Unit 4

Probability Distributions

4.0 LEARNING OUTCOMES

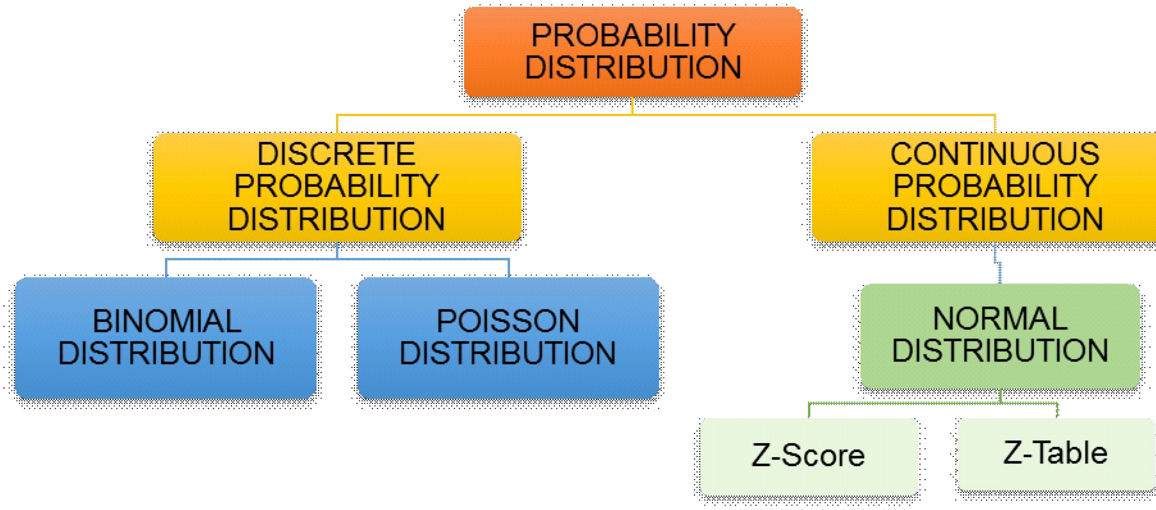
After completion of this unit the students will be able to

- Understand the concept of Random Variables
- Distinguish between discrete and continuous random variable
- Understand and apply the concept of Probability Distribution
- Write probability distribution of discrete random variable
- Calculate the mathematical expectation and variance of a discrete random variable
- Understand and apply the concept of Binomial Distribution
- Calculate the mathematical expectation and variance for a binomial distribution
- Understand and apply the concept of Poisson Distribution
- Calculate the mathematical expectation and variance for a Poisson distribution
- Understand and apply the concept of Normal Distribution
- Calculate the mathematical expectation and variance for a normal distribution
- Calculate Z-Score and Use Z-Table to interpret normal distribution data set

4.0.0 BEFORE YOU START, YOU SHOULD KNOW

1. Random experiment, Sample space, Event associated with a sample space
2. Mutually exclusive and mutually exhaustive events
3. Independent and dependent events
4. Multiplication theorem of Probability
5. Addition theorem of Probability
6. Total Probability
7. Bayes' theorem

4.1 CONCEPT MAP



4.2 INTRODUCTION

Suhani has two black sweaters and a white sweater in her cupboard. She takes out a sweater at random, notes the colour and puts it back in the cupboard. She repeated the process once more before making up her mind.

What shall be the sample space of the situation stated above?

Let us consider the sweaters as B_1, B_2 and W_1

For the selection of two assignments,

the sample space is $S = \{ B_1B_1, B_1 B_2, B_2B_1, B_2B_2, B_1W_1, B_2W_1, W_1B_1, W_1B_2, W_1W_1\}$

Clearly these draws are of a random experiment with random outcomes that cannot be predicted.

Let X represent the number of white sweaters drawn in this situation, in that case what can you say about the value of X?

Here, $X (B_1B_1) = X (B_1 B_2) = X (B_2B_1) = X(B_2B_2) = 0$ as the sample element does not have any white sweater.

Also, $X(B_1W_1) = X(B_2W_1) = X(W_1B_1) = X(W_1B_2) = 1$ as the sample element has one white sweater

And, $X(W_1W_1) = 2$ as the sample element has two white sweaters

$\Rightarrow$ X can take values 0, 1 or 2

Here, X is a function whose domain is the set of possible outcomes (or sample space) of a random experiment. Also, the variable X take any real value, therefore, its co-domain is the set of real numbers

In such a case X is considered as a random variable

Definition: A random variable is a real valued function whose domain is the sample space of a random experiment

Let us consider an experiment of tossing a coin two times in succession.

Clearly the sample space of this experiment is S = {HH, HT, TH, TT}.

If X represents the number of heads obtained in this situation, Then X(HH) = 2

X (HT) = X (TH) = 1

and X (TT) = 0.

let Y represent the number of tails minus the number of heads for each random outcome of the above sample space S

Then Y(TT) = 2

Y(TH) = Y(HT) = 0

And Y(HH) = -2

In this case, Y is a random variable which can take values 2, 0 or - 2

Please note that more than one random variable can be defined on a given sample space. In both the situations above, we shall assume that each random outcome is equally likely to be selected.

Example 1

Rajat is playing a game of rolling a die with his friends. According to the game rules, he will win Rs 5 for rolling an even number and for getting an odd digit on the die, he looses ₹ 2. If X represents the amount of money Rajat wins or loses. Show that X is a random variable and also represent it as a function on the sample space of the game play.

Solution: Sample space of the game play S = {1, 2, 3, 4, 5, 6}

As X represents the amount of money Rajat wins or loses ⇒ X is a function whose values are defined on the basis of random outcomes, therefore it is a random variable.

X (2) = X (4) = X (6) = 1× 5 = Rs 5 as Rajat wins Rs 5 when he rolls an even digit on the die, X(1) = X(3) = X(5) = 1× (-2) = -2 as Rajat loses Rs 2 on rolling an odd digit on the die

Thus, for each element of the sample space, X takes a unique value, hence, X is a function on the sample space whose range is {+5, -2}

4.2.1 Discrete and Continous Random Variables

Recall that a variable is a quantity that keeps varying.

Let us consider toss of a fair coin and let X be the random variable defined as

$$X = \begin{cases} 0, \textit{if coin toss result in head} \\ 1, \textit{if coin toss result in tail} \end{cases}$$

Here, the random variable is taking two distinct and countable (measurable) values.

Hence X in this case has distinct and countable outcomes with no number in between these values, therefore it is a *discrete random variable.*

A wrist watch with only hour and minutes display shows time as 12:00, then 12:01, 12:02, and so on and there is no time shown in between. In this case the random time change is distinct and countable. Therefore the change in time in this case is discrete random variable

Each possible value of the discrete random variable can be associated with a non-zero probability.

Whereas a wrist watch displaying the seconds count as well shows time 22:31:25 pm and 22:32:17 pm and the elapsed time in between as well. A random variable whose value is obtained by measuring and it takes many values between two values, is called a *continuous random variable*.

In other words, a continuous random variable is a random variable with a set of possible infinite and uncountable values (known as the range).

4.2.2 Probability Distribution of Discrete Random Variable

Let us consider another random variable X defined as sum of digits on rolling of two dice.

The following grid shows the sample space of this random experiment:

+	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Clearly, X will take values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12 which are distinct and countable, hence X is a discrete random variable in this case.

Now let us find the probability for each random outcome

X →	2	3	4	5	6	7	8	9	10	11	12
P(X) →	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Table (i)

Observe that in table (i); for all possible values of the discrete random variable X, all elements of the sample space are covered.

This table of possible outcomes and their respective probability is called *Probability distribution table* for the given random variable X. A probability distribution table links every possible outcome of the random experiment with the probability of the event to occur.

In a probability distribution table, the sum of all the probabilities is one. (Refer table (i))

Example 2

A coin is tossed thrice and outcomes are recorded. Prepare the probability distribution table for the number of heads.

Solution: Let a random variable X denote the number of heads in three throws of a die

Here the sample space S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

In a single toss of a fair coin: P (head) = $\frac{1}{2}$, and P (tail) = $\frac{1}{2}$

Which means that X can attain values 0 for no heads, 1 for one head and two tails, 2 for two heads and one tail or 3 for three heads

$$\Rightarrow X = 0, 1, 2 \text{ and } 3$$

The probability of no heads i.e. P (TTT) which can also be written as

$$P(X = 0) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

(Recall multiplication theorem of probability from class XI)

probability of obtaining one head and two tails i.e. P(HTT, THT, TTH) is denoted by

$$P(X = 1) = 3 \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{3}{8}$$

probability of obtaining two head and one tail i.e. P(HHT, HTH, THH) is denoted by

$$P(X = 2) = 3 \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{3}{8}$$

And, probability of obtaining three heads i.e. P (HHH) is written as

$$P(X = 3) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$$

> In a probability distribution the sum of all the probabilities is always one

Therefore, the probability distribution table is:

$x_i \rightarrow$	0	1	2	3
$P(x_i) = p_i \rightarrow$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

4.3 MATHEMATICAL EXPECTATION OF DISCRETE PROBABILITY DISTRIBUTION

Recall that mean is a measure of central tendency as it locates a rough estimation of a middle or average value of a random variable in an experiment.

