CHAPTER SIX
ELEMENTS OF PROBABLITY

SPECIFIC OBJECTIVES

At the end of the topic the trainee should be able to:

- Discuss the basic concepts of probability;
- Apply the techniques of counting and set theory;
- Apply the laws of probability;
- Apply probability distribution concepts to decision problems.

Introduction

Probability is the ratio of the number of favorable cases to the total number of equal likely cases. Probability is a very popular concept in business management. This is because it covers the risks which may be involved in certain business situations. It is a fact that when a business investment is being arranged, the outcome is usually uncertain. Therefore the concept of probability may be used to describe the **degree of uncertainty of a particular** business **outcome.** Probability may therefore be defied as the chances of a given event occurring. Numerically, probability values range between 0 and 1. a probability of 0 implies that the event cannot occur at all. A probability of 1 implies that the event will certainly occur.

Therefore other events have their probabilities with values lying between 0 and 1

The formula used to determine probability is as follow

$$\text{Probability (x)} = \frac{r}{n} = \frac{\text{Favourable outcomes}}{\text{Total outcomes}}$$

Uses of Probability in Business

1. Business games of chance e.g. Raffles Lotteries e.t.c.
2. Insurance firms: this is usually done when a new client or property is being insured. The company has to be certain about the chances of the insured risks occurring.
3. Business decision making regarding viability of projects thus the projects with a greater probability has greater chances.

Example

A bag contains 80 balls of which 20 are red, 25 are blue and 35 are white. A ball is picked at random what is the probability that the ball picked is:

(i) Red ball
(ii) Black ball
(iii) Red or Blue ball

Solution

(i) Probability of a red ball = $\frac{\text{Number of red balls in the bag}}{\text{Total number of balls in the bag}} = P(R)$

$$= \frac{20}{80} = \frac{1}{4}$$

(ii) Probability of black ball = $\frac{\text{Number of black balls in the bag}}{\text{Total number of balls}} = P(B)$

$$= \frac{0}{80} = 0$$

(iii) P(R or B) = $\frac{20}{80}$ or $\frac{25}{80} = \frac{20}{80} + \frac{25}{80}$

$= \frac{9}{16}$

Note: in probability or is replaced by a plus (+) sign. See addition rule.

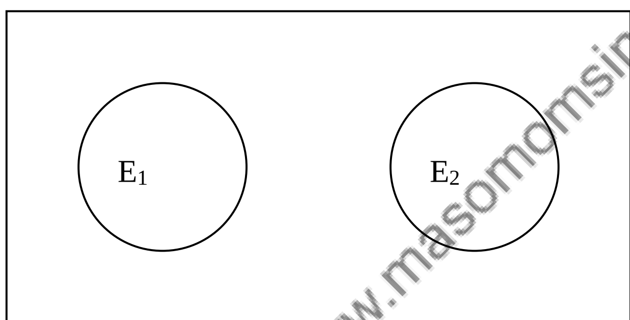
Common terms

Events: an event is a possible outcome of an experiment or a result of a trial or an observation.

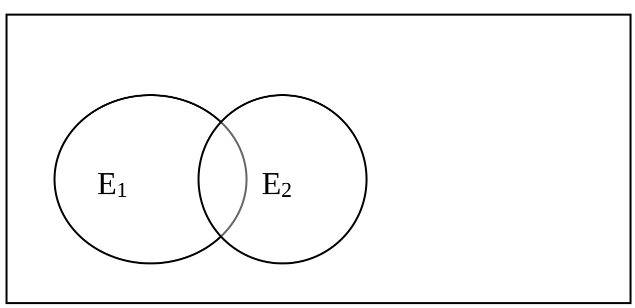
Mutually exclusive events

A set of events is said to be mutually exclusive if the occurrence of any one of the events precludes the occurrence of any of the other events e.g. when tossing a coin, the events are a head or a tail these are said to be mutually exclusive since the occurrence of heads for instance implies that tails cannot and has not occurred.

It can be represented in Venn diagram as.



$E_1 \cap E_2 = \emptyset$



Non-mutually exclusive events (independent events)

$E_1 \cap E_2 \neq \emptyset$

Consider a survey in which a random sample of registered voters is selected. For each voter selected their sex and political party affiliation are noted. The events "KANU" and "woman" are not mutually exclusive because the selection of KANU does not preclude the possibly that the voter is also a woman.

