 ...$ are also events.

5. Elementary event - An event has a single element or outcome. Element and outcome are interchangeable in probability theory.

In probability, we might not want to assign probability to every collection of outcomes. For example, it might not be possible to assign probabilities to all events if S has infinite elements. Another reason is that we want to derive other probabilities from some events, not all events, from S.

The universal set is the certain event since it contains all possible outcomes of the experiment, denoted as S.

The empty set is an impossible event since a trial of an experiment must always have one outcome, denoted as φ.

N.B. : If an event has zero probability, the event may not be the impossible event because it may have elements. If an event has unity probability, then it may not be the certain event because it may not have all elements.

1. Consider now a sequence of m independent trails of a random experiment which has n possible outcomes. The occurrence of the outcome is not in any way influenced by the occurrence of any other. Notice also that we are considering trials of a given random experiment. In the sequence of m trials, the event A occurs m_A times, the event B occurs m_B times, and so on.

Thus

$$m_A + m_B + ... + m_n = m \quad (1)$$

and

$$(m_A/m) + (m_B/m) + ... + (m_n/m) = 1. \quad (2)$$

The term m_A/m is defined as the relative frequency of the occurrence of event A.

As m -> ∞, the term m_A/m tends towards a fixed value. This value is called the probability of the event A and is denoted by $P(A)$.

2. Consider now a sequence of m independent trials of a random experiment which has n possible outcomes. The occurrence of the outcome is not in any way influenced by the occurrence of any other. In the sequence of m trials, the outcome a occurs m_a times, the outcome b occurs m_b times, and so on. Thus

$$m_a + m_b + \ldots + m_n = m$$

and

$$(m_a/m) + (m_b/m) + \ldots + (m_n/m) = 1.$$

As m -> ∞, the term m_a/m tends towards a fixed value again. This value is called the probability of the outcome a and is denoted by $P(a)$.

It can be seen that

$$0 \leq m_A/m \leq 1 \qquad (3)$$

$$0 \leq P(A) \leq 1 \qquad (4)$$

$$P(S) = 1 \qquad (5)$$

$$P(\varphi) = 0 \qquad (6)$$

Consider the events A and B in m number of trials of a given experiment. Suppose that

1. The event $A\bar{B}$ occurs m_1 times,

2. The event $B\bar{A}$ occurs m_2 times,

3. The event AB occurs m_3 times,

4. The event $\overline{A + B}$ occurs m_4 times.

The above four events are mutually exclusive, so that no more than one of them can occur at any one trial, and their union is the certain event so that at least one of them must occur at any trial. Thus

$$m_1 + m_2 + m_3 + m_4 = m. \qquad (7)$$

Also

$$m_A = m_1 + m_3 \qquad (8)$$

$$m_B = m_2 + m_3 \qquad (9)$$

$$m_{AB} = m_3 \qquad (10)$$

$$m_{A+B} = m_1 + m_2 + m_3. \quad (11)$$

Thus

$$(m_{A+B}/m)=(m_A/m)+(m_B/m)-(m_{AB}/m) \quad (12)$$

so that

$$P(A+B) = P(A) + P(B) - P(AB). \quad (13)$$

$P(AB)$ is called the joint probability of the event AB. If A and B are mutually exclusive,

$$P(AB) = 0 \quad (14)$$

and

$$P(A+B) = P(A) + P(B). \quad (15)$$

Summary

1. $P(A) \geq 0$

2. $P(S) = 1$

3. $P(A+B) = P(A) + P(B) - P(AB)$

4. $P(AB) = 0$ for mutually exclusive events A and B.
   $P(A+B) = P(A) + P(B)$

1, 2 and 4 are axioms. 3 can be deducted from these axioms, although we did the reverse here.

Conditional Probability

In many applications, we often perform experiment that consists of sub-experiments.

Suppose a random experiment E consists of two sub-experiments E_1 and E_2. We take the event A from experiment E_1 and the event B from experiment E_2, and consider the event AB as the event from experiment E.

AB is the intersection of events from experiments E_1 and E_2 and $P(AB)$ is called the joint probability of the event AB in the experiment E.

Quite often, the occurrence of event B may depend on the occurrence of event A. As before, consider the events A and B in m number of trials of a given experiment.

