Dr Jezdimir Knezevic

Mechanics of Probability

Published by MIRCE Science Limited

Woodbury Park, Exeter, EX5 1JJ, UK.
Phone: +44 (0) 1395 233 856

ISBN 978-1-904848-05-9
EAN 9781904848059

Preface

This book is written for mechanical, electrical, aeronautical, marine, civil, chemical and all other deterministically educated engineers who are daily creating airplanes, cars, ships, trains and all other machines. However, this book is not written to tell them how to deal with their functionality performance but to give them the fundamentals of the mathematical knowledge required dealing with their functionability[1] performance. In simplest terms, at the design stages of any machine the functionality performance, like speed, acceleration, stopping distance, reaction time, take off power and similar have to be predicted, but at the same time answers to the following questions have to be provided:

Will a given machine be functionable until a given instant of time?
Will a given machine be functionable at a given instant of time?
Will a given machine state be tested by a given instant of time?
Will functionability resources be provided by a given instant of time?
Will a given machine stay functionable during a stated period of time?
Will a given machine be functionable at a given instant of time?

Although these questions are related to very different properties of a machine, for which different types of engineers are responsible, they all have answers of the common nature, which is;

"Yes, with a probability of x." or "No, with a probability of y"

However, answers of this type cause huge and fundamental difficulties, among mechanical, electrical, aeronautical, civil and other types of engineers, in understanding and predicting the numerical values of x and *y*. The reason for that is the fact that all types of engineering are based n scientific disciplines in which relationships are governed

[1] Functionability, n, ability to function, introduced by Dr Knezevic in the book "Reliability, Maintainability and Supportability – a probabilistic approach, McGraw Hill, London, 1993. [1]

by rules and laws of a deterministic nature. Consequently, the measures that describe quantitatively describe the functionality performance of machines have well understood and uniquely defined values. For example, speed, acceleration, weight, volume, capacity, voltage, length, width, shape, road clearance, size of memory, centre of gravity, conductivity and many others are single-value numbers for a machine under consideration. Thus, classical engineers are people whose minds are orientated towards a deterministic way of thinking. However, when it is necessary to provide the probabilistic answers, as it is the case with questions raised above, it becomes clear that the deterministic approach does not work. Consequently, the concept of probability has to be embraced into engineering vocabulary, before it enters into engineering methods and practices.

Probability is an abstract entity that obtains a physical meaning only when behaviour of a large sample of a given machine is considered. Hence, understanding and predicting the numerical values of "x" is reduced to the physical observations and analysis of the functionability performance of a large sample of a given machine, resulting from the occurrences of physically observable and measurable functionability events.

Consequently, to use the concept of probability in the dealings with the functionability properties of machines, it is necessary to learn the terminology, definitions and rules of probability theory, in the manner understandable to, deterministically educated and trained engineers.

However, it is important to stress that this book is not intended to be a rigorous treatment of all relevant theorems and proofs. The intention is to provide an understanding of the main concepts behind probability theory and to show practical applications of the existing theorems and rules on the observable functionability measures of a machine in-service life. At the same time, this is not a book that explains how to generate probability functions but how to use it once they are known.

As a mechanical engineer who has spent over thirty years trying to mathematically and scientifically understand and apply the concept of probability to the concept of functionability, which is the foundation of Mirce Mechanics, I would be extremely happy if my experience, as summarised in this book, can be of some value to all existing durability, reliability, maintainability, supportability and similar type of engineers and managers and students alike who have chosen to become a part of this exciting and challenging profession.

Dr J. Knezevic
Woodbury Park
26th March 2013

P.S. I wish to emphasise that whatever is covered by this book is mathematically and scientifically correct but there is much more that has not been addressed. However, all that is not covered could not be understood without understanding material presented in this book.

1. Introduction

"We do not know how to predict what would happen in any given circumstances, and we believe now that it is impossible, that the only thing that can be predicted is the probability of different events." R. Feynman

According to Einstein "Everything that the human race has done and thought is concerned with the satisfaction of felt needs". During the history of the human civilisation an endless number of machines have been created to satisfy endless human needs. Hence, humans have created ships, airplanes, tractors, computers, cars, radars, and other machines. The designed-in capability of any machine to satisfy felt needs by delivering the required function, with a physically measurable performance like speed, acceleration, power, fuel consumption, breaking distance and many others, is known as functionality. The functionality performance of any machine results from well known and understood physical processes, all of which are accurately predictable by the laws of science. The essential features of these laws are determinism, reversibility and independence of time, location and human impacts.

However, to deliver inherent functionality machines live their lives in a natural environment governed by human decisions and action. Thus, a life of machines is continuously shaped by very rich interactions that generate physically observable phenomena like failures, accidents, services, inspections, repairs, modification, replacements, cannibalisation, demands for spare parts, necessary training, transportation delays, storage damage, and so forth. These phenomena determine the characteristic of a machine known as functionability[2], which is measured through reliability, punctuality, availability and similar characteristics. The functionability of machine is an emerging property of rich interactions between internal components of machines and the interaction of a machine with the natural world and human actions. The essential features of these interactions are indeterminism, irreversibility and dependence of time, space and human impacts, and as such they are not predictable by the laws of science used for the predictions of their functionality.

Consequently, to scientifically understand the mechanics that generate functionability events in time that shape the functionability performance of a machine I established the MIRCE Akademy at Woodbury Park, in 1999. Our staff, students, fellows and members study mechanisms that generate emerging functionability events, in scale from 10^{-10} to 10^{10} metre, through the life of machines like aircraft, racing cars, ships, trains and similar to understand their complexity and dynamics. Our research established that:

1. Each Machine has unique functionability pattern in respect to events and their timing
2. There is a functional and spatial interdependence between the parts of a machine.
3. Natural and human environment shape emerging functionability events

Having taken onboard the above observed facts; I have created a mathematical scheme for the prediction of the motion of the functionability events through the life of a given machine. This has given birth to Mirce Mechanics, whose axioms, mathematical formulas, rules and computational methods enable predictions of a functionability performance to be made with a probabilistic regularity.

Consequently, the satisfaction of human needs through time depends on combined effects of functionality and functionability performance of machines. Thus, deterministically educated engineers are needed for creating functionality performance and probabilistically educated engineers are needed for the creation of functionability performance and their interactions will determine the reliability, cost and effectiveness of the satisfaction of human needs.