Independent Events

Events are said to be independent when the occurrence of any of the events does not affect the occurrence of the other(s).

E.g. the outcome of tossing a coin is independent of the outcome of the preceding or succeeding toss.

Example

From a pack of playing cards what is the probability of;

(i) Picking either a 'Diamond' or a 'Heart' → mutually exclusive

(ii) Picking either a 'Flower' or an 'Ace' → independent events

Solution.

(i) P (Diamond or Heart)

= P (Diamond) + P (Heart)

$= \frac{13}{52} + \frac{13}{52} = \frac{26}{52}$

= 0.5

(ii) P (Flower or Ace)

= P (Flower) + P (Ace) – P (Flower and Ace)

$= \frac{13}{52} + \frac{4}{52} - \frac{1}{52}$

$= \frac{4}{52}$ = 0.31

Note: that the formula used incase of independent events is different to the one of mutually exclusive.

There are mainly four schools of thought on probability

i) the classical to priors approach

ii) the relative frequency of empirical approach

iii) the axiomatic approach

iv) the personalistic approach

The classical approach

This school of thought assumes that all the possible outcomes of an experiment are mutually exclusive and equally likely.

P (E) = $\frac{\text{number of outcomes favorable to the occurrence of event E}}{\text{Total number of outcome}} = \frac{a}{a+b}$

Relative axiomatic approach

The axiomatic approach theory of probability is an honest attempt at constructing a theory probability large tree from the inadequacies of both the classical and empirical approaches.

Personality approach

According to the personality or subjective concept the probability of an event is the degree of confidence placed in the occurrence of an event by a particular individual based on the evidence available to him.

Relative frequency approach

The relative frequency theoreticians agree that the only valid procedure for determining event probabilities is through repetitive experiments

P (E) = a/n

Rules of Probability

(a) Additional Rule – This rule is used to calculate the probability of two or more mutually exclusive events. In such circumstances the probability of the separate events must be added.

Example

What is the probability of throwing a 3 or a 6 with a throw of a die?

Solution

P (throwing a 3 or a 6) = $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$

(b) Multiplicative rule

This is used when there is a string of independent events for which individual probability is known and it is required to know the overall probability.

Example

What is the probability of a 3 and a 6 with two throws of a die?

Solution

P (throwing a 3) and P (6)

= P (3) and P (6) = $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$

Note: 1) In probability 'and' is replaced by 'x' – multiplication.

2) $P(x)$ and $P(y) \neq P(x \text{ and } y)$ note that these two are different. The first implies P(x) happening and P(y), but if the order of which happened first is unimportant then we have p(x and y).

In the example above:

$$P(3) \text{ and } P(6) = \frac{1}{36}$$

But

P (3 and 6) = P (3 followed by 6) or P (6 followed by 3)

$$= [P(3)\, P(6)] \text{ or } [P(6)\, P(3)]$$

$$= \frac{1}{36} + \frac{1}{36} = \frac{1}{18}$$

Conditional Probability

This is the probability associated with combinations of events but given that some prior result has already been achieved with one of them.

It's expressed in the form of

P (x|y) = Probability of x given that y has already occurred.

$$P(x|y) = \frac{P(xy)}{P(y)} \rightarrow \text{conditional probability formula.}$$

Example:

In a competitive examination, 30 candidates are to be selected. In all 600 candidates appear in a written test, and 100 will be called for the interview.

(i) What is the probability that a person will be called for the interview?

(ii) Determine the probability of a person getting selected if he has been called for the interview?

(iii) Probability that person is called for the interview and is selected?

Solution:

Let event A be that the person is called for the interview and event B that he is selected.