Suppose that

1. The event $A\bar{B}$ occurs m_1 times,

2. The event $B\bar{A}$ occurs m_2 times,

3. The event AB occurs m_3 times,

4. The event $\overline{A + B}$ occurs m_4 times.

The relative frequency of the event A, on the condition that B has occurred, is the number of times both A and B occur divided by the number of times B occurs.

$$m_B = m_2 + m_3 \tag{16}$$

$$m_{AB} = m_3. \tag{17}$$

Therefore,

$$P(A/B) = m_3 /(m_2 + m_3) \tag{18}$$

$= $ conditional probability of A given B.

Similarly,

$$P(B/A) = m_3 /(m_2 + m_3) \tag{19}$$

$$= m_{BA}/m_A$$

From the assumption concerning A and B,

$$P(A) = m_A / m = (m_1 + m_3) / m$$

$$P(B) = m_B / m = (m_2 + m_3) / m$$

$$P(AB) = m_{AB} / m = m_3 / m$$

$$P(A/ B) = m_{AB} / m_B = m_3 / (m_2 + m_3)$$

$$P(B/ A) = m_{BA} / m_A = m_3 / (m_1 + m_3)$$

It is also assumed that $P(A) \neq 0$ and $P(B) \neq 0$. Thus

$$P(B)P(A/ B) = P(AB)$$

or

$$P(A/ B) = P(AB) / P(B) \tag{20}$$

and

$$P(A)P(B/ A) = P(AB)$$

or

$P(B/A) = P(AB) / P(A)$. (21)

If $P(AB) = P(A)P(B)$, (22)

then the events A and B are statisticallyindependent events. Now from equations (20) and (22),

$P(A/B) = P(A)$ (23)

and from eqns. (21) and (22)

$P(B/A) = P(B)$. (24)

Eg.1 If A and B have no common elements (i.e. mutually exclusive). Under these conditions $AB = \varphi$ and $P(A/B) = [P(AB) / P(B)] = 0$.

Eg.2 Given Figure 1, find $P[(A+B)/M]$.

Figure 1

$AB = \varphi$ $(AM)(BM) = \varphi$

$P[(A+B)M]$ $= P[AM + BM]$

$= P(AM)+P(BM)-P[(AM)(BM)]$

$= P(AM) + P(BM)$

$P[(A+B)/M]$ $= P[(A+B)M]/P(M)$

$= P(A/M) + P(B/M)$

Total Probability

Consider n mutually exclusive events $A_1,\ A_2,\ \ldots,\ A_n$ where

$A_1 + A_2 + \ldots + A_n = S$. (25)

Figure 2

From equation (25)

$B = BS$ $= B(A_1 + A_2 + \ldots + A_n)$

$= BA_1 + BA_2 + \ldots + BA_n$

and

$$P(B) = P(BA_1 + BA_2 + \dots BA_n)$$

$$= P(BA_1) + P(BA_2) + \dots + P(BA_n).$$

Since events BA_i are mutually exclusive, but

$$P(BA_i) = P(B/A_i)P(A_i).$$

Thus

$$P(B) = P(B/A_1)P(A_1) + \dots + P(B/A_n)P(A_n)$$

$$= \sum_{i=1}^{n} P(B/A_i)P(A_i). \quad (26)$$

Bayes' Theorem

It has been shown that for two events A and B,

$$P(AB) = P(A/B)P(B)$$
$$= P(B/A)P(A).$$

Thus

$$P(A/B) = P(B/A)P(A) / P(B). \quad (27)$$

To calculate $P(A_i/B)$ where A_1, A_2, ..., A_n are n mutually exclusive events, we simply substitute equations (26) into (27) and replace A by A_i. Thus

$$P(A_i/B) = P(B/A_i)P(A_i) / \sum_{i=1}^{n} P(B/A_i)P(A_i) \quad (28)$$

$P(A)$ - a priori probability

$P(A_i/B)$ - a posteriori probability

Summary

1. In probability theory, outcomes or elements of a clearly specified experiment form a set S called space or certain event.
2. Events are subsets of S.
3. Events are mutually exclusive if they have no common elements.
4. Assignment of probability to each event must satisfy the 3 axioms.
5. P(AB) = $P(A)P(B)$ --> statistically independent events.

6. $P(A/B) = P(AB)P(B)$ - conditional probability of A.

7. $P(B/A) = P(AB)/P(A)$ - conditional probability of B.

8. See equation (28) for n mutually exclusive events A_1, A_2, ..., A_n where $A_1 + A_2 + \ldots + A_n = S$.

Combined Experiments

As we have already seen, one often deal with 2 or more (sub-)experiments or repeated trials of an experiment. To analyse such problems in a proper manner, we must conceive a combined experiment whose outcomes are multiples of objects belonging to the (sub-)experiments.