Mechanical, electrical, civil, aerospace, marine, manufacturing and other types of engineers are very familiar with the above statements, apart from the expression

---

[22] Functionability is emerging characteristic of a machine life that defines the ability to function through time. in Knezevic, J., Reliability, Maintainability and Supportability – A probabilistic Approach, Text and Software package, pp. 291, McGraw Hill, London 1993. ISBN 0-07-707691-5

“probabilistic regularity”. That is concept that does not exist in their vocabularies. Daily used laws and formulas developed by the great scientist like Newton, Maxwell, Faraday, Hook, La place, Hamilton, Bernoulli and many others are related to deterministic regularity, which is one where all initial parameters and conditions are defined the future always deliver precisely predicted results. However, in Mirce Mechanics the future always delivers different results, which cannot be predicted precisely.

Consequently, the concept of probability has to be embraced. It is an abstract entity that obtains a physical meaning only when behaviour of a large sample of a given machine is considered. Hence, understanding the answers to the questions raised, is reduced to the physical observations and analysis of the trajectories of the motion of a large number of given machines through functionability states, resulting from the occurrences of physically observable and measurable functionability events.

Consequently, to use the concept of probability in the dealings with the functionability properties of machines, it is necessary to learn the terminology, definitions and rules of probability theory, in the manner understandable to, deterministically educated and trained engineers and managers.

1.1 The Nature of Probability Theory

Probability theory is a mathematical discipline with aims similar to those, for example, of geometry or analytical mechanics. In each field the following three aspects of the theory must be distinguished

- The formal logical content
- The intuitive background
- The applications.

The character of probability theory cannot be understood and appreciated without considering all three aspects in their dependencies.

1.1.1 Formal Logical Content

Mathematics is a human creation that is axiomatically concerned with relations among undefined things. For example, geometry does not even try to discuss what a point or a line "really are." They remain undefined entities, as the axioms of geometry defy the relations among them. For example the first axiom states that two points determine a line. Thus, axioms are the rules, and there is nothing magical about them. Hence, in mathematics any statement is true if and only if it is a logical outcome of the basic axioms. Different forms of geometry are based on different sets of axioms, and the logical structure of non-Euclidean geometries is independent of their relation to reality. Situation is the same in physics where physicists study the motion of bodies under laws of attraction different from Newton's, and such studies are meaningful even if Newton's law of attraction is accepted as true in nature.

1.1.2 Intuitive Background

The axioms of geometry and of mechanics have an intuitive background. In fact, geometrical intuition is so strong that it is prone to run ahead of logical reasoning. The extent to which logic, intuition, and physical experience are interdependent is a problem beyond the scope of this book. However, what is certain is the fact that intuition can be trained and developed. All of us were novices in mathematics and as such we progress cautiously and relied on individual rules, whereas at later stages, due to experience accumulated, some people developed a natural feeling for concepts such as four-dimensional space.

Even the collective intuition of mankind appears to progress. Newton's notions of a field of force and of action at a distance and Maxwell's concept of electromagnetic waves were at first decried as "unthinkable" and "contrary to intuition." However, today's technology manifested through radio, phone, computer, microwave oven and similar machines have popularised these notions to such an extent that they formed a part of the ordinary vocabulary. Similarly, the modern student has no appreciation of the modes of thinking, the prejudices, and other difficulties against which the theory of probability had to struggle when it was new.

Nowadays newspapers report on samples of public opinion and the magic of statistics embraces all phases of life to the extent that every day we get the statistics of the chances of rain during each day. Thus everyone has acquired a feeling for the meaning of statements such as “the chances of rain are 40 %.” Vague as it is, this intuition serves as background and guide for the first step. It will be developed as the theory progresses and acquaintance is made with more sophisticated applications.

1.1.3 Applications

The concepts of geometry and mechanics are in practice identified with certain physical objects, but the process is so flexible and variable that no general rules can be given. The notion of a rigid body is fundamental and useful, even though no physical object is rigid. Whether a given body can be treated as if it were rigid depends on the circumstances and the desired degree of approximation. A rubber is certainly not rigid, but in discussing the motion of automobiles on ice textbooks usually treats the rubber tyres as rigid bodies. Depending on the purpose of the theory, the atomic structure of matter is usually disregarded and it is typically treated as a single mass point.

In applications, the abstract mathematical models serve as tolls to mechanical, electrical, aeronautical and other types of engineers. In some cases different models can be used to describe the same empirical situation. The manner in which mathematical theories are applied does not depend on preconceived ideas; it is a purposeful technique depending on, and changing with, experience. A philosophical analysis of such techniques is a legitimate study, but it is not within the realm of mathematics, physics, or statistics. The philosophy of the foundations of probability must be divorced from mathematics and statistics, exactly as the discussion of our intuitive space concept is now divorced from the mathematical concept of geometry.

1.2 The Statistical Probability

It is a historical fact that probability theory was originally developed to address games of chance. However, by now it is clear that intuitive probability is insufficient for scientific purposes. The observable, or "natural," probability distribution seemed perfectly clear to everyone and has been accepted without hesitation by physicists.

The modern mathematical theory of probability is limited to one particular aspect of "chance" that that might be called physical or statistical probability. Generally speaking, this concept may be characterised as a concept of probabilities that does not refer to judgements but to possible outcomes of a conceptual experiment. Thus, from the outset the existence of the possible outcomes of an experiment (known as the sample space) and the probabilities associated with them. This is analogous to the procedure in mechanics where fictitious models involving two, three, or seventeen mass points are introduced, but all of them are deprived from individual properties. Similarly, in analysing the statistical experiment, probability theory is concerned with the accidental circumstances of an actual experiment. Hence, the object of study is sequences (or arrangements) of possible outcomes. There is no place in our system for speculations concerning the probability that the sun will rise tomorrow.