Definition: In an experiment, for a given random variable X whose possible finite values x_1, x_2, x_3, ..., x_n occur with probabilities p_1, p_2, p_3, …,p_n respectively such that $\sum p_i = 1$

Then the mathematical expectation is the weighted average of the possible values of X given by:

$$E(X) = x_1p_1 + x_2p_2 + x_3p_3 + \cdots + x_np_n = \sum_{i=1}^{n}(x_ip_i)$$

In a nutshell, the *mathematical expectation*, also known as *expected value* for a random variable X is the summation of product of all possible values for the given random variable X and their respective probabilities.

Example 3

A coin is tossed twice and outcomes are recorded. Prepare the probability distribution table for random variable X which represents the number of heads in the experiment. Also calculate the mathematical expectation of X.

Solution: Let a random variable X denote the number of heads in two throws of a die

$\Rightarrow$ the sample space S = { HH, HT, TH,TT}

Clearly X = 0, 1 and 2

The probability of occurrence of a head = probability of occurrence of a tail = $\frac{1}{2}$

The probability distribution table-

$x_i \rightarrow$	0	1	2
Sample event	TT	HT, TH	HH
$P(x_i) = p_i \rightarrow$	$\frac{1}{4}$	$2\times\frac{1}{4}=\frac{1}{2}$	$\frac{1}{4}$
$x_i p_i \rightarrow$	$0\times\frac{1}{4}=0$	$1\times\frac{1}{2}=\frac{1}{2}$	$2\times\frac{1}{4}=\frac{1}{2}$

Note that $\sum p_i = 1$

Therefore, $E(X) = \sum_{i=1}^{n}(x_i p_i) = 0 + \frac{1}{2} + \frac{1}{2} = 1$

Example 4

In a manufacturing unit inspection, from a lot of 20 baskets which include 6 defectives, a sample of 2 baskets is drawn at random without replacement. Prepare the probability distribution of the number of defective baskets. Also calculate E(X) for the random variable X.

Solution:

As X denotes the number of defective baskets in a draw of 2 *without replacement*

$\Rightarrow X = 0, 1$ and 2

Therefore, in a draw of two baskets;

X	0	1	2
x_i	No defective baskets	One defective basket	Two defective baskets
$P(x_i) = p_i$	$\frac{14}{20}\times\frac{13}{19}=\frac{182}{380}$	$2\times\frac{14}{20}\times\frac{6}{19}=\frac{168}{380}$	$\frac{6}{20}\times\frac{5}{19}=\frac{30}{380}$
$x_i p_i$	$0\times\frac{182}{380}=0$	$1\times\frac{168}{380}=\frac{168}{380}$	$2\times\frac{30}{380}=\frac{60}{380}$

Note that $\sum p_i = 1$

Therefore, $E(X) = \sum_{i=1}^{n}(x_i p_i) = 0 + \frac{168}{380} + \frac{60}{380} = \frac{228}{380} = 0.6$

E(X) of a random variable X, is the theoretical mean of X. It is not based on sample data but on the distribution of it

So, the mean expectation value is a parameter and not a statistic.

Sometimes it is also represented by use of Greek letter ***mu*** (μ) as well.

Also random variables with different probability distributions can have equal means. Let us take an example to study this statement in detail

Have a look at probability distribution of two different random variables X and Y as given below:

X	1	2	3	4
$P(x_i) = p_i$	$\frac{1}{7}$	$\frac{3}{7}$	$\frac{2}{7}$	$\frac{1}{7}$
$x_i p_i$	$1\times\frac{1}{7}=\frac{1}{7}$	$2\times\frac{3}{7}=\frac{6}{7}$	$3\times\frac{2}{7}=\frac{6}{7}$	$4\times\frac{1}{7}=\frac{4}{7}$

Here, $E(X) = \sum_{i=1}^{n}(x_i p_i) = \frac{1}{7}+\frac{6}{7}+\frac{6}{7}+\frac{4}{7}=\frac{17}{7}=2.71$

Y	-1	0	4	5
$P(y_i) = p_i$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{3}{7}$	$\frac{1}{7}$
$y_i p_i$	$-1\times\frac{1}{7}=\frac{-1}{7}$	$0\times\frac{2}{7}=0$	$4\times\frac{3}{7}=\frac{12}{7}$	$5\times\frac{1}{7}=\frac{5}{7}$

Here, $E(Y) = \sum_{i=1}^{n}(x_i p_i) = \frac{-1}{7}+0+\frac{12}{7}+\frac{5}{7}=\frac{16}{7}=2.29$

Clearly the random variables X and Y with different probability distributions can have equal means. In such cases, we need a technique to check variability and extent to which the values of random variable are spread out.

4.4 VARIANCE OF DISCRETE PROBABILITY DISTRIBUTION

While the mean is a central tendency; known as the average of a group of data, the variance measures the average degree to which each number in the data is different from the mean value. The extent or scope of the variance correlates to the size of the overall range of the given sample

> Variance enables us to study the variability of random variable from the mean expectation
>
> When there is a narrower range among the sample elements in a given sample space; that means that the value of the random variable is close to mean expectation and hence the variance is less
>
> And, when there is wide range among the sample elements, it means that the value of the random variable is far from the mean expectation and thus the variance is high.

Basically, the variance measures the average degree to which each sample element differs from the mean of the sample space

In a probability distribution of a discrete random variable X, the variance denoted by *Var (X)*; is the summation of the product of the squared deviations of x_i from the mean E(X) and the corresponding probabilities p_i.

Definition: Let X be a discrete random variable whose possible finite values $x_1, x_2, x_3, \ldots, x_n$ occur with probabilities $p_1, p_2, p_3, \ldots, p_n$ respectively.

$$\sigma_x^{\ 2} = \text{Var}(X) = \sum_{i=1}^{n} x_i^{\ 2} p_i - \left(\sum_{i=1}^{n} x_i p_i\right)$$

In other words, Var (X) = E(X^2) – [E(X)]^2, where E (X^2) = $\sum_{i=1}^{n} x_i^{\ 2} p_i$ –

And, the standard deviation denoted by σ_x is given by:

$$\sigma_x = \sqrt{Var(X)}$$

Example 5

A class XII has 20 students whose marks (out of 30) are 14, 17, 25, 14, 21, 17, 17, 19, 18, 26, 18, 17, 17, 26, 19, 21, 21, 25, 14 and 19 years. If random variable X denotes the marks of a selected student given that the probability of each student to be selected is equally likely.

a) Prepare the probability distribution of the random variable X.

b) Find mean, variance and standard deviation of X.

Solution: Based on the given data, let us prepare a table

Marks	14	17	18	19	21	25	26
frequency	3	5	2	3	3	2	2

As probability of a selection of a student is equally likely

That means P (a student to be selected) = $\frac{1}{20}$

Therefore, the probability distribution is:

Marks = x_i	14	17	18	19	21	25	26
Frequency = f_i	3	5	2	3	3	2	2
P(x_i) = p_i	$\frac{3}{20}$	$\frac{5}{20}$	$\frac{2}{20}$	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{2}{20}$	$\frac{2}{20}$
$x_i p_i$	$\frac{42}{20}$	$\frac{85}{20}$	$\frac{36}{20}$	$\frac{57}{20}$	$\frac{63}{20}$	$\frac{50}{20}$	$\frac{52}{20}$
$x_i^2 p_i$	$\frac{147}{5}$	$\frac{289}{4}$	$\frac{162}{5}$	$\frac{1083}{20}$	$\frac{1323}{20}$	$\frac{125}{2}$	$\frac{338}{5}$

Here, E(X) = $\sum_{i=0}^{n}(x_i p_i) = \frac{42}{20} + \frac{85}{20} + \frac{36}{20} + \frac{57}{20} + \frac{63}{20} + \frac{50}{20} + \frac{52}{20} = \frac{385}{20} = 19.25$

And, $\sum_{i=0}^{n} x_i^{\ 2} p_i = \frac{147}{5} + \frac{289}{4} + \frac{162}{5} + \frac{1083}{20} + \frac{1323}{20} + \frac{125}{2} + \frac{338}{5} = \frac{7689}{20}$

$$\Rightarrow \text{Var}(X) = \sum_{i=1}^{n} x_i^2 p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = \frac{7689}{20} - \left(\frac{385}{20}\right)^2$$

$$= \frac{7689}{20} - \frac{148225}{400}$$

$$= \frac{153780 - 148225}{400}$$

$$= \frac{5555}{400} = 13.9$$

And standard deviation, $\sigma_x = \sqrt{Var(X)} = \sqrt{13.9} = 3.7$

Example 6

Let X denote the number of hours a person watches television during a randomly selected day. The probability that X can take the values x_i, has the following form, where k is some unknown constant.

$$P(X = x_i) = \begin{cases} 0.2, & \text{if } x_i = 0 \\ kx_i, & \text{if } x_i = 1 \text{ or } 2 \\ k(5 - x_i), & \text{if } x_i = 3 \\ 0 & \text{otherwise} \end{cases}$$

a) Find the value of k.

b) What is the probability that the person watches two hours of television on a selected day?

c) What is the probability that the person watches at least two hours of television on a selected day?

d) What is the probability that the person watches at most 2 hours of television on a selected day?