Product of Spaces

Given n experiments E_1, E_2, ..., E_n, the combined experiment resulting from all the experiments is defined as

$$E = E_1 \text{ x } E_2 \text{ x } \ldots \text{ x } E_n \qquad (29)$$

If S_i is the space of E_i for $i = 1, 2, \ldots, n$, the space S of E is the cartesian product of the spaces S_1, S_2, …, S_n. We write in the form

$$S = S_1 \text{ x } S_2 \text{ x } \ldots \text{ x } S_n \qquad (30)$$

It is important to note that the elements of a cartesian product are ordered objects. For example, {1, 2} is different from {2, 1}.

The events of the combined experiment are sets of the form A_1 x A_2 x ... x A_n where A_i belongs to S_i. Note that all events in E must form a Borel field so that the sums and products of the events are also events in E.

In many applications, events are independent. For independent events, it can be shown that the probability of the event A in E is

$$P(A) = P_1(A_1)\, P_2(A_2) \ldots P_n(A_n) \qquad (31)$$

where $P_i(A_i)$ is the probability of event A_i in E_i. It can be seen that $P(A)$ cannot be determined from the probabilities $P_1(A_1)$, $P_2(A_2)$, ..., $P_n(A_n)$ of each experiment alone unless all events are independent events.

Sum of Spaces

Given n experiments E_1, E_2, ..., E_n, the combined experiment resulting from all the experiments is defined as

$$E = E_1 + E_2 + \ldots + E_n \qquad (32)$$

If S_i is the space of E_i for $i = 1, 2, \ldots$, n, the space S of E is

$$S = S_1 + S_2 + \ldots + S_n \tag{33}$$

Let A_i be the event of the space S_i in E_i. Thus, the events of the combined experiment are sets of the form

$$A = A_1 + A_2 + \ldots + A_n \tag{34}$$

If S_i and S_j are mutually exclusive spaces for $i \neq j$, we have

$$P(S) = P(S_1) + P(S_2) + \ldots + P(S_n) \tag{35}$$

and

$$P(A) = P(A_1) + P(A_2) + \ldots + P(A_n) \tag{36}$$

We need to find $P(A_i)$ in terms of $P_i(A_i)$ and $P(S_i)$ of E_i . Since A_i S_i, the probability of event A_i in E_i is

$$P(A_i / S_i) = P_i (A_i) \tag{37}$$

From conditional probability, we have

$$P(A_i) = P(A_i / S_i)P(S_i) \tag{38}$$

Therefore,

$$P(A_i) = P_i(A_i)P(S_i) \tag{39}$$

and

$$P(A) = P_1(A_1)P(S_1) + P_2(A_2)P(S_2) + \ldots + P_n(A_n)P(S_n) \tag{40}$$

It can be seen that $P(A)$ can be determined from the probabilities $P_1(A_1)$, $P_2(A_2)$, ..., $P_n(A_n)$, $P(S_1)$, $P(S_2)$, ..., $P(S_n)$ of each experiment alone if all events are mutually exclusive.

References

[1] Taub, H. and Schilling, D. L., ‘Principles of Communication Systems’, 2/e, McGraw-Hill, 1987.

[2] Shanmugam, K. S., ‘Digital and Analog Communication Systems’, John Wiley & Sons, 1979.

[3] Papoulis, A, ‘Probability, Random Variables, and Stochastic Processes’, McGraw-Hill, 1965.

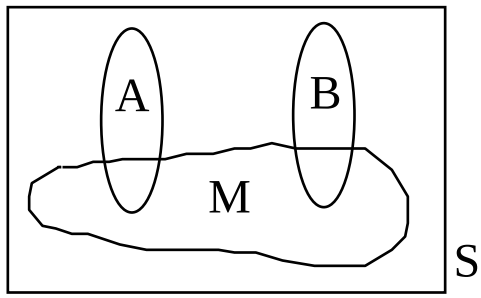

Figure 1

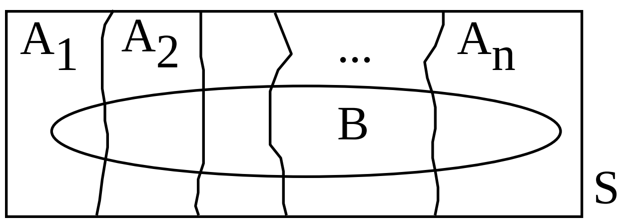

Figure 2