The astronomer speaks of measuring the temperature at the centre of the sun or of travel to Sirius. These operations seem impossible, and yet it is not senseless to contemplate them. By the same token, it is unnecessary to worry whether or not a mentally conceived experiment can be performed; as it is perfectly possible to analyse abstract models. In the back of our minds we keep an intuitive interpretation of probability, which gains operational meaning in certain applications. We imagine the experiment performed a great many times. An event with probability 0.6 should be expected, in the long run, to occur sixty times out of a hundred. This description is deliberately vague but supplies a picturesque intuitive background sufficient for the more elementary applications. As the theory proceeds and grows more elaborate the operational meaning and the intuitive picture will become more concrete.

Probabilities, in Mirce Mechanics, play the same role as masses in classical mechanics. The motion of the planetary system can be discussed without knowledge of the individual masses and without contemplating methods for their actual measurements. Even models for non-existent planetary systems may be the subjects of a beneficial and challenging study. Similarly, practical and useful probability models may refer to non-observable worlds. Probability theory would be effective and useful even if not a single numerical value were accessible. [2]

2. Experiments

The word experiment is used to describe any process that can be repeated under given conditions in order to obtain some measurement. For example in chemistry laboratory it is possible to determine that the boiling point of water is 100^{o} C. Given that the experimental conditions remain the same, each result will be always obtained.

However, there are experiments in which the results vary in spite of all efforts to keep the experimental condition the same. For example, the duration of a pit stop of a formula one car clearly demonstrate the point made. The same car, driven by the same driver, coming to the stop at the same place, where the same mechanics, using the same tools, perform the same task, wheels replacement, always take a different amount of time.

2.1 Experiments and Events

An experiment is a well-defined act or process that leads to a single well-defined outcome.

This definition is generally accepted terminology in probability theory, to represent any process, trial, action or activity related to a real life situation. Thus, every experiment must: be capable of being described, so that the observer knows when it occurs. Hence each experiment has one and only one outcome, so that the set of all possible outcomes can be specified.

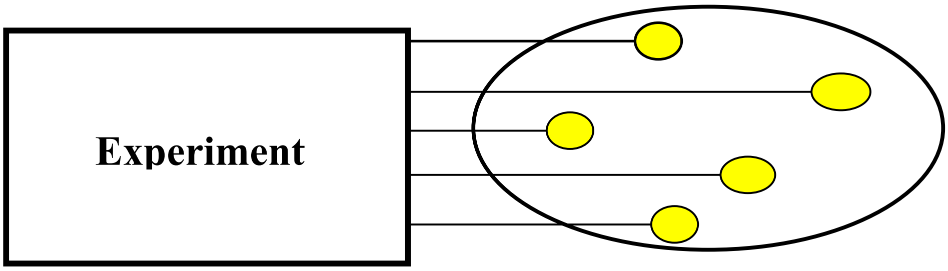


Figure 2.1 Graphical Representation of an Experiment and its Outcomes.

Generally speaking experiments can be divided into following two categories:

- Deterministic Experiments, where the same result is obtained each time the same experiment is repeated.
- Statistical Experiments, where different results are obtained each time the same experiment is repeated.

…………………………………………………………………………

3. THE CONCEPT OF THE PROBABILITY SYSTEM

"Who start in certainty end in doubt"

The physical manifestations of the Probability theory could be seen in everyday life situations where the outcome of a repeated process, experiment, test, or trial is a priority unknown and a prediction has to be made.

To use the concept of probability in everyday scientific and engineering practices it is necessary to learn the terminology, definitions and rules of probability theory. It is important to understand that this monograph is not intended to a rigorous treatment of all relevant theorems and proofs. The intention is to provide an understanding of the main concepts behind probability theory and to show practical applications of the existing theorems, rules and definitions in the scientific and engineering applications.

3.1 The Probability Function

The theory of probability is developed from axioms in the same way as algebra and geometry. In practice this means that its elements have been defined together with several axioms that govern their relations. All other rules and relations are derived from them. The full derivation of elementary rules and axioms can be found in Kolmogorov[3].

In cases where the outcome of an experiment is uncertain, it is necessary to assign some measure that will indicate the chances of occurrence of a particular event. Such a measure of events is called the *probability of the event* and symbolized by *P(.),* and for event *A,* is *P(A)*. The function which associates each event *A* in the sample space *S,* with the probability measure *P(A)*, is called the *probability function*[3] - the probability of that event. A graphical representation of the probability function is given in Figure 3.1.

[3] Function is a relation where each member of the domain is paired with only one member of the range.

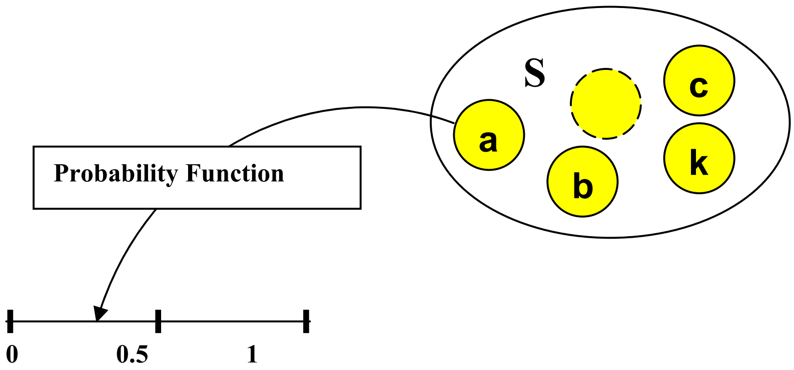


Figure 3.1 Graphical Representation of Probability Function

Formally, the probability function is defined as: ……………………………………..

4. The Concept of the Random Variable

A function that assigns a number (usually a real number) to each sample point in the sample space S is a random variable[4].

Outcomes of experiments may be expressed in numerical and non-numerical terms. In order to compare and analyse them it is much more convenient to deal with numerical terms. So, from the point of view of practicality, it is necessary to assign a numerical value to each possible elementary event in a sample space S. Even if the elementary events themselves are already expressed in terms of numbers, it is possible to reassign a single number to each elementary event. Thus, each elementary event in S can be associated with one, and only one, real number. The function that achieves this is known as the random variable.

Suppose that the symbol X is used to stand for any particular number assigned to any given elementary event. Thus, X is a variable since there will be some set of elementary

[4] The variable is only a placeholder, which can *always* be replaced by any particular element from a set of possibilities, which means that wherever they appear in mathematical expressions they can be replaced by *one* element from some specified set.

events assigned to this value. In other words, a random variable is a real-valued function defined in a sample space. Usually it is denoted with capital letters, such as X, Y and Z, whereas small letters, such as x, y, z, a, b, c, and so on, are used to denote particular values of random variables, see Figure 4.1.