e) Calculate mathematical expectation

f) Find variance and standard deviation of random variable X

Solution:

x_i	0	1	2	3
$P(x_i) = p_i$	0.2	k	2k	2k

a) As $\sum p_i = 1$

$\Rightarrow 0.2 + k + 2k + 2k = 1$

$\Rightarrow 5k = 0.8$

$\Rightarrow k = \frac{4}{25}$

b) Probability that the person watches two hours of television

$= P(x_i = 2)$

$= 2k = 2 \times \frac{4}{25} = \frac{8}{25}$

c) Probability that the person will watch at least two hours of television

$= P\ (x_i = 2, 3)$

$= 2k + 2k = 4k = 4 \times \frac{4}{25} = \frac{16}{25}$

d) Probability that the person will watch at most two hours of television

$= P\ (x_i = 0, 1, 2)$

$= 0.2 + k + 2k = 0.2 + 3k = 0.2 + 3 \times \frac{4}{25} = \frac{17}{25}$

e)

x_i	0	1	2	3
$P(x_i) = p_i$	$0.2 = \frac{1}{5}$	$\frac{4}{25}$	$\frac{8}{25}$	$\frac{8}{25}$
$x_i p_i$	0	$\frac{4}{25}$	$\frac{16}{25}$	$\frac{24}{25}$
$x_i^2 p_i$	0	$\frac{4}{25}$	$\frac{32}{25}$	$\frac{72}{25}$

$E(X) = \sum_{i=1}^{n}(x_i p_i) = 0 + \frac{4}{25} + \frac{16}{25} + \frac{24}{25} = \frac{44}{25} = 1.76$

f) $\sum_{i=1}^{n} x_i^2 p_i = 0 + \frac{4}{25} + \frac{32}{25} + \frac{72}{25} = \frac{108}{25}$

$$\Rightarrow \text{Var}(X) = \sum_{i=1}^{n} x_i^2 p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = \frac{108}{25} - \left(\frac{44}{25}\right)^2 = \frac{108}{25} - \frac{1936}{625}$$

$$= \frac{2700 - 1936}{625} = \frac{764}{625} = 1.22$$

And standard deviation, $\sigma_x = \sqrt{\text{Var}(X)} = \sqrt{1.22} = 1.1$

4.5 BINOMIAL DISTRIBUTION

When you toss a coin, it either shows a 'head' or a 'tail'. When you are asked to calculate 3 + 4, the answer is either 'correct' or 'incorrect'. In such similar experiments the likely outcome is either a 'success' or a 'failure'.

Note that in each trial, the probability of success or failure remains constant, the outcome of any trial is independent of the outcome of any other trial.

In the case of discrete random variable X denoting a prime number on throw of a die, we can say that numbers 2, 3 and 5 will be considered as 'success', while 1, 4 and 6 will be counted as 'failure' in the experiment.

Each time you roll a die or perform any experiment in probability, it is called a *trial*. If in an experiment, a die is rolled thrice, then the number of trials is counted as 3, each trial having exactly two outcomes, namely, *success or failure*.

Independent trials which have only two outcomes usually referred as 'success' and 'failure' are called Bernoulli trials. Here, the probability of success and failure remains same.

Definition: In a random experiment, a collection of trials is called *Bernoulli trials*, if:

a) The number of trials is finite.

b) The trials are independent by nature.

c) Each trial has exactly two outcomes defined as success and failure.

d) The probability of success remains the same in each trial.

Recall example 4 where the random variable X denotes the number of defective baskets in a draw of two baskets *without replacement*

What would happen if 5 baskets are drawn *with replacement*?

$$\Rightarrow X = 0, 1, 2, 3, 4 \text{ and } 5$$

In this case, probability of drawing a defective basket will be considered a success, usually denoted by p and probability of drawing a non-defective basket will be a failure, denoted by q

Also, the draw of a basket will be called a trial and since we are drawing 5 baskets; the number of trials is 5

Do you think that these trials qualify as *Bernoulli trials*?

here $p = \frac{6}{20}$ and $q = 1 - p = 1 - \frac{6}{20} = \frac{14}{20}$, Since probability of success remains same in all the trials, hence we can say these are binomial trials.

When the drawing is done *without replacement*, the probability of success (i.e., drawing a defective basket) in first trial is $\frac{6}{20}$, in 2nd trial it will be $\frac{5}{19}$ and so on.

Clearly, the probability of success is not same for all trials, hence the trials in example 4 are *not* Bernoulli trials.

Probability of 'r' successes in 'n' Bernoulli trials is given by:

$$P(\text{'r' successes}) = C_r^n\, p^r q^{n-r} = \frac{n!}{r!\,(n-r)!}\, p^r q^{n-r}$$

Where n = number of trials

r = number of successful trials = 0, 1, 2, 3, …, n

p = probability of a success in a trial

q = probability of a failure in a trial

And, $p + q = 1$

Clearly, P ('r' successes), is the (r + 1)th term in the binomial expansion of $(q + p)^n$.

The probability distribution of number of successes for a random variable X can be written as:

X = r_i	0	1	2	3	…	r	…	n
P(r_i) = p_i	$C_0^n\, p^0 q^{n-0}$	$C_1^n\, p^1 q^{n-1}$	$C_2^n\, p^2 q^{n-2}$	$C_3^n\, p^3 q^{n-3}$	…	$C_r^n\, p^r q^{n-r}$	…	$C_n^n\, p^n q^{n-n}$

This probability distribution is called *Binomial distribution* with parameters n and p.

The binomial distribution with n Bernoulli trials and success p is also denoted by B(n, p)

Example 7:

Prepare the Binomial distribution B (4, $\frac{2}{3}$)

Solution: Here total number of trials = n = 4 and p = $\frac{2}{3}$

As $p + q = 1 \Rightarrow q = 1 - \frac{2}{3} = \frac{1}{3}$

Now, number of successes = r = 0, 1, 2, 3 or 4

The binomial distribution can be given as:

$X = r_i$	0	1	2	3	4
$P(r_i) = p_i$	$C_0^4\left(\frac{2}{3}\right)^0\left(\frac{1}{3}\right)^{4-0}$ $= \frac{1}{81}$	$C_1^4\left(\frac{2}{3}\right)^1\left(\frac{1}{3}\right)^{4-1}$ $= \frac{8}{81}$	$C_2^4\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^{4-2}$ $= \frac{24}{81}$	$C_3^4\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^{4-3}$ $= \frac{32}{81}$	$C_4^4\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^{4-4}$ $= \frac{16}{81}$

Now let us calculate the mean expectation, Variance and Standard Deviation for example 7

Recall that $E(X) = \sum_{i=1}^{n}(x_i p_i) = 0\times\frac{1}{81} + 1\times\frac{8}{81} + 2\times\frac{24}{81} + 3\times\frac{32}{81} + 4\times\frac{16}{81} = \frac{216}{81} = 2.67$

Also see that $np = 4\times\frac{2}{3} = \frac{8}{3} = 2.67$

$$\Rightarrow \text{Var}(X) = \sum_{i=1}^{n} x_i^2 p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2 = \left(0 + \frac{8}{81} + \frac{96}{81} + \frac{288}{81} + \frac{256}{81}\right) - \left(\frac{216}{81}\right)^2 = \frac{5832}{6561} = 0.89$$

Also see that $npq = 4\times\frac{2}{3}\times\frac{1}{3} = \frac{8}{9} = 0.89$

And Standard deviation = $\sigma_x = \sqrt{Var(X)} = \sqrt{0.9} = \sqrt{npq} = 0.95$

In a binomial distribution having 'n' number of Bernoulli trials where p denotes the probability of success and q denotes the probability of failure then,

Mean = np

Variance = npq

Standard Deviation = $\sqrt{npq}$

Example 8

If a fair coin is tossed 9 times, find the probability of

a) exactly five tails

b) At least five tails

c) At most five tails

Solution:

Repeated tosses of a fair coin qualify as Bernoulli's trails

Let X denote the number of tails in an experiment of 9 such trials and hence is the binomial distribution

Here, n = 9, $p = \frac{1}{2}$ and $q = 1 - p = \frac{1}{2}$

As P ('r' successes) = $C_r^n p^r q^{n-r} = \frac{n!}{r!(n-r)!} p^r q^{n-r}$

a) Probability of exact 5 successes in 9 trials = P(X = 5) = $C_5^9 p^5 q^{9-5}$

$$= \frac{9!}{5!\,(9-5)!} \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^{9-5}$$

$$= \frac{9!}{5!\,(9-5)!} \cdot \left(\frac{1}{2}\right)^9$$

$$= \frac{63}{256}$$

b) Probability of at least 5 successes in 9 trials = P (X ≥ 5)

$$= \left(\frac{1}{2}\right)^9 [C_5^9 + C_6^9 + C_7^9 + C_8^9 + C_9^9] = \frac{256}{512}$$

c) Probability of at most 5 successes in 9 trials = P (X ≤ 5) = 1 - P(X > 5)

$$= 1 - \left(\frac{1}{2}\right)^9 [C_6^9 + C_7^9 + C_8^9 + C_9^9]$$

$$= 1 - \frac{131}{512} = \frac{382}{512}$$

Example 9

In a manufacturing unit inspection, from a lot of 20 baskets which include 6 defectives, a sample of 2 baskets is drawn at random with replacement. Prepare the binomial distribution of the number of defective baskets. Also find E(X) and Var(X) for the random variable X

Solution: Here, X denotes the number of defective baskets in a draw of 2 baskets with replacement

Clearly, the trials are Bernoulli trials

And X = 0, 1 and 2

Also number of trials = n = 2

If drawing a defective basket is considered a success,

then $p = \frac{6}{20} = \frac{3}{10} \Rightarrow q = \frac{7}{10}$

X	**0**	**1**	**2**
r_i	No defective baskets	One defective basket	Two defective baskets
$P(r_i) = p_i$	$C_0^2 \left(\frac{3}{10}\right)^0 \left(\frac{7}{10}\right)^{2-0} = \frac{49}{100}$	$C_1^2 \left(\frac{3}{10}\right)^1 \left(\frac{7}{10}\right)^{2-1} = \frac{42}{100}$	$C_2^2 \left(\frac{3}{10}\right)^2 \left(\frac{7}{10}\right)^{2-2} = \frac{9}{100}$

E(X) = np = $4 \times \frac{3}{10} = 1.2$ and Var(X) = npq = $4 \times \frac{3}{10} \times \frac{7}{10} = 0.84$

Example 10

The probability that Rohit will hit a shooting target is $\frac{2}{3}$. While preparing for an international shooting competition, Rohit aims to achieve the probability of hitting the target at least once to be 0.99. What is the minimum number of chances must he shoot to attain this probability?