(i) $\therefore P(A) = \frac{100}{600} = \frac{1}{6}$

(ii) $P(B|A) = \frac{30}{100} = \frac{3}{10}$

(iii) $P(AB) = P(A) \times P(B|A)$

$$= \frac{1}{6} \times \frac{3}{10} = \frac{3}{60} = \frac{1}{20}$$

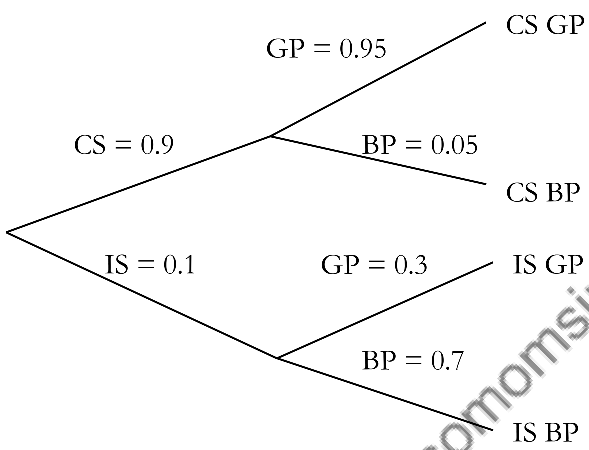
Example:

From past experience a machine is known to be set up correctly on 90% of occasions. If the machine is set up correctly then 95% of good parts are expected but if the machine is not set up correctly then the probability of a good part is only 30%.

On a particular day the machine is set up and the first component produced and found to be good. What is the probability that the machine is set up correctly.

Solution:

This is displayed in the form of a probability tree or diagram as follows:



CS – Correct Setting

IS – Incorrect Setting

GP – Good Product

BP – Bad Product

P (CSGP) = 0.9 × 0.95 = 0.855
P (CSBP) = 0.9 × 0.05 = 0.045
P (ISGP) = 0.1 × 0.3 = 0.03
P (ISBP) = 0.1 × 0.7 = 0.07
1.00

Probability of getting a good part (GP) = CSGP or ISGP
= CSGP + ISGP
= 0.855 + 0.03 = 0.885

Note: Good parts may be produced when the machine is correctly set up and also when it's incorrectly setup. In 1000 trials, 855 occasions when its correctly setup and good parts produced (CSGP) and 30 occasions when it's incorrectly setup and good parts produced (ISGP).

Probability that the machine is correctly set up after getting a good part.

$$= \frac{\text{Number of favourable outcomes}}{\text{Total possible outcomes}} = \frac{P(CSGP)}{P(GP)} = \frac{0.855}{0.885} = 0.966$$

Or

$= P(CS|GP) = \frac{P(CSGP)}{P(GP)} = \frac{0.855}{0.885} = 0.966$

Example

In a class of 100 students, 36 are male and studying accounting, 9 are male but not studying accounting, 42 are female and studying accounting, 13 are female and are not studying accounting.

Use these data to deduce probabilities concerning a student drawn at random.

Solution:

	Accounting A	Not accounting $\bar{A}$	Total
Male M	36	9	45
Female F	42	13	55
Total	78	22	100

$P(M) = \frac{45}{100} = 0.45$

$P(F) = \frac{55}{100} = 0.55$

$P(A) = \frac{78}{100} = 0.78$

$P(\bar{A}) = \frac{22}{100} = 0.22$

P(M and A) = P(A and M) = $\frac{36}{100}$ = 0.36

P(M and $\bar{A}$) = 0.09

P(F and $\bar{A}$) = 0.13

These probabilities can be express differently as;

P(M) = P(M and A) or P(M and $\bar{A}$)
= 0.36 + 0.09 = 0.45

P(F) = P(F and A) or P(F and $\bar{A}$)
= 0.42 + 0.13 = 0.55

P(A) = P(A and M) + P(A and F) = 0.36 + 0.42 = 0.78

$P(\bar{A})$= P($\bar{A}$ and M) + P($\bar{A}$ and F) = 0.09 + 0.13 = 0.22

Now calculate the probability that a student is studying accounting given that he is male.

This is a conditional probability given as P(A|M)

$$P(A|M) = \frac{P(A \text{ and } M)}{P(M)} = \frac{0.36}{0.45} = 0.80$$

From the formula above we get that,
P(A and M) = P(M) P(A|M) ……………….. (i)
Note that P(A|M) ≠ P(M|A)

Since $P(M|A) = \frac{P(A \text{ and } M)}{P(A)}$ this is known as the Bayes' rule.

Bayes' rule/Theorem

This rule or theorem is given by

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

It's used frequently in decision making where information is given the in form of conditional probabilities and the reverse of these probabilities must be found.

Example
Analysis of questionnaire completed by holiday makers showed that 0.75 classified their holiday as good at Malindi. The probability of hot weather in the resort is 0.6. If the probability of regarding holiday as good given hot weather is 0.9, what is the probability that there was hot weather if a holiday maker considers his holiday good?