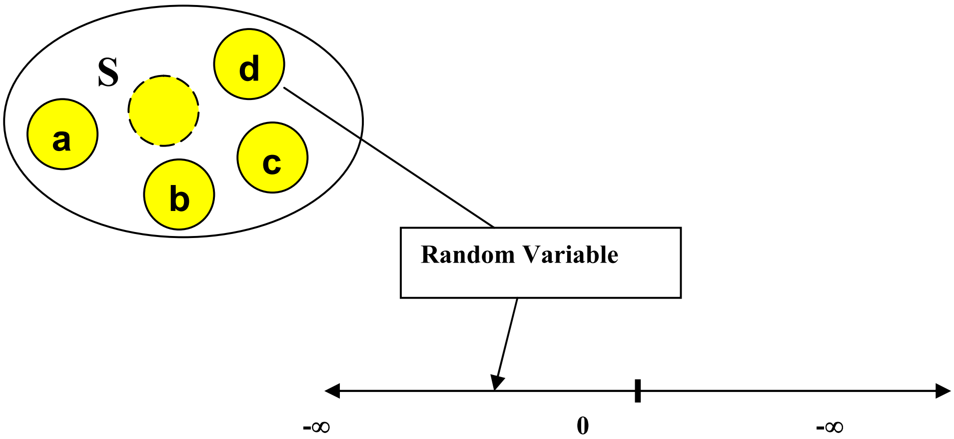


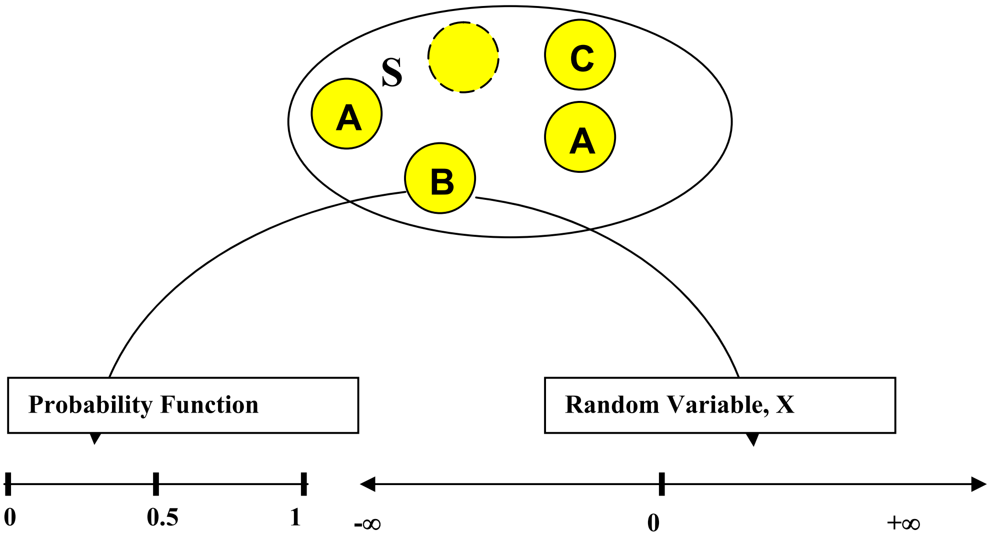
Figure 4.1 Graphical Representation of Random Variable

If X is a random variable and r is a fixed real number, it is possible to define the event A to be the subset of S consisting of all sample points 'a' to which the random variable X assigns the number r, $A = (a : X(a) = r)$. On the other hand, the event A has a probability $p = P(A)$. The symbol p can be interpreted, generally, as the probability that the random variable X takes on the value r, $p = P(X = r)$. Thus, the symbol $P(X = r)$ represents the probability function of a random variable.

Therefore, by means of the random variable it is possible to assign probabilities to real numbers, although the original probabilities were only defined for events of the set S, as shown in Figure 4.2. …………………………………………………………..

5. The Concept of the Probability Distribution of Random Variables

Taking into account the concept of the probability distribution, and the concept of the random variable, given in chapters 3 and 4 respectively, it could be said that the probability distribution of the random variable is a set of pairs, $\{r_i, P(X = r_i),\ i = 1, n\}$ as shown in Figure 5.1.



S	A	B	C					…………………………………
X	x_1	x_2	x_3					
P	p_1	p_2	p_3					
Probability Distribution of a Random Variable								

Figure 5.1 Probability Distribution of a Random Variable

The easiest way to present this set is to make a list of all its members. If the number of possible values is small, it is easy to specify a probability distribution. On the other hand, if there are a large number of possible values, a listing may become very difficult. In the extreme case, where an infinite number of possible values are concerned (for example, all real numbers between zero and one), it is clearly impossible to make a listing. For those cases, in mathematics, there are other methods that could be used for specifying a probability distribution of a random variable, namely: ………………………

6. Discrete Theoretical Probability Distributions

In probability theory, there are several rules that define the functional relationship between the possible values of random variable X and their probabilities, P(X). Rules that have been developed by mathematicians will be analysed here. As they are purely theoretical, i.e. they do not exist in reality, they are called theoretical probability distributions. Instead of analysing the ways in which these rules have been derived, the analysis in this chapter concentrates on their properties.

It is necessary to emphasize that all theoretical distributions represent the family of distributions defined by a common rule through unspecified constants known as parameters of distribution. The particular member of the family is defined by fixing numerical values for the parameters that define the distribution. The probability distributions most frequently used in engineering are examined in the next chapter.

Among the family of theoretical probability distributions that are related to discrete random variables, the Binomial distribution and the Poisson distribution are relevant to the objectives set by this book. A brief description of each now follows.

…………………………………………………….

6.1 Bernoulli Trials

The simple probability distribution is one with only two event classes. For example, a car is tested and one of two events, pass or fail, must occur, each with some probability. The type of experiment consisting of series of independent trials, each of which can eventuate in only one of two outcomes are known as Bernoulli Trials, and the two event classes and their associated probabilities a Bernoulli Process. In general, one of the two events is called a "success" and the other a "failure" or "non-success". These names serve only to tell the events apart, and are not meant to bear any connotation of "goodness" of the event. The symbol p, stands for the probability of a success, q for the probability of failure (p + q =1). If 5 independent trials are made (n = 5), how many different sequences of their outcomes could be observed?