Solution: Let the number of chances Rohit shoots the target be n

Here, the trials are Bernoulli with p be the probability of success to hit the target = $\frac{2}{3}$ and q be the probability of failure to hit the target be $1 - p = \frac{1}{3}$

Then P(r number of successes) = $C_r^n \left(\frac{2}{3}\right)^r \left(\frac{1}{3}\right)^{n-r} = \frac{n!}{r!(n-r)!} \left(\frac{2}{3}\right)^r \left(\frac{1}{3}\right)^{n-r}$

As Rohit wants to hit the target at least once with the probability of 0.99

$\Rightarrow$ P(r = 1,2,3,…) $\geq$ 0.99

$\Rightarrow 1 - P(r = 0) \geq 0.99$

$\Rightarrow 1 - \frac{n!}{0!(n-0)!} \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^{n-0} \geq 0.99$

$\Rightarrow 1 - 0.99 \geq \left(\frac{1}{3}\right)^n$

$\Rightarrow 0.01 \geq \left(\frac{1}{3}\right)^n$

$\Rightarrow 100 \leq (3)^n$

As $3^5 \geq 100 \Rightarrow n \geq 5$

$\Rightarrow$ Rohit should hit the target at least 5 times to achieve his target.

Example 11

Sonal and Anannya are playing a game by throwing a die alternatively till one of them gets a '1' and wins the game. Find their respective probabilities of winning, if Sonal starts first

Solution: Clearly the trials are Bernoulli's with $n \longrightarrow \infty$

Getting a 1 on a single throw of the die is considered a success

$\Rightarrow p = \frac{1}{6}$ and $q = 1 - p = \frac{5}{6}$

Sonal starts the game by throwing the die first

P(Sonal to win in the first throw) = $\frac{1}{6}$

When will Sonal get a chance to win next?

Sonal will get to try winning in third throw; when Sonal fails in first throw and then Anannya fails to win in second throw

P (Sonal to win in third throw) = $\left(\frac{5}{6}\right)^2 \times \frac{1}{6}$

Next time Sonal will get to win is fifth throw

P (Sonal to win in fifth throw) = $\left(\frac{5}{6}\right)^4 \times \frac{1}{6}$ and so on

Hence, P(Sonal will win) = P(in first throw) + P(in third throw) + P(in fifth throw) + …

$= \frac{1}{6} + \left(\frac{5}{6}\right)^2 \times \frac{1}{6} + \left(\frac{5}{6}\right)^4 \times \frac{1}{6} + \cdots \infty$

It is an infinite geometric series with $a = \frac{1}{6}$, $r = \left(\frac{5}{6}\right)^2 < 1$

As $S_\infty = \frac{a}{1-r}$, $r < 1 \Rightarrow S = \frac{\frac{1}{6}}{1-\left(\frac{5}{6}\right)^2} = \frac{6}{11}$

And, P(Anannya to win) = 1 – P(Sonal to win) = $1 - \frac{6}{11} = \frac{5}{11}$

Example 12

A die is thrown again and again until three 5*s*′ are obtained. Find the probability of obtaining the third 5 in the seventh throw of the die

Solution: Clearly the trials are Bernoulli with n = 6

P (a 5 on a single throw of die) = $\frac{1}{6} \Rightarrow p = \frac{1}{6}$ and $q = 1 - p = \frac{5}{6}$

For finding the probability of third six in the seventh throw of the die, we know that there must have been two 5*s*′ on previous six throws

$\Rightarrow$ P (third 5 on seventh throw of die) = P (two 5*s* on six throws) $\times$P (a 5 on the next single throw of die)

$= C_2^6 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{6-2} \times \frac{1}{6} = 15 \times \frac{1}{36} \times \frac{625}{1296} = \frac{3125}{15552}$

4.6 POISSON DISTRIBUTION

Let us consider the car sales of a car dealer showroom X in a city, on a given day.

Do you think that the number of cars sales on a given day will make for a random variable?

Assuming that each car sale is an independent event, meaning that sale of one car sale gives no information about when the next sale will happen. And the probability of one car sale in a given length of time, does not change over time.

Source courtesy:
http://clipart-library.com/free-car-images.html

Theoretically, the rate at which the car sales are occurring is not changing through time.

Therefore, we can conclude that the events defined as car sales in such a case are occurring randomly and independently.

Based on these conditions, a random variable X, representing the number of events in a given length of time has a *Poisson distribution.*

A discrete probability distribution that expresses the probability of a given number of events occurring over a fixed period of time or space is called a Poisson distribution if:

1. The events occur with a known constant mean rate
2. The events are independent of the time from the occurrence of the previous event.

3. The rate of occurrence of events is constant and not based on time
4. The probability of an event is proportional to the length of the period of time

The Poisson distribution can also be used for the number of events in other specified intervals such as distance, area or volume.

DEFINITION: Let X be the discrete random variable which represents the number occurrence of events over a period of time.

If X follows the Poisson distribution, then the probability of occurrence of '*k*' number of events over a period of time is given by:

$$P\,(X = k) = f(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Where e is Euler's number (e = 2.71828…)

'*k*' is the number of occurrences of the event such that k = 0, 1,2, …

And = E(X) = Var(X), is a positive real number

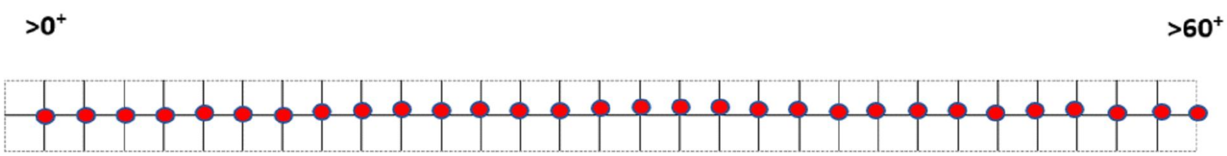
With existence conditions:

1. $\sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = 1$
2. for k = 0, 1, 2, …

Formula courtesy: https://en.wikipedia.org/wiki/Poisson_distribution

A restaurant is doing booming business. It was recorded that during their peak business time, an average of 30 customers per hour arrive at the restaurant. Can we develop a Poisson probability distribution model for the arrivals of customers, if 30 customers arrive in an interval of 1 hour on an average?

You might say that arrival average is 1 customer every 2 minutes



But the thing to remember here is that arrival time of each customer is random and hence this approach is inappropriate

Let us try another approach and divide each one-minute interval along an interval of an hour so that each customer arrival is equally-likely

During each minute, let us consider one customer's arrives in the middle of that time interval. As probability for a customer to arrive is $\frac{1}{2}$ ⇒ this is going to be a binomial distribution B $(60,\frac{1}{2})$

Thus the process will average at $E(X) = np = 60 \times \frac{1}{2} =$ 30 arrivals during an hour

But then again, we cannot assume that the customers are arriving at uniform pace and at regular time intervals

What if we divide the time interval in seconds and consider that probability for a customer to arrive is not equally- likely but biased at $\frac{1}{120}$. In such a case, the binomial distribution will be B $(3600, \frac{1}{120})$

And the process will average at $E(X) = np = 3600 \times \frac{1}{120}$ = 30 arrivals during an hour

To summarize this process, we can say that:

a) as $n \to \infty$, the time intervals getting larger and larger, p is smaller than before while E(X) = np is kept constant at 30 customers per hour

b) In the limit as $n \longrightarrow \infty$, the number of customers arriving during an hour is $\sim$ Poisson (30)

c) With the width $w = \frac{1}{3600}$ of an hour with arrival rate $\lambda = 30$ customer per hour:

   i) Probability of an arrival during the interval = $\lambda w = 30 \times \frac{1}{3600} = \frac{1}{120}$

   ii) Probability of more than one arrival during a time interval is 0

   iii) Probability of an arrival during a time interval is independent of the previous arrivals

One of the famous historical study shows an application of Poisson distribution, to estimate the numbers of Prussian cavalry soldiers killed due to horse-kicks in a year!

Picture credits: https://mindyourdecisions.com/blog/2013/06/21/what-do-deaths-from-horse-kicks-have-to-do-with-statistics/

Example 13

As the story goes, the Prussian soldiers monitored 10 cavalry corps over a period of 20 years. The annual number of recorded deaths due to horse-kick '*k*' observations is as shown in the table:

k	0	1	2	3	4	Total
Number of deaths	109	65	22	3	1	200

Does this data provide adequate description of Poisson distribution?