Solution

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

Let H = hot weather
G = Good

$P(G) = 0.75 \quad P(H) = 0.6$ and $P(G|H) = 0.9$ (Probability of regard holiday as good given hot weather)

Now the question requires us to get

$P(H|G)$ = Probability of (there was) hot weather given that the holiday has been rated as good).

$$= \frac{P(H)P(G|H)}{P(G)} = \frac{(0.6)(0.9)}{0.75}$$

$= 0.72$.

Application

1. A machine comprises of 3 transformers A, B and C. The machine may operate if at least 2 transformers are working. The probabilities of each transformer working are given as shown below;

$P(A) = 0.6, \quad P(B) = 0.5, \quad P(C) = 0.7$

A mechanical engineer went to inspect the working conditions of those transformers. Find the probabilities of having the following outcomes

i. Only one transformer operating
ii. Two transformers are operating
iii. All three transformers are operating
iv. None is operating
v. At least 2 are operating
vi. At most 2 are operating

Solution

$P(A) = 0.6$	$P(\bar{A}) = 0.4$	$P(B) = 0.5$	$P(\sim B) = 0.5$
$P(C) = 0.7$	$P(\bar{C}) = 0.3$		

i. P(only one transformer is operating) is given by the following possibilities

	1st	2nd	3rd	
P	(A	$\bar{B}$	$\bar{C}$)	= 0.6 x 0.5 x 0.3 = 0.09
P	($\bar{A}$	B	$\bar{C}$)	= 0.4 x 0.5 x 0.3 = 0.06
P	($\bar{A}$	$\bar{B}$	C)	= 0.4 x 0.5 x 0.7 = 0.14

∴ P(Only one transformer working)

= 0.09 + 0.06 + 0.14 = 0.29

ii. P(only two transformers are operating) is given by the following possibilities.

	1st	2nd	3rd	
P	(A	B	$\bar{C}$)	= 0.6 x 0.5 x 0.3 = 0.09
P	(A	$\bar{B}$	C)	= 0.6 x 0.5 x 0.7 = 0.21
P	($\bar{A}$	B	C)	= 0.4 x 0.5 x 0.7 = 0.14

∴ P(Only two transformers are operating)

$= 0.09 + 0.21 + 0.14 = 0.44$

iii. P(all the three transformers are operating).

$= P(A) \times P(B) \times P(C)$

$= 0.6 \times 0.5 \times 0.7$

$= 0.21$

iv. P(none of the transformers is operating).

$= P(\bar{A}) \times P(\bar{B}) \times P(\bar{C})$

$= 0.4 \times 0.5 \times 0.3$

$= 0.06$

v. P(at least 2 working).

= P(exactly 2 working) + P(all three working)

$= 0.44 + 0.21$

$= 0.65$

vi. P(at most 2 working).

= P(Zero working) + P(one working) + P(two working)

$= 0.06 + 0.29 + 0.44$

$= 0.79$

POISSON PROBABILITY DISTRIBUTION

This is a set of probabilities which is obtained for discrete events which are described as being rare. Occasions similar to binominal distribution but have very low probabilities and large sample size.

Examples of such events in business are as follows:

i. Telephone congestion at midnight
ii. Traffic jams at certain roads at 9 o'clock at night
iii. Sales boom
iv. Attaining an age of 100 years (Centureon)

Poisson probabilities are frequently applied in business situations in order to determine the numerical probabilities of such events occurring.
The formula used to determine such probabilities is as follows

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

Where x = No. of successes
λ = mean no. of the successes in the sample (λ = np)
e = 2.718

Example 1

A manufacturer assures his customers that the probability of having defective item is 0.005. A sample of 1000 items was inspected. Find the probabilities of having the following possible outcomes

i. Only one is defective
ii. At most 2 defective
iii. More than 3 defective

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

(λ = np = 1000 × 0.005) = 5

i. P(only one is defective) = P(1) = P(x = 1)

$$= \frac{2.718^{-5} \times 5^{1}}{1!} \qquad \text{Note that } 2.718^{-5} = \frac{1}{2.718^{5}}$$

$$= \frac{5}{2.718^{5}}$$

$$= \frac{5}{148.33}$$

$$= 0.0333$$

ii. P(at most 2 defective) = P(x ≤ 2)
= P(0) + P(1) + P(2)