Applying Rule 1 (Section 2), the answer is, $2^5 = 32$. However, it is not necessarily true that all sequences will be equally probable. The probability of a given sequences depends upon p and q, the probability of the two events. Fortunately, since trials are independent, it is possible to compute the probability of any sequence.

For example, let us find the probability of the particular sequence of events (S, S, F, F, S), where S stands for success and F for failure. The probability of first observing an S is p. If the second observation is independent of the first, then:

Probability of $(S,S) = p \times p = p^2$

The probability of an F on the third trial is, so that probability of (S, S) followed by F is p^2q. In the same way probability of $(S,S,F,F) = p^2q^2$ and that of the entire sequence is: $p^2q^2p = p^3q^2$. The same argument shows that the probability of the sequences (S, F, F, F, F) is pq^4, that $(S,S,S,S,S) = p^5$, of $(F,S,S,S,F) = p^3q^2$, and so on.

If all possible sequences and their probabilities are written down the following fact emerges: The probability of any given sequences of n independent Bernoulli Trials depends only on the number of successes and p. This is regardless of the order in which successes and failure occur in sequence, the probability is: $p^r q^{n-r}$ whereas the number of successes, and $n-r$ is the number of failures.

……………………….

6.2 The Binomial Distribution

The theoretical probability distribution, which pairs the number of successes in n trials with its probability, is called the binominal distribution.

This probability distribution is related to experiments, which consist of a series of independent trials, each of which can result in only one of two outcomes: success and or failure. These names are used only to tell the events apart. By convention the symbol p stands for the probability of a success, q for the probability of failure $(p+q=1)$.

The number of successes, x in n trials is a discrete random variable which can take on only the whole values from 0 through n. The formal rule for the probability mass function of the discrete random variable X is:

$$f(x,n) = P(X=x) = \binom{n}{x} p^x q^{n-x}, \qquad 0 < x < n \qquad 6.1$$

where: $$\binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x} \qquad 6.2$$

The binomial distribution expressed in cumulative form, representing the probability that X falls at or below a certain value 'a' is defined by the following equation:

$$P(X \le a) = \sum_{i=o}^{a} P(X = x_i) = \sum_{i=0}^{a} \binom{n}{i} p^i q^{n-i} \qquad 6.3$$

As an illustration of the binomial distribution, the PMF and CDF are shown in Figure 6.1 with parameters n = 10 and p = 0.3.

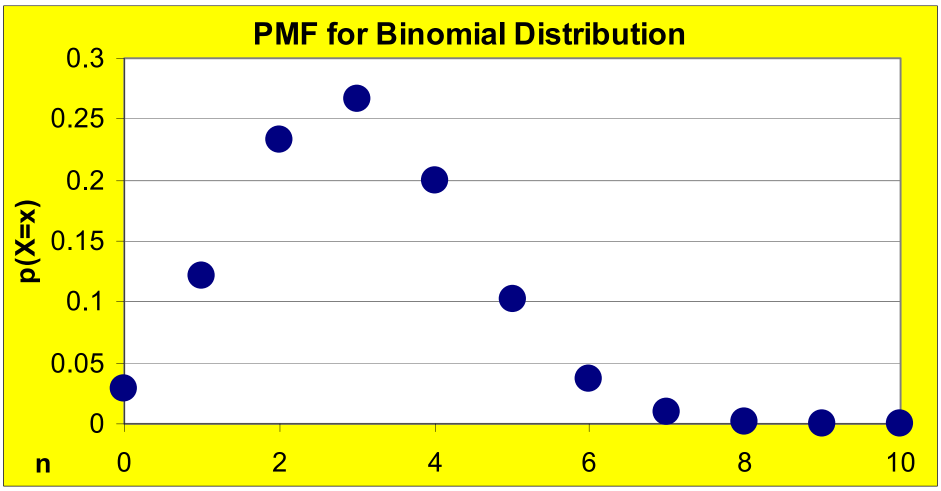


Figure 6.1 PMF for Binomial Distribution, n = 10, p = 0.3

……………………………………………………………..

Example 6.1

Consider a large fleet of cars, where 80 per cent of them have an engine, which can only run on leaded petrol, and 20 per cent of them can only use unleaded petrol. What is the probability that an engineer checking three cars will find one with an unleaded petrol engine?

Solution 6.1

This may be regarded as a Bernoulli process with unleaded being a success and leaded a failure, with corresponding probabilities p = 0.20 and q = 0.80. The required probability can be determined by making use of Equation (6.1), where n = 3, x =1, thus:

$$F(1) = P(X = 1) = \binom{3}{1}(0.20)^1(0.80)^{3-1} = \frac{3\times 2\times 1}{1\times(2\times 1)}\times 0.2\times 0.64 = 0.38$$

Example 6.2

According to past information the probability of producing a defective component is 0.05. Maintaining the same production process in the following six trials determine:

(a) The probability of no defectives
(b) The probability of two defectives
(c) The mean and the standard deviation of the number of defectives

Solution 6.2:

On the assumption that the probability of occurrence of a defective component remains constant from trial to trial, the problem can be viewed as a series of six Bernoulli trials and the required answers could be obtained from the binomial distribution with n = 6, p = 0.05 and q = 0.95. Thus:

(a)

$$P(X=0)=\frac{n!}{x!(n-x)!}p^x q^{n-x}=\frac{6!}{0!(6-0)!}0.05^0\times0.95^6$$

$$=\frac{6\times5\times4\times3\times2\times1}{1\times(6\times5\times4\times3\times2\times1)}0.05^0\times0.95^6=1\times1\times0.74=0.74$$

(b)

$$P(X=2)=\frac{6!}{2!(6-2)!}0.05^2\times0.95^4=0.0305$$

(c) Mean number of defectives $M(X)=np=6\times0.05=0.30$

Standard deviation $SD(X)=\sqrt{npq}=\sqrt{6\times0.05\times0.95}=0.53$

Example 6.3

A product is claimed to be 90 per cent free of defects. What is the expected value and standard deviation of the number of defects in a sample of four?