Solution:

Here $E(X) = \frac{0 \times 109 + 1 \times 65 + 2 \times 22 + 3 \times 3 + 4 \times 1}{200} = \frac{122}{200} = 0.61 = \lambda$

As k = 0, 1, 2, 3, 4

By Poisson distribution formula: $P(X=k) = \frac{0.61^k e^{-0.61}}{k!}$

Then Poisson model progression will be

$P(k = 0) = \frac{0.61^0 e^{-0.61}}{0!} = 0.54$

$P(k = 1) = \frac{0.61^1 e^{-0.61}}{1!} = 0.33$

$P(k = 2) = \frac{0.61^2 e^{-0.61}}{2!} = 0.1$

$P(k = 3) = \frac{0.61^3 e^{-0.61}}{3!} = 0.02$

$P(k = 4) = \frac{0.61^4 e^{-0.61}}{4!} = 0.003$

Now Poisson prediction will be: 200 × P(k)

k	0	1	2	3	4
Number of deaths	109	65	22	3	1
Poisson predicted deaths	200×0.54 =108	200×0.33 =66	200×0.1 =20	200×0.02 =4	200×0.003 =0.6~1

Yes, the Poisson predictions are adequate for the given data

Example 14

A traffic engineer records the number of bicycle riders that use a particular cycle track. He records that an average of 3.2 bicycle riders use the cycle track every hour. Given that the number of bicycles that use the cycle track follow a Poisson distribution, what is the probability that:

a) 2 or less bicycle riders will use the cycle track within an hour?

b) 3 or more bicycle riders will approach the intersection within an hour?

Also write the mean expectation and variance for the random variable X

Solution:

For this problem, $E(X) = Var(X) = \lambda = 3.2$

a) The goal is to find $P(X \leq 2)$

As $P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}$

$\Rightarrow P(X = 0) = \frac{3.2^0 e^{-3.2}}{0!} = \frac{1}{e^{3.2}} = 0.041$

$P(X = 1) = \frac{3.2^1 e^{-3.2}}{1!} = \frac{3.2}{e^{3.2}} = 0.13$

$P(X = 2) = \frac{3.2^2 e^{-3.2}}{2!} = \frac{5.12}{e^{3.2}} = 0.21$

Therefore, $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.041+0.13+0.21 = 0.381$

b) The goal is to find P(X $\geq$ 3)

The probability that there are 3 or more bicycle riders using the track within an hour has no upper limit on the value of '*k*', which means that this probability cannot be calculated directly

But, using the rule of complement we can say that

P(X $\geq$ 3) = 1 - P(X $\leq$ 2) = 1 – 0.381 = 0.619

In the given Poisson distribution, E(X) = Var(X) = λ = 3.2

For the calculation of Euler's number: http://eguruchela.com/math/calculator/e-power-x

Example 15

A particular river near a small-town floods and overflows twice in every 10-years on an average. Assuming that the Poisson distribution is appropriate, what is the mean expectation. Also calculate the probability of 3 or less overflow floods in a 10-year interval.

Solution: As the average event of flood overflow, in every 50-years is two

$\Rightarrow$In the given Poisson distribution, λ = 2

The goal is to find P(X $\leq$ 3)

As $P\ (X = k\ overflow\ floods) = \frac{\lambda^k e^{-\lambda}}{k!}$

$\Rightarrow$ P (X = 0) = $\frac{2^0 e^{-2}}{0!} = \frac{1}{e^2} =$ 0.14

P (X = 1) = $\frac{2^1 e^{-2}}{1!} = \frac{2}{e^2} =$ 0.27

P (X = 2) = $\frac{2^2 e^{-2}}{2!} = \frac{2}{e^2} =$0.27

P (X = 3) = $\frac{2^3 e^{-2}}{3!} = \frac{4}{3e^2} =$0.18

Therefore, P(X $\leq$ 3) = P(X = 0) + P(X = 1) + P(X = 2) +P(X = 3)

= 0.14+0.27+0.27+0.18 = 0.86

Example 16:

For a Poisson distribution model, if arrival rate of passengers at an airport is recorded as 30 per hour on a given day. Find:

a) The expected number of arrivals in the first 10 minutes of an hour

b) The probability of exactly 4 arrivals in the first 10 minutes of an hour

c) The probability of 4 or fewer arrivals in the first 10 minutes of an hour

d) The probability of 10 or more arrivals in an hour given that there are 8 arrivals in the first 10 minutes of that hour

Picture courtesy: https://englishlive.ef.com/blog/career-english/travel-english-key-words-for-the-airport/

Solution:

a) As 10 minutes = $\frac{1}{6}$ th of an hour

$\Rightarrow$ In the given Poisson distribution, X is defined as number of arrivals in the first 10 minutes or $w = \frac{1}{6}$ th hour is the width of time interval

Here, λ = 30 as number of arrivals is 30 per hour

Therefore, E(X) = $\lambda w = 30 \times \frac{1}{6} = \mathbf{5}$ for the first 10-minute of the hour

and P(X = k) = $\frac{5^k e^{-5}}{k!}$, where k = 0, 1, 2, 3, …

b) Probability of exactly 4 arrivals in the first 10 minutes of an hour = P(k = 4) = $\frac{5^4 e^{-5}}{4!}$ = 0.176

c) The probability of 4 or fewer arrivals in the first 10 minutes of an hour = P(k $\leq$ 4) = P(k = 0)+P(k = 1)+ P(k = 2) + P(k = 3) +P(k= 4)

$= \frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!} + \frac{5^2 e^{-5}}{2!} + \frac{5^3 e^{-5}}{3!} + \frac{5^4 e^{-5}}{4!}$

= 0.007+0.03+0.08+0.14+0.18 $\approx$ 0.44

d) We are given that there have been 8 arrivals in the first 10 minutes (=$\frac{1}{6}th$ hour)

And we need to find probability of 10 or more arrivals in an hour
That means, we need to find probability of 10 – 8 = 2 arrivals in 60 – 10 = last 50 minutes (=$\frac{5}{6}th$ hour)

Therefore, E(X) = $\lambda w = 30 \times \frac{5}{6} = \mathbf{25}$ for the last 50-minute of the hour

And, in this case P(X = k) = $\frac{25^k e^{-25}}{k!}$, where k = 0, 1, 2, 3, …

$\Rightarrow$ probability of 10 or more arrivals in an hour given that there are 8 arrivals in the first 10 minutes of that hour = P(k = 2, 3, 4, 5, …∞)

= 1 – P (k = 0, 1)

$= 1 - (\frac{25^0 e^{-25}}{0!} + \frac{25^1 e^{-25}}{1!})$

4.7 NORMAL DISTRIBUTION

In this module, the distributions discussed up till now are applicable when the random variable is discrete by nature. In case of a continuous random variable like heights or weights; as we have infinite number of values between two distinct values; thus it becomes very difficult to distribute the total probability among all these values.

Therefore, a continuous random variable X is defined in terms of its probability density function $f(x)$ also known as *PDF*.

In such a case the probability density function $f(x)$ is defined as:

A continuous random variable X is designed to follow normal distribution with constant parameters μ = E(X) and Var(x) = σ^2 and written as X ~ N (μ , σ^2)

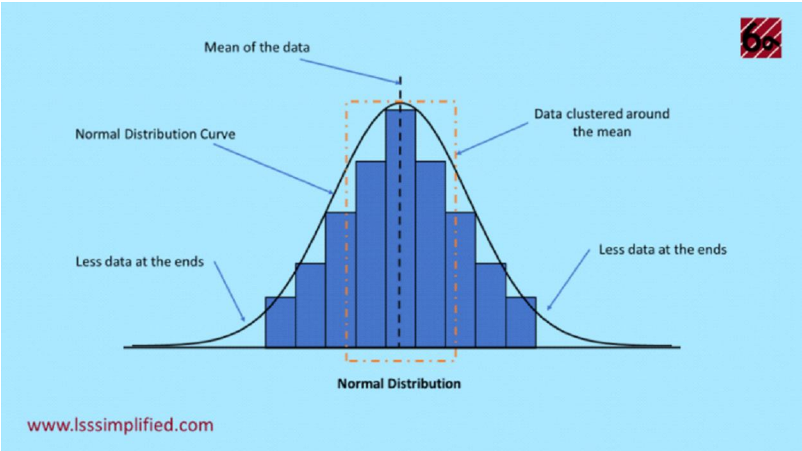
$$f(x) \geq 0, \forall\, x \in (-\infty, \infty)$$

such that $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} . e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad \text{——— (i)}$$

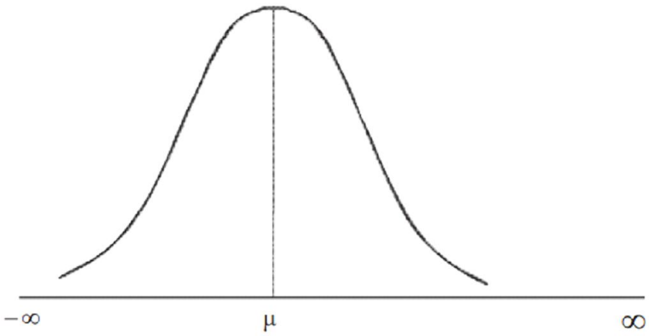
where, $\mu \in (-\infty, \infty)$ is the mean of normal distribution
$\sigma > 0$ is the standard deviation