$P(x = 0) = \frac{e^{-5}5^0}{0!}$

$= 2.718^{-5}$

$= \frac{1}{2.718^{+5}}$

$= \frac{1}{148.336}$

$= 0.00674$

$P(1) = 0.0337$

$P(2) = \frac{2.718^5\ 5^2}{2!}$

$= \frac{25}{2 \times 148.336}$

$= 0.08427$

$P(x \leq 2) = 0.00674 + 0.0337 + 0.08427$

$= 0.012471$

iii. P(more than 3 defective) = $P(x > 3)$

$= 1 - [P(0) + P(1) + P(2) + P(3)]$

<u>POISSON MATHEMATICAL PROPERTIES</u>

1. The mean or expected value = np = λ

   Where; n = Sample Size

   p = Probability of success

2. The variance = np = λ
3. Standard deviation = $\sqrt{np} = \sqrt{\lambda}$

<u>Example</u>

The probability of a rare disease striking a given population is 0.003. A sample of 10000 was examined. Find the expected no. suffering from the disease and hence determine the variance and the standard deviation for the above problem

<u>Solution</u>

Sample size n = 10000

P(*a person suffering from the disease*) = 0.003 = p

$\therefore$ expected number of people suffering from the disease

Mean = λ = 10000 × 0.003

= 30

= np = λ

Variance = np = 30

Standard deviation = $\sqrt{np}$ = λ

$= \sqrt{30}$

= 5.477

NORMAL PROBABILITY DISTRIBUTION

In a continuous distribution, the variable can take any value within a specified range, e.g. 2.21 or 1.64 compared to the specific values taken by a discrete variable e.g. 1 or 3. The probability is represented by the area under the probability density curve between the given values.

The uniform distribution, the normal probability distribution and the exponential distribution are examples of a continuous distribution

The normal distribution is a probability distribution which is used to determine probabilities of continuous variables

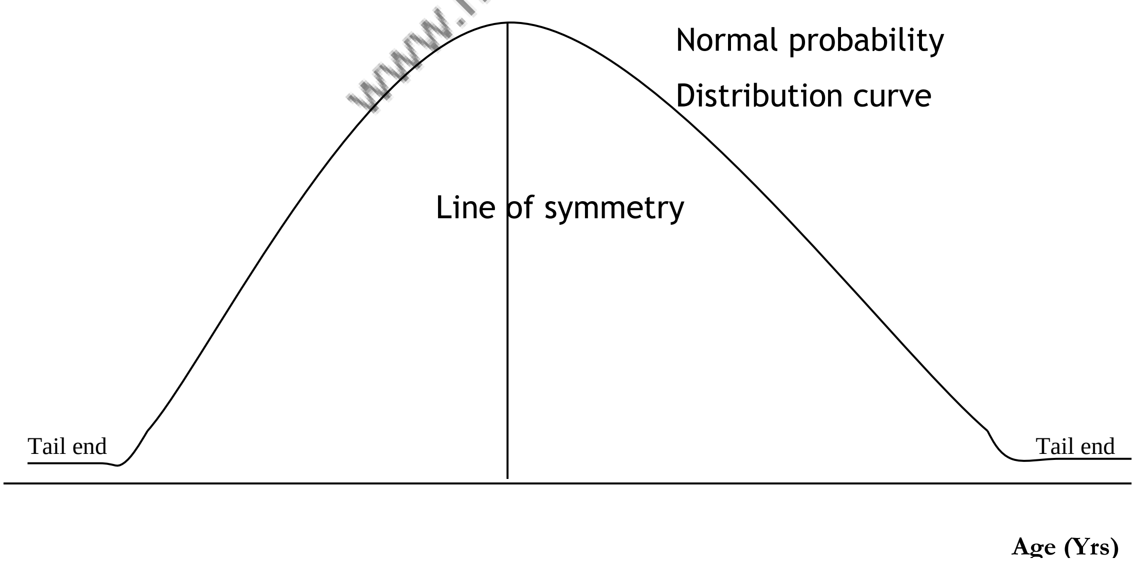
Examples of continuous variables are

- Distances
- Times
- Weights
- Heights
- Capacity e.t.c

Usually continuous variables are those, which can be measured by using the appropriate units of measurement.