Solution 6.3:

In this example *n = 4, p = 0.10* and *q = (1 - 0.1) = 0.9*, $E(X)=4\times0.10=0.40$ and $SD(X)=\sqrt{4\times0.10\times0.90}=0.60$

The same results can be achieved in a more tedious way by the direct application of Equation (6.4) for E(X) and (6.5) for V(X), as shown in Table 6.1.

Table 6.1 Solutions to Example 6.3

Defects	Individual probability	$E(X)$	$E(X^2)$
0	0.6561	0.0000	0.0000
1	0.2916	0.2916	0.2916

2	0.0486	0.0972	0.1944
3	0.0036	0.0108	0.0324
4	0.0001	0.0004	0.0016
	1.000	0.4000	0.5200

Therefore, the expected number of defects in the sample of four is E[X] = 0.4 with a standard deviation $SD(X) = \sqrt{0.52 - 0.4^2} = 0.6$.

7. CONTINUOUS THEORETICAL PROBABILITY DISTRIBUTIONS

In probability theory, there are several rules, which define the functional relationship between the possible values of random variable X and their probability rules, $P(X)$. As mathematicians on theoretical bases have developed these rules, they are called theoretical probability distributions. It means that they do not exist in physical reality, but they could be used to represent physical phenomena. Instead of analysing the ways in which these rules have been derived, the analysis in this chapter concentrates on their properties.

It is necessary to emphasize that all theoretical distributions are developed as a family of distributions defined by a common rule through a generic constants known as parameters of distribution. The particular member of the family is defined by fixing numerical values for the parameters that define the distribution. The probability distributions most frequently used in Mirce Mechanics to measure reliability, maintainability and supportability properties of machines are examined in this chapter. The most frequently used rules for distribution functions are exponential, normal, lognormal, and Weibull. Each of them will be discussed in this chapter.

Each of the above mentioned rules define a family of distribution functions. Each member of the family is defined with a few parameters that in their own way control the distribution. All parameters can be classified in the following three categories:

Scale parameter, A, which defines the location of the distribution on the horizontal scale.

Shape parameter, B, which controls the shape of the distribution curves.

Source parameter, C, which defines the origin or the minimum value that a random variable can have.

Thus, individual members of a specific family of the probability distribution are defined by fixing numerical values for the above parameters.

…………………………………

Example 7.1

On average, a machine breaks down once every 10 days. Find the chance that less than five full days will elapse between breakdowns.

Solution 7.1

Based on the available data, the time to failure event, T of the machine considered is distributed in accordance to the exponential probability function with a scale parameter of 10 days. Accordingly, the required probability is determined by the value of the CDF where x =5, thus:

$$P(T \le 5) = F(5) = 1 - \exp^{\left[-\left(\frac{5}{10}\right)\right]} = 1 - 0.61 = 0.39$$

Example 7.3

A catering department installed 2000 coffee machines, which have an average life of 2000 hours with a standard deviation of 400 hours.

How many machines might be expected to fail in the first 1000 operating hours?

What is the probability of a coffee machine failing between 1800 and 2400 operating hours?

After how many hours would one expect 5 percent of the machines to have failed?

Solution 7.3

The operating life of this particular machine can be represented with a random variable, T, , whose probability distribution is defined, as N (2000,400).

$P(T \le 1000) = F(1000) = ?$, According to Equation (7.7), $z = \left(\frac{1000 - 2000}{400} \right) = -2.5$

From Table T1 (given in the appendix) the required probability is $\Phi(-2.5) = 0.00621$. Thus, the expected number of failed coffee machines is $2000 \times 0.00621 = 12.4 = 12$.

$P(1800 \le T \le 2400)$ =?. The required probability can be determined by making use of Equation (7.10), thus: $F(2400) = \Phi(1.0) =$ 0.84135 and. $F(1800) = \Phi(-0.5) =$ 0.3085 Therefore, the probability of a coffee machine failing in the specified interval of operating time is $P(1800 \le T \le 2400) = 0.84 - 0.31 = 0.53$

Here, the task is to determine the value of t that corresponds to a probability of 0.05, which could be expressed in the following way, $F(?) = P(T \le ?) = 0.05$.

According to Table 1, for the Standardise Normal Distribution, the cumulative distribution function, F(z) for z = -1.64 corresponds to the specified probability of 0.05. Thus, the task is to find the numerical value of t for which z = -1.64.

$$z = -1.64 \equiv \Phi\left(\frac{t-2000}{400}\right)$$

Making use of Equation 7.7, the solution becomes: $\frac{t-2000}{400} = -1.64 \rightarrow t = 1344 \ \ hr$

Example 7.5

If we are interested in the age distribution of motor vehicles in the UK that possess a valid MOT certificate, the distribution function will be defined by three parameters because the certificate is not needed for vehicles younger than three years. Thus, the random variable T is defined as LN(1.75, 0.57, 3). What percentage of the vehicles is less than five years old?

Solution 7.5

$$P(\text{Age of motor vehicle with certificate} \leq 5) = F(5) = \Phi\left(\frac{\ln(5-3)-1.75}{0.57}\right) = 0.03$$

Thus, only three percent of vehicles with a valid MOT certificate are less than 5 years old.

Example 7.6

Assuming that the operational life of a certain component can be represented by the Weibull distribution with B = 4, A =2000, and C =1000, find the probability that the component will not fail in the first 1500 hours.

Solution 7.6

The required probability can be calculated by applying Equation (7.27), thus:

$$P(T > 1500) = 1 - [1 - P(T \le 1500)] = F(1500) = \exp\left[-\left(\frac{1500-1000}{2000-1000}\right)^4\right] = 0.939$$

Example 7.7

A large number of identical relays have the time to first failure that follow a Weibull distribution with the parameters, A = 16 years and B = 0.5.
What is the probability that a relay will fail during year 1 of operation?
What is the probability that a relay will fail during year 5 of operation?
What is the expected mean time to failure?