When a random variable can take on any value within a given range where the probability distribution is continuous (refer 4.1.1), it is called a *normal distribution* or *Gaussian* distribution. A random variable with a Gaussian distribution is said to be normally distributed, and is called a normal deviate

In the normal distribution function given by (i), the curve known as probability curve is bell-shaped with one peak point as shown below:



Picture courtesy: https://www.lsssimplified.com/normal-distribution-for-lean-six-sigma/



The normal distribution is used in the cases where we need to make inferences by taking random samples; and distribution of random variable is not known. This type of distribution is applied to fit the actual observed frequency distribution on many phenomena like weights and heights

A Normal distribution have key features that are easy to spot in graphs:

1. The mean, median and mode of the sample space are exactly the same.
2. The bell-shaped probability curve has one peak point, it means that the normal distribution has a unique mode

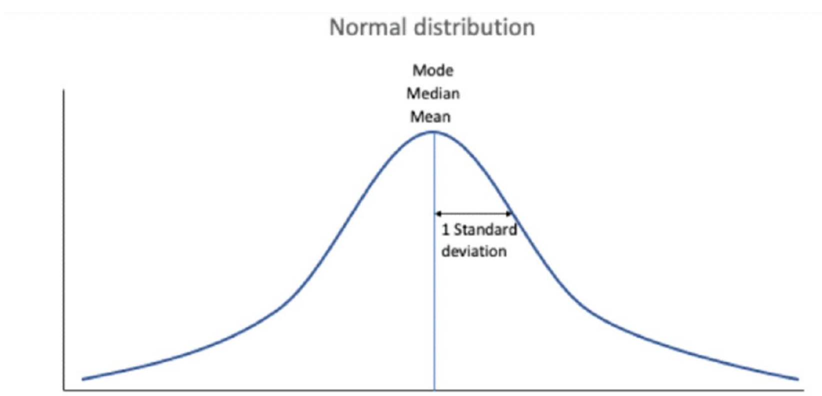
3. The area below the curve $f(x) \quad \forall\, x \in (-\infty, \infty)$ has two tails of the curve extended on both sides and never touch the axis. As the line through x = μ is dividing the normal curve into two equal parts in all aspects which means that the normal curve is symmetrical about x = μ as half the values fall below the mean and half above the mean.
4. In a normal distribution curve the total area below the curve is always equal to 1 unit; i.e., $\int_{-\infty}^{\infty} f(x) = 1$
5. The distribution can be described by two values: the mean and the standard deviation

Galton board is a real-life example of how normal distribution can show the probability distribution https://www.mathsisfun.com/data/quincunx.html

Picture credits: https://en.wikipedia.org/wiki/Bean_machine

4.7.1 Standard Normal Distribution

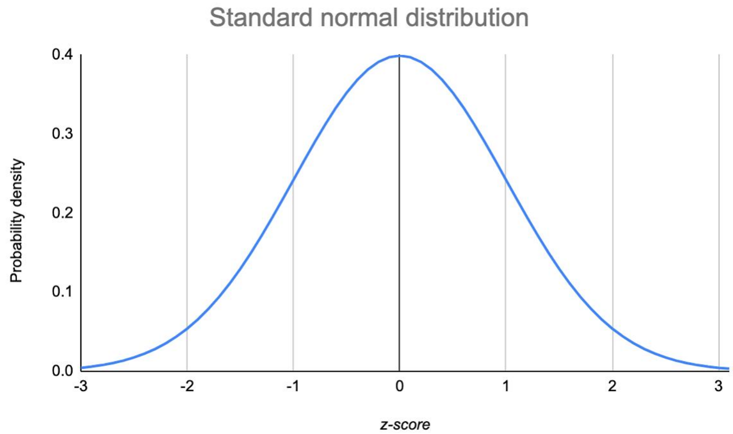
As discussed above, X is a normal variate based on two parameters namely, mean (μ) and standard deviation(σ)



But in a real-life situation, there can be a data set with a mean as 50 and standard deviation of 3 while there can be another data set with a mean of 100 and a standard deviation of 5. How do we compare such different normally distributed data sets?

4.7.2 Z-Score of Normal Distribution

When mean (μ) = 0 and standard deviation(σ) = 1 for a data set, then the normal distribution is called as standard normal distribution



Picture courtesy: https://www.scribbr.com/statistics/normal-distribution/

We make use of data by converting it into a standard normal variate. All normal variate can be converted to standard normal distribution. In order to do so, we calculate the standard score or *Z-score* for each of the data value in the normal variate therefore enabling to compare information since they are on the same scale. This distribution is also called a Z-Distribution.

Basically, the *Z-score* in a standard normal distribution represents how far the said data point from the mean (μ) .

How do we find the Z-Score?

Recall that a normal distribution function is given by $f(x) = \frac{1}{\sigma\sqrt{2\pi}}.e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

In this case Z = $\frac{x-\mu}{\sigma}$ is called the Z-Score.

Example 17:

Calculate Z-Score for a normal distribution of length of 7 rare species of Indian butterfly that you have in your garden

Butterfly	1	2	3	4	5	6	7
Length (in cm)	2	2	3	2	5	1	6

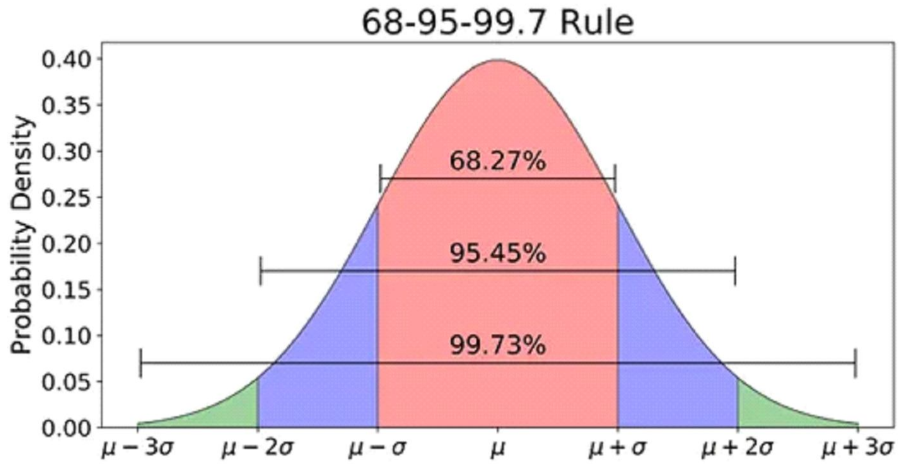
Solution: Here mean = $\mu = \frac{2+2+3+2+5+1+6}{7} = 3$

and $\sigma = \sqrt{\frac{(2-3)^2+(2-3)^2+(3-3)^2+(2-3)^2+(5-3)^2+(1-3)^2+(6-3)^2}{7}} = 1.69$ [as $\sigma = \sqrt{\sum\frac{(x_i-\mu)^2}{n}}$]

Butterfly	1	2	3	4	5	6	7
Length (in cm)	2	2	3	2	5	1	6
Z-Score = $\frac{x-\mu}{\sigma}$	$\frac{2-3}{1.69}$ ≈ –0.59	$\frac{2-3}{1.69}$ ≈ –0.59	$\frac{3-3}{1.69} = 0$	$\frac{2-3}{1.69}$ ≈ –0.59	$\frac{5-3}{1.69}$ ≈ 1.18	$\frac{1-3}{1.69}$ ≈ 1.18	$\frac{6-3}{1.69}$ ≈ 1.78

Notice that butterfly number 3 has the Z-Score = 0, it means that this data point is the mean of the data.

Also note that the Z-score is positive if the data point lies above the mean, and negative if it lies below the mean.



Picture courtesy: https://www.simplypsychology.org/normal-distribution.html

As shown in the graph above, if the data values in a normal distribution are converted into z-scores in a standard normal distribution, then the percentage of the data that fall within specific numbers of standard deviations (σ) from the mean (μ) for bell-shaped curve

1. Data points are symmetrical along the mean (μ)
2. Z-score describes the position of each data point in terms of its distance from the mean, when measured in standard deviation units.
3. The Z-score is positive if the data point lies above the mean, and negative if it lies below the mean.
4. there is a 68.27% probability of randomly selecting a Z-score between -1 and +1 standard deviations from the mean.

   $\Rightarrow \int_{\mu-\sigma}^{\mu+\sigma} f(x)dx$ has probability 68.27%

   where $f(x) = \frac{1}{\sigma\sqrt{2\pi}}.e^{-\frac{1}{2}(z)^2}$

5. 95.45% probability of randomly selecting a score between -2 and +2 standard deviations from the mean.

   $\Rightarrow \int_{\mu-2\sigma}^{\mu+2\sigma} f(x)dx$ has probability 95.45%

   where $f(x) = \frac{1}{\sigma\sqrt{2\pi}}.e^{-\frac{1}{2}(z)^2}$

6. 99.73% probability of randomly selecting a score between -3 and +3 standard deviations from the mean.