Following are the properties of the normal distribution:

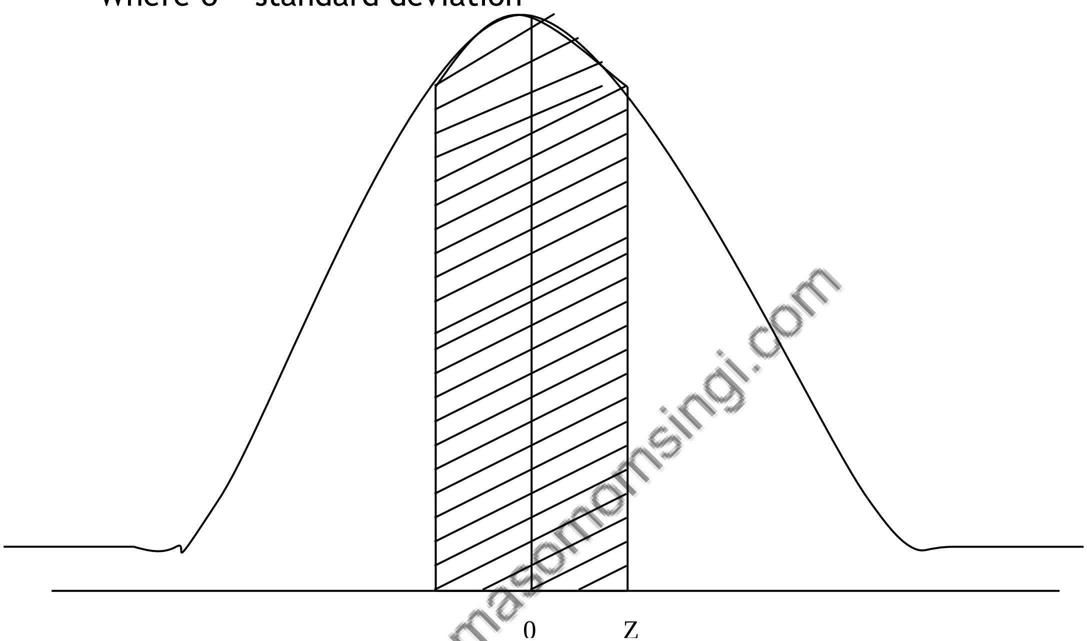
1. The total area under the curve is = 1 which is equivalent to the maximum value of probability



2. The line of symmetry divides the curve into <u>two equal halves</u>
3. The two ends of the normal distribution curve continuously approach the horizontal axis but they <u>never cross it</u>
4. The values of the mean, mode and median are all <u>equal</u>

NB: The above distribution curve is referred to as normal probability distribution curve because if a frequency distribution curve is plotted from measurements of a given sample drawn from a normal population then a graph similar to the normal curve must be obtained.
It should be noted that 68% of any population lies within one standard deviation, $\pm 1\sigma$
95% lies within two standard deviations $\pm 2\sigma$
99% lies within three standard deviations $\pm 3\sigma$

Where σ = standard deviation



STANDARDIZATION OF VARIABLES

Before we use the normal distribution curve to determine probabilities of the continuous variables, we need to standardize the original units of measurement, by using the following formula.

$$Z = \frac{\chi - \mu}{\sigma}$$

Where x = Value to be standardized
Z = Standardization of x
μ = population mean
σ = Standard deviation

Example

A sample of students had a mean age of 35 years with a standard deviation of 5 years. A student was randomly picked from a group of 200 students. Find the probability that the age of the student turned out to be as follows

i. Lying between 35 and 40
ii. Lying between 30 and 40
iii. Lying between 25 and 30
iv. Lying beyond 45 yrs

v. Lying beyond 30 yrs
vi. Lying below 25 years

Solution

(i). The standardized value for 35 years

$$Z = \frac{\chi - \mu}{\sigma} = \frac{35 - 35}{5} = 0$$

The standardized value for 40 years

$$Z = \frac{\chi - \mu}{\sigma} = \frac{40 - 35}{5} = 1$$

$\therefore$ the area between Z = 0 and Z = 1 is 0.3413 (These values are checked from the normal tables see appendix)
The value from standard normal curve tables.
When z = 0, p = 0
And when z = 1, p = 0.3413
Now the area under this curve is the area between z = 1 and z = 0
= 0.3413 – 0 = 0.3413