Solution 7.7

a.) $F(1\ \ Year) \Rightarrow P(x \le 1\ year) = 1 - \exp\left[-\left(\frac{1}{16}\right)^{0.5}\right] = 1 - \exp[-0.25] = 0.2212$

b.) $F(5\ \ Year\ ) \Rightarrow P(x \le 5\ year) = 1 - \exp\left[-\left(\frac{5}{16}\right)^{0.5}\right] = 1 - \exp[-0.559] = 0.572$

c.) Using the table T2, when B = 0.5, $\frac{E[T]}{A} = 2.00 \Rightarrow$ E[T] = 16 x 2 = 32 years

8. Jointly Distributed Random Variables

Thus far, probability distributions of single random variables have been addressed. However, we are often interested in probability statements concerning two or more random variables. In order to deal with them, with X and Y two continuous random variables will be denoted. This generic approach that is valid for discrete variables, where integrals are replaced by Sum symbols.

The expression presented below are also applicable to the more general case of n random variables $X_1, \ldots, X_n$.

• Joint cumulative distribution probability function is defined as:

$F(x, y) = P(X \le x, Y \le y)$

• Joint probability density function density, for a given cumulative function, F(x, y), is given by $f(x, y) = \dfrac{\partial^2 F(x, y)}{\partial x \partial y}$

In general, a joint density function is any (integrable) function f(x, y) satisfying the properties

$f(x, y) \ge 0, \qquad \iint f(x, y)\,dxdy = 1$

Usually, f(x, y) will be given by an explicit formula, along with a range (a region in the xy-plane) on which this formula holds. In the general formulas below, if a range of integration is not explicitly given, the integrals are to be taken over the range in which the density function is defined.

…………………………………………………………

9, **Convolution of Independent Random Variables**

In this chapter we turn to the important question of determining the distribution of a sum of independent random variables in terms of the distributions of the individual constituents.

9.1 Convolution of Discrete Random Variables

In this section we consider only sums of discrete random variables, reserving the case of continuous random variables for the next section. We consider here only random variables whose values are integers. Their distribution functions are then defined on these integers. We shall find it convenient to assume here that these distribution functions are defined for all integers, by defining them to be 0 where they are not otherwise defined.

Suppose X and Y are two independent discrete random variables with probability mass functions $m_1(x)$ and $m_2(x)$. Let $Z = X + Y$. We would like to determine the distribution function $m_3(x)$ of Z. To do this, it is enough to determine the probability that Z takes on the value z, where z is an arbitrary integer. Suppose that $X = k$, where k is some integer. Then Z=z if and only if Y=z-k. Hence, the event Z=z is the union of the pair wise disjoint events $(X = k)$ and $(Y = z- k)$, where k runs over the integers. Since these events are pair wise disjoint, we have

$$P(Z = z) = \sum_{-\infty}^{\infty} P(X = k) \times P(Y = z - k)$$

Thus, we have found the distribution function of the random variable *Z.* This leads to the following definition.

Let X and Y be two independent integer-valued random variables, with distribution functions $m_1(x)$ and $m_2(x)$ respectively. Then the convolution of $m_1(x)$ and $m_2(x)$ is the distribution function $m_3 = m_1 \times m_2$ given by

$$m_3(j) = \sum_k m_1(k) \times m_2(j-k) \quad \text{for } j = ..., -2, -1, 0, 1, 2, ...$$

The function $m_3(x)$ is the distribution function of the random variable $Z = X + Y$

…………………………………………………………………..

9.2 Convolution of Continuous Random Variables

In this section we consider the continuous version of the problem posed in the previous section: How are sums of independent random variables distributed?

Let X and Y be two continuous random variables with density functions $f(x)$ and $g(y)$, respectively. Assume that both $f(x)$ and $g(y)$ are defined for all real numbers. Then the convolution $f \times g$ of f and g is the function given by

$$(f*g)(z)=\int_{-\infty}^{\infty} f(z-y)g(y)dy=\int_{-\infty}^{\infty} g(z-y)f(x)dx$$

This definition is analogous to the definition, given above of the convolution of two discrete distribution functions. Thus it should not be surprising that if X and Y are independent, then the density of their sum is the convolution of their densities. This fact is stated as a theorem below

Theorem*: Let X and Y be two independent random variables with density functions $f_X(x)$ and $f_Y(y)$ defined for all x. Then the sum Z = X + Y is a random variable with density function $f_Z(z)$, where f_Z is the convolution of f_X and f_Y.*

To get a better understanding of this important result, some examples will be used.

Example 9.4

Suppose that two numbers are chosen independently at random from the interval [0; 1] with uniform probability density. What is the probability density of their sum?

Solution 9.4

Let X and Y be random variables describing our choices and $Z = X + Y$ their sum. Thus

$$f_X(a)=f_Y(a)=\begin{cases}1 & \text{if } 0<a<1\\ 0 & \text{otherwise}\end{cases}$$

and the density function for the sum is given by $f_{X+Y}(a)=\int_0^1 f_X(a-y)dy$.

For the values of a between 0 and 1 , $0\le a\le 1$ it follows that $f_{X+Y}(a)=\int_0^a dy=a$

For the values of a between 1 and 2 , 1<a<2 it follows that $f_{X+Y}(a) = \int_{a-1}^{1} dy = 2 - a$

Finally:

$$f_Z(z) = \begin{cases} z & \text{if } 0 \le z \le 1 \\ 2 - z & \text{if } 1 < z < 2 \\ 0 & \text{otherwise} \end{cases}$$

9. 3/ Sum of Two Independent Exponential Random Variables

Suppose that two numbers are chosen at random from the interval [0;1) with an exponential density with parameter λ.. What is the density of their sum?