   $\Rightarrow \int_{\mu-3\sigma}^{\mu+3\sigma} f(x)dx$ has probability 99.73%

   where $f(x) = \frac{1}{\sigma\sqrt{2\pi}}.e^{-\frac{1}{2}(z)^2}$

Example 18:

Given that mean of a normal variate X is 12 and standard deviation is 4, then find:

a) Find the Z-Score of data point 20

b) The data point if its Z-Score is 5

c) Data point if its Z-Score is -2

Solution:

a) As $\mu = 12$ and $\sigma = 4$ and $x = 20$

$Z = \frac{x-\mu}{\sigma} \Rightarrow \frac{20-12}{4} = 2$

b) As $\mu = 12$ and $\sigma = 4$ and Z = 5

Then $Z = \frac{x-\mu}{\sigma} \Rightarrow 5 = \frac{x-12}{4}$

$\Rightarrow 20 = x - 12$

$\Rightarrow x = 32$

c) As $\mu = 12$ and $\sigma = 4$ and Z = -2

Then $Z = \frac{x-\mu}{\sigma} \Rightarrow -2 = \frac{x-12}{4}$

$\Rightarrow -8 = x - 12$

$\Rightarrow x = 4$

4.7.3 Z-Test for Normal Distribution

If a drug company announces one day that they had found a new drug that cures diabetes, you would want to be sure how true is their claim. Techniques and various hypothesis test are able to tell you if it's probably true, or probably not true. Some of the popular hypothesis tests used in probability distribution are f-test, chi-square test, t-test and Z-test

We are going to discuss one of these tests used in normal distribution data set. To use Z-test, we need to see that:

a) sample size **is greater than 30**.

b) d**ata points should be** independent from each other.

c) data should be randomly selected from a population, where each data point has equally likely of being selected.

d) sample sizes **should be equal** if at all possible.

Let us now see how to use Z-test in a given normal distribution of data set

Example 19

In a district, exam scores of 300 student of class XII are recorded at the end of the session.

a) Ramesh scored 800 marks in total out of 1000. The average score for the batch was 700 and the standard deviation was calculated to be 180. Find out how has Ramesh scored compared to his batch mates in the whole district.

b) Sudha scored 420 marks in the same batch. What can you say about her performance as compared to the batch of 300 students?

c) How much has Abhay scored if he has done better than 44.83% of his batchmates?

Solution:

a) Firstly, we need to find Ramesh's Z-Score and use the respective z-table before we determine how well he has performed as compared to his batch mates

As $\mu = 700$ and $\sigma = 180$ and x = 800

$$Z = \frac{x-\mu}{\sigma} \Rightarrow Z = \frac{800-700}{180} = 0.56$$

Once you have the Z-Score, the next step is choosing between the two Z- Tables. (Refer Appendix at the end)

In the Z-table, go vertically down on the leftmost column to find the value of the first two digits of your Z Score (0.5 in this case) and then go alongside on the topmost row to find the value of the digits at the second decimal position (.06 in this case). Once you have mapped these two values, the intersection of the row of the first two digits and column of the second decimal point in the table gives the value 0.7123 i.e. the area on the left of ordinate corresponding to Z = 0.56. This area also represents the probability of scoring < 800 marks.

Lastly, to get this as a percentage we multiply that number with 100 i.e. 0.7123 x 100 = 71.23%.

Hence, we can say that Ramesh did better than 71.23% of students in the district.

b) In the case of Sudha, $\mu = 700$ and $\sigma = 180$ and x = 420

$$Z = \frac{x-\mu}{\sigma} \Rightarrow Z = \frac{420-700}{180} = -1.56$$

Looking at the Z-Table we can say that it maps to 0.0594 and hence we can say that Sudha did better than 100$\times$ 0.0594 =5.94% of students in the district.

c) If Abhay has done 44.83% better than his batchmates, then his score on Z-Table is 44.83 ÷ 100 = 0.4483 which corresponds to Z-Score = - 0.13

Here $\mu = 700$ and $\sigma = 180$ and Z = -0.13

Therefore $Z = \frac{x-\mu}{\sigma} \Rightarrow -0.13 = \frac{x-700}{180}$

$\Rightarrow -23.4 = x - 700$

$\Rightarrow x = 676.6 \approx 677$

Which means that Abhay has scored approximately 677 marks out of 1000

Example 20:

Given that the scores of a set of candidates on an IQ test are normally distributed. If the IQ test has a mean of 100 and a standard deviation of 10, what is the probability that a candidate who takes the test will score between 90 and 110?

Solution: $P(90 < X < 110) = P(X < 110) - P(X < 90)$

$$\Rightarrow P(90 < X < 110) = P(-1 < Z < 1) = P(Z < 1) - P(Z < -1)$$

$$= 0.8413 - 0.1587 = 0.6826$$

4.8 CHECK YOUR PROGRESS

1. State which of the following are not the probability distributions of a random variable. Give reasons for your answer

a.

X	-1	0	1	2
P(X)	0.1	0.8	0.001	0.2

b.

X	1	2	3	4	5
P(X)	0.1	0.4	0.05	-0.2	0.2

c.

X	-2	2	5
P(X)	0.5	0.2	0.3

2. A lady's bag contains 2 black and 1 red pens. One pen is drawn at random and then put back in the box after noting its colour. The process is repeated again. If X denotes the number of red pens recorded in the two draws. Describe X.
3. What is the mean of the numbers obtained on throwing a die having written 1 on three faces, 2 on two faces and 5 on one face?
4. Raheem tossed a fair coin 10 times, find the probability of (i) exactly six heads (ii) at least six heads (iii) at most six heads.
5. Find the probability of getting 5 exactly twice in 7 throws of a fair die.
6. Let X denote the number of hours a class XII student studies during a randomly selected school day. The probability that X can take the values x_i, for an unknown constant 'k'

$$P(X = x_i) = \begin{cases} 0.1, & if\ x_i = 0 \\ kx_i, & if\ x_i = 1 or 2 \\ k(5 - x_i), & if\ x_i = 3, 4 \end{cases}$$

   a. Find the value of k.
   b. What is the probability that the student studied for at least two hours? Exactly two hours? At most two hours?
7. How many times must Sumit toss a fair coin so that the probability of getting at least one head is more than 90%?
8. A pair of dice is thrown and the random variable X represents the sum of the numbers that appear on the two dice. Calculate the mathematical expectation of X.
9. Find the variance of a bernoulli random variable whose probability of success is 0.6.
10. If the mean and variance of a binomial distribution are 4/3 and 8/9 resp. find P(x=1)
11. What is the expected value of number of tails on a throw of a fair coin?
12. A customer care company receives an average of 4.5 calls every 5 minutes. If each customer executive can handle one of these calls over the 5-minute period. But if an executive is not unavailable to take the call, then the call is put on hold. Assuming that the calls received by the customer care company follows a Poisson distribution, what is the minimum number of customer executives are needed on duty so that calls received are placed on hold for at the most 10% of the time?

13. A statistician records the number of trucks approaching a particular intersection to analyze the flow of traffic. He observes that on an average 1.6 trucks approach the intersection every minute. Assuming that the number of trucks approaching the intersection, follow a Poisson distribution model, what is the probability that 3 or more trucks will approach the intersection withing a minute?

14. A computer disk manufacturer tests disk quality on random basis before approving it. The approval is based on the number of errors in a test area on each disk and follows Poisson distribution with $\lambda = 0.2$. What is the percentage of test areas having two or a smaller number of errors?

15. In a Poisson distribution, if mean is 2, what is the variance?

16. It is given that 3% defective electric bulb are manufactured by a company. Using Poisson distribution, find the probability of 100 bulbs will contain no defective bulbs. (Use $e^{-3} = 0.05$)

17. The mortality rate for a certain disease is 0.007. Using Poisson distribution, calculate the probability for 2 deaths in a group of 400 people

18. An ice-cream parlour receives a customer at an average rate of 4 per minute. If the number of customers received by the parlour follows a Poisson distribution, what is the approximate probability that 16 customers will be coming to the parlour in a particular 4-minute period on a given day?

19. **Using Z-Table, Calculate**

    a) **P $(Z < 1.20)$**

    b) **P $(Z \leq 1.20)$**

    c) **P $(-0.5 \leq Z \leq 1.0)$**

    d) **P $(-1.0 \leq Z \leq 1.0)$**

20. A company conducted an IQ test for randomly select 50 employees. Volunteer A scored 74 out of the possible 120 points. If the average IQ test score was recorded as 62 and the standard deviation was 11. How well did volunteer A perform on the test compared to the other volunteers?

21. An average ceiling fan manufactured by the Jagdeep Corporation lasts 300 days with a standard deviation of 50 days. Assuming that the ceiling fan;s life is normally distributed, what is the probability that a ceiling fan will last at most 365 days?

22. In a survey of daily travel time (in minutes) of students to reach their school was recorded as follows:

    26, 33, 65, 28, 34, 55, 25, 44, 50, 36, 26, 37, 43, 62, 35, 38, 45, 32, 28, 34

    If the mean travelling time is 38.8 minutes and the standard deviation is 11.4 minutes. Convert the travel time of each student into a Z-Score

23. In an examination, 2000 students appeared and the mean of the normal distribution of marks is 30 with standard deviation as 6.25. Find out how many students are expected to score.

    i. between 20 and 40 marks.

    ii. less than 25 marks

24. In a normal distribution, 31% of the articles are under 45 and 8% are over 64. Calculate the mean and standard deviation of the distribution

4.9 QUICK RECAPITULATION

1. A random variable is a real valued function whose domain is the sample space of a random experiment, denoted by X
2. More than one random variable can be defined on a given sample space.
3. The random variable X can take distinct and countable (measurable) values in between values for it, then it is called a *discrete random variable.*
4. When a random variable whose value is obtained by measuring and it takes many values between two values, is called a *continuous random variable*
5. The table of possible outcomes and their respective probability is called *Probability distribution table* for the given random variable X
6. In a probability distribution the sum of all the probabilities is always one
7. In an experiment, for a given random variable X whose possible finite values x_1, x_2, x_3, …., x_n occur with probabilities p_1, p_2, p_3, …,p_n respectively such that

   $\sum p_i = 1$

   Then the mathematical expectation is the weighted average of the possible values of X given by $E(X) = x_1p_1 + x_2p_2 + x_3p_3 + \cdots + x_np_n = \sum_{i=1}^{n}(x_ip_i)$
8. Mean expectation value is a parameter and not a statistic.
9. Variance enables us to study the variability of random variable from the mean expectation
10. Let X be a discrete random variable whose possible finite values x_1, x_2, x_3,..., x_n occur with probabilities p_1, p_2, p_3, …,p_n respectively.