$\therefore$ the probability age lying between 35 and 40 yrs is 0.3413

(ii). 30 and 40 years

$$Z = \frac{\chi - \mu}{\sigma} = \frac{30 - 35}{5} = \frac{-5}{5} = -1$$

$$Z = \frac{\chi - \mu}{\sigma} = \frac{40 - 35}{5} = 1$$

$\therefore$ the area between Z = -1 and Z = 1 is
= 0.3413 (lying on the positive side of zero) + 0.3413 (lying on the negative side of zero)
P = 0.6826

$\therefore$ the probability age lying between 30 and 40 yrs is 0.6826

(iii). 25 and 30 years

$$Z = \frac{\chi - \mu}{\sigma} = \frac{25 - 35}{5} = \frac{-10}{5} = -2$$

$$Z = \frac{\chi - \mu}{\sigma} = \frac{30 - 35}{5} = -1$$

$\therefore$ the area between Z = -2 and Z = -1

Probability area corresponding to Z = -2
= 0.4772 (the z value to check from the tables is 2)

Probability area corresponding to Z = -1
= 0.3413 (the z value for this case is 1)

$\therefore$ the probability that the age lies between 25 and 30 yrs
= 0.4772 – 0.3413 (The area under this curve)
P = 0.1359

iv). P(beyond 45 years) is determined as follow = P(x > 45)

$$Z = \frac{\chi - \mu}{\sigma} = \frac{45 - 35}{5} = \frac{+10}{5} = +2$$

Probability corresponding to Z = 2 = 0.4772 = probability of between 35 and 45

$\therefore$ P(Age > 45yrs) = 0.5000 – 0.4772
= 0.0228

PRACTICE QUESTION

Question One

The quality controller, Mr. Brooks, at Queensville Engineers has become aware of the need for an acceptance sampling programme to check the quality of bought-in components. This is of particular importance for a problem the company is currently having with batches of pump shafts bought in from a local supplier. Mr. Brooks proposes the following criteria to assess whether or not to accept a large batch of pump shafts from this supplier.

From each batch received take a random sample of 50 shafts, and accept that batch if no more than two defectives are found in the sample.

Mr. Brooks needed to calculate the probability of accepting a batch $P_{a,}$ when the proportion of defectives in the batch, p, is small (under 10%, say)

Required:

a) Explain why the Poisson distribution is appropriate to investigate this situation.
b) Using the Poisson distribution, determine the probability of accepting a batch $P_{a,}$ containing p=2% defectives if the method is used.
Determine $P_{a,}$ for p = 0%, 2%, 5%, 10%, 15%

Question Two

A woven cloth is liable to contain faults and is subjected to an inspection procedure. Any fault has a probability of 0.7 that it will be detected by the procedure, independent of whether any other fault is detected or not.

Required:

a) If a piece of cloth contains three faults, A, B and C,
   i) Calculate the probability that A and C are detected, but that B is undetected;
   ii) Calculate the probability that any two of A, B and C be detected, the other fault being undetected;
   iii) State the relationship between your answers to parts (i) and (ii) and give reasons for this.

b) Suppose now that, in addition to the inspection procedure given above, there is a secondary check which has a probability of 0.6 of detecting each fault missed by the first inspection procedure. This probability of 0.6 applies independently to each and every fault undetected by the first procedure.
   i) Calculate the probability that a piece of cloth with one fault has this fault undetected by both the inspection procedure and the secondary check;
   ii) Calculate the probability that a piece of cloth with two faults has one of these faults detected by either the inspection procedure or the secondary check, and one fault undetected by both;
   iii) Of the faults detected, what proportions are detected by the inspection procedure and what proportion by the secondary check?

Question Three

A company has three production sections S_1, S2 and S_3 which contribute 40%, 35% and 25%, respectively, to total output. The following percentages of faulty units have been observed:

S_1	2%	(0.02)
S_2	3%	(0.03)
S_3	4%	(0.04)

There is a final check before output is dispatched. Calculate the probability that a unit found faulty at this check has come from section 1, S_1

Question Four

Assuming a Binomial Distribution what is the probability of a salesman making 0, 1, 2,3,4,5 or 6 sales in 6 visits if the probability of making a sale on a visit is 0.3?