Example 9.5

It is an interesting and important fact that the convolution of two normal densities with means μ_1 and μ_2 and variances σ_1 and σ_2 is again a normal density, with mean $\mu_1+\mu_2$ and variance $\sigma_1^2 + \sigma_2^2$

We will show this in the special case that both random variables are standard normal. The general case can be done in the same way, but the calculation is more complicated. Suppose X and Y are two independent random variables, each with the standard normal density. We have

$$f_X(x) = f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad \text{then}$$

$$f_Z(z) = f_X(x) \times f_Y(y)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{(z-y)^2}{2}} e^{-\frac{y^2}{2}} dy = \frac{1}{2\pi} e^{-\frac{z^2}{4}} \int_{-\infty}^{\infty} e^{-\left(\frac{y-z}{2}\right)^2} dy = \frac{1}{2\pi} e^{-\frac{z^2}{4}} \sqrt{\pi} \left[\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{y-z}{2}\right)^2} dy \right],$$

The expression in the square brackets is equal to 1, since it is the integral of the normal density function with μ = 0 and $\sigma = \sqrt{2}$, This is follows:

$$f_Z(z) = \frac{1}{\sqrt{4\pi}} e^{-\frac{z^2}{4}}$$

9.4 Independent Trials

We now consider briefly the distribution of the sum of n independent random variables, all having the same density function. If $X_1, X_2, \ldots, X_n$ are these random variables and $S_n = X_1 + X_2 + \ldots + X_n$ is their sum, then we will have

$$f_{Sn}(x) = \left(f_{X_1} \times f_{X_2} \times \ldots\ldots \times f_{X_n}\right)(x)$$

where the right-hand side is an n-fold convolution. It is possible to calculate this density for general values of n in certain simple cases.

10. Limit Theorems

The most important theoretical results in probability theory are limit theorems. From all of them the most important are those classified as "laws of large numbers" or under the heading of "central limit theorem".

10.1 Markov's Inequality

If X is a random variable that takes only nonnegative values, then for any value a>0

$$P(X \geq a) \leq \frac{E(X)}{a}$$

10.2 Chebishev's Inequality

If X is a random variable with finite mean μ and variance σ^2, then for any value k>0

$$P\left\{\left(X-\mu\right)\geq k\right\}\leq\frac{\sigma^2}{k^2}$$

The importance of these inequalities is that they enable us to derive bounds on probabilities when only the mean or both the mean and variance, of the probability distributions are known. Of course, if the actual distributions were known, then the desired probabilities concerned could be exactly computed and there would be no need to seek bounds.

Example 10.1

Assume that it is known that the number of parts produced in a specific factory during a week is a random variable with mean of 50.

a) What is the probability that the next week the number of parts produced will exceed 50?
b) If the variance of weekly production is 25, what is the probability that the next week production will be between 40 and 60?

Solution 10.1:

Let X be the number of parts that will be produced next week.

a) By Markov's inequality $P(X\geq 75)\leq\frac{E(X)}{75}=\frac{50}{75}=0.66$

b) By Chebishev's inequality $P\left\{\left(X-50\right)\geq 10\right\}\leq\frac{25}{10^2}=0.25$

Thus, $P\left\{\left(X-50\right)<10\right\}\geq 1-\frac{1}{4}=0.75$, which means that the probability that the next week production of parts considered will be between 40 and 60 is at least 0.75.

As Chebishev's inequality is valid for all distributions of the random variable X, it is not to be expected that the bounds of probability is always close to the real values.

Proposition 3. If Var(X)=0, then $P\left[X = E(X)\right] = 1$

This practically means that the only random variable having variance equal to 0 are these that are constant with probability 1.

10.3 The Weak Law of Large Numbers

Let $\mathrm{X}_1, \mathrm{X}_2, \ldots\ldots \mathrm{X}_n$, be a sequence of independent and indetically distrubuted random variables, each having finite mean E[X]=μ. Then, for any ε>0,

$$P\left\{\left\|\frac{\mathrm{X}_1 + \mathrm{X}_2 + \mathrm{X}_3, \ldots\ldots \mathrm{X}_n}{n} - \mu\right| > \varepsilon\right\} \to 0 \qquad n \to \infty$$

The weak law of a large numbers was originally proved by Jacob Bernoulli for the special case where the X are 0-1, which is a Bernuolli, are random variables. The general form of this theorem was proved by the Russian mathematician A. Khintchine (1894-1959)

10.4 The Central Limit Theorem

This is one of the most remarkable results in probability theory. Loosely put, it states that the sum of a large number of independent random

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables, each having finite mean and variance. Then the variable $\dfrac{X_1 + X_2 + X_3 + \ldots\ldots X_n - n\mu}{\sigma\sqrt{n}}$ tends to be a standard normal variable as, thus:

$$P\left\{\frac{X_1+X_2+X_3+.....X_n-n\mu}{\sigma\sqrt{n}}\leq a\right\}\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a}e^{-\frac{x^2}{2}}dx,\qquad n\to\infty$$

The first version of the central limit theorem was proved by DeMoirve around 1733 for the special case where the X are Bernoulli random variables with p=1/2. This was extended by Laplace to the case of arbitrary p. However, his proof was not rigours. A truly rigorous proof of the central limit theorem was first made by Russian mathematician Liapounoff in the period of 1901-02.

10.5 The Strong Law of Large Numbers

Probably, this is the best-known result in the probability theory. It states that the average of a sequence of independent random variables, having a common distribution, will, with the probability of 1, converge to the mean of that distribution.

Let $X_1, X_2, ...X_n,$ be a sequence of independent and indetically distrubuted random variables, each having finite mean $\mu = E[X]$. Then, with probability 1,

$$\frac{X_1+X_2+X_3+.....X}{n}\to\mu\qquad n\to\infty$$

That is, the strong law of large numbers states that

$$P\left\{\lim_{n\to\infty}\left(X_1+X_2+....+X_n\right)/n=\mu\right)\right\}=1$$

As an application of SLLN, suppose that a sequence of independent trials of some experiment is performed. Let E be a fixed event of the experiment and denote P(E) the probability that E occurs on any particular trial. Letting:

$$X=\begin{cases}1 & \text{if E occurs on the i}^{\text{th}}\text{ trial}\\0 & \text{if E does not occurs on the i}^{\text{th}}\text{ trial}\end{cases}$$

We have by the string law of large numbers that with probability 1,

$$\frac{X_1 + X_2 + X_3 + ...X}{n} \to E(X) = P(E)$$

Since $X_1 + X_2 + .. + X_n$ represents the number of times that the event E occurs in the first n trials, we may interpret the above expression as a stating that, with probability 1, the limiting proportion of time that the event E occurs is just *P(E).*

The strong law of large numbers was originally proved, in the case of Bernuolli random variables, by the French mathematician E. Borel (1871-1955). The general form of the strong law, presented above, was proved by the Russian mathematician A.N. Kolmogorov (1903-1987).