    Let μ = E (X) be the mean of X. Then the variance of X, denoted by Var (X) or σ_x^2 is given by $\sigma_x^2 \Rightarrow \text{Var}(X) = \sum_{i=1}^{n} x_i^2 p_i - \left(\sum_{i=1}^{n} x_i p_i\right)^2$

    In other words, Var (X) = $E(X^2) - [E(X)]^2$, where $E(X^2) = \sum_{i=1}^{n} x_i^2 p_i$
11. The standard deviation denoted by $\sigma_x = \sqrt{Var(X)}$
12. In a random experiment, a collection of trials is called *Bernoulli trials*, if:
    i. The number of trials is finite.
    ii. The trials are independent by nature.
    iii. Each trial has exactly two outcomes defined as success and failure.
    iv. The probability of success remains the same in each trial.
13. Probability of '*r*' successes in '*n*' Bernoulli trials is given by

    P ('r' successes) = $C_r^n\, p^r q^{n-r} = \frac{n!}{r!(n-r)!}\, p^r q^{n-r}$

    Where n = number of trials

    r = number of successful trials = 0, 1, 2, 3, …, n

p = probability of a success in a trial

q = probability of a failure in a trial

And, $p + q = 1$

14. P('r' successes) is the (r+1)th term in the binomial expansion of $(q + p)^n$
15. The binomial distribution with n Bernoulli trials and Probability of success p is also denoted by B(n, p)
16. In a binomial distribution having 'n' number of Bernoulli trials where p denotes the probability of success and q denotes the probability of failure, then
    i. Mean = np
    ii. Variance = npq
    iii. Standard Deviation = $\sqrt{npq}$
17. Let X be the discrete random variable which represents the number occurrence of events over a period of time. If X follows the Poisson distribution, then the probability of occurrence of 'k' number of events over a period of time is given by

$$P\,(X = k) = f(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Where e is Euler's number (e = 2.71828…)

'k' is the number of occurrences of the event such that k = 0, 1,2, …

And λ = E(X) = Var(X), is a positive real number

With existence conditions:

i. $\sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = 1$

ii. $f(k) \geq 0$, for k = 0, 1, 2, …

18. A continuous random variable X is defined in terms of its probability density function $f(x)$ also known as *PDF* as well
19. A continuous random variable X is designed to follow normal distribution with constant parameters μ = E(X) and Var(x) = σ^2 and written as X ~ N (μ , σ^2)

$$f(x) \geq 0, \forall\, x \in (-\infty, \infty)$$

such that $f(x) = \frac{1}{\sigma\sqrt{2\pi}} . e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ ———————— (i)

where $\mu \in (-\infty, \infty)$ is the mean of normal distribution and $\sigma > 0$ is the standard deviation

20. When a random variable can take on any value within a given range where the probability distribution is continuous, it is called a *normal distribution* or *Gaussian* distribution.
21. A random variable with a normal/Gaussian distribution is said to be normally distributed, and is called a normal deviate
    a. The mean, median and mode of the sample space are exactly the same.
    b. The bell-shaped probability curve has one peak point, it means that the normal distribution has a unique mode

c. The area below the curve $f(x) \quad \forall x \in (-\infty, \infty)$ has two tails of the curve extended on both sides and never touch the axis. As the line through x = μ is dividing the normal curve into two equal parts in all aspects which means that the normal curve is symmetrical about x = μ as half the values fall below the mean and half above the mean.

d. In a normal distribution curve the total area below the curve is always equal to 1 unit; i.e., $\int_{-\infty}^{\infty} f(x) = 1$

e. The distribution can be described by two values: the mean and the standard deviation

22. When mean (μ) = 0 and standard deviation(σ) = 1 for a data set, then the normal distribution is called as standard normal distribution

23. In a normal distribution of data, the Z-score is given by Z = $\frac{x-\mu}{\sigma}$

24. When the Z-score is positive if the data point lies above the mean, and negative if it lies below the mean

25. When the data values in a normal distribution are converted into Z-scores in a standard normal distribution, then the percentage of the data that fall within specific numbers of standard deviations (σ) from the mean (μ) for bell-shaped curve is constant.

i. Data points are symmetrical along the mean (μ)

ii. Z-score describes the position of each data point in terms of its distance from the mean, when measured in standard deviation units.

iii. The Z-score is positive if the data point lies above the mean, and negative if it lies below the mean.

iv. There is a 68.27% probability of randomly selecting a Z-score between -1 and +1 standard deviations from the mean.

a. $\Rightarrow \int_{\mu-\sigma}^{\mu+\sigma} f(x)$dx has probability 68.27%

b. where $f(x) = \frac{1}{\sigma\sqrt{2\pi}} . e^{-\frac{1}{2}(z)^2}$

v. 95.45% probability of randomly selecting a score between -2 and +2 standard deviations from the mean.

a. $\Rightarrow \int_{\mu-2\sigma}^{\mu+2\sigma} f(x)$dx has probability 95.45%

b. where $f(x) = \frac{1}{\sigma\sqrt{2\pi}} . e^{-\frac{1}{2}(z)^2}$

vi. 99.73% probability of randomly selecting a score between -3 and +3 standard deviations from the mean.

a. $\Rightarrow \int_{\mu-3\sigma}^{\mu+3\sigma} f(x)$dx has probability 99.73%

b. where $f(x) = \frac{1}{\sigma\sqrt{2\pi}} . e^{-\frac{1}{2}(z)^2}$

26. Some of the popular hypothesis tests used in probability distribution are f-test, chi-square test, t-test and Z-test

27. To use Z-test, we need to see that:

    i. sample size **is greater than 30**.

    ii. d**ata points should be** independent from each other.

    iii. data should be randomly selected from a population, where each data point has equally likely of being selected.

    iv. sample sizes **should be equal** if at all possible.

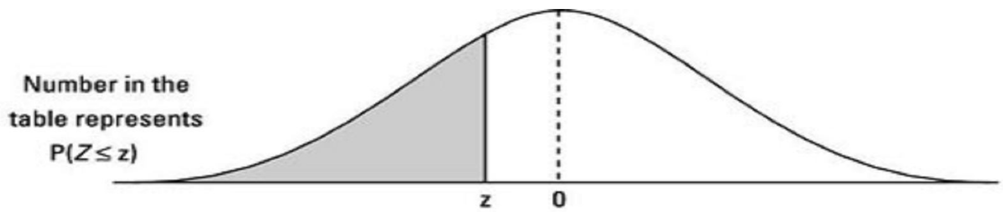
    v. for convenient calculation of Z-Score, we use Z-Table to interpret normal distribution data set

4.10 ANSWER KEY TO CHECK YOUR PROGRESS

1. (a) no, $\Sigma p \neq 1$ (b) no, $p < 0$ (c) yes **2.** X = 0,1,2 **3.** 2 **4.** 105/512, 193/512, 53/64 **5.** $C_2^7 \frac{5^5}{6^7}$
6. k = 0.15, 0.75, 0.3, 0.55 **7.** 4 or more times **8.** 7 **9.** 0.24 **10.** 32/81 **11.** 0.5 **12.** 7 **13.** 0.217 **14.** 99.89% **15.** 2 **16.** 0.05 **17.** 0.235 **18. 0.099 19. (a)** 0.11507 , (b) 0.88493, (c) **0.5328, (d) 0.6826 20.** better than 43 other volunteers **21.** 90% **23.** (i) 1781, (ii). 424 **24.** Mean = 50, SD = 10

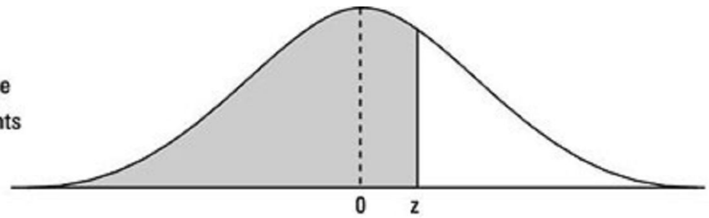
4.11 APPENDIX

Z- SCORE TABLE



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
–3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
–3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
–3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
–3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
–3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
–3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
–3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
–2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
–2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
–2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
–2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
–2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
–2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
–2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
–2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
–2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
–2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
–1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
–1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
–1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
–1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
–1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
–1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
–1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
–1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
–1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
–1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
–0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
–0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
–0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
–0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
–0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
–0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
–0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
–0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
–0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
–0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Number in the table represents P(Z ≤ z)



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

Table Courtesy: https://www.dummies.com/education/math/statistics/how-to-use-the-z-table/

❑❑❑