**UNIT-III**

1.1 DISCRETE UNIFORM DISTRIBUTION

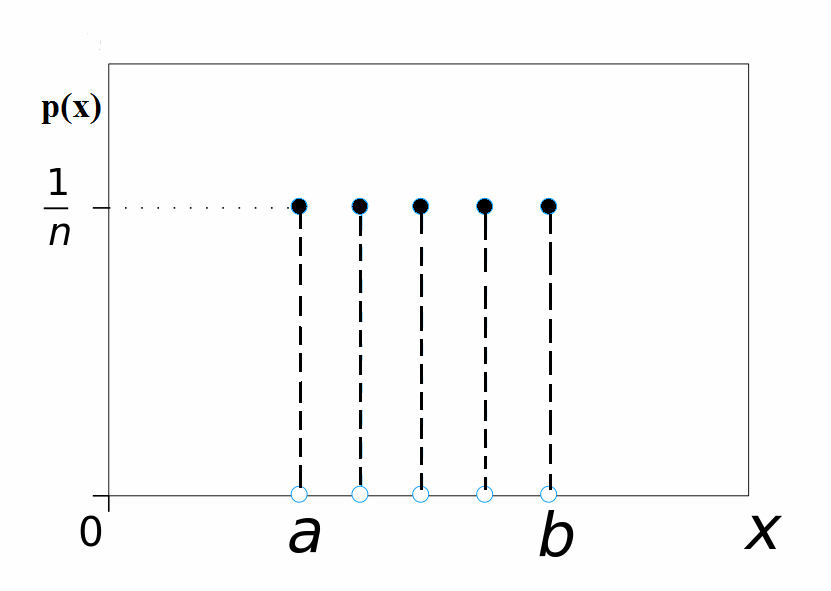
A random variable has a uniform distribution when each value of the random variable is equally likely, and values are uniformly distributed throughout some interval. Uniform distributions can be discrete or continuous, but here we are considering only the discrete case.

Definition: A random variable X is said to have a discrete uniform distribution over the range [1, n] if its probability mass function is expressed as follows:

 (1.1)

Here n is known as the parameter of the distribution and lies in the set of all positive integers. Equation (1.1) is also called a **discrete rectangular distribution**.

Such distributions can be conceived in practice if under the given experimental conditions, the different values of the random variable become equally likely. Thus for a die experiment, Rolling a single die is one example of a discrete uniform distribution; a die roll has six possible outcomes: 1,2,3,4,5, or 6. There is a 1/6 probability for each number being rolled, for an experiment with a deck of cards such distribution is appropriate.



1.1.1Moments of rectangular distribution

Let a discrete random variable X have discrete uniform distribution over the range [1, n], then by definition





Also





Now variance is given by







1.1.2 Moment generating function

By definition







1.1.3 Bernoulli trials

Let us perform a sequence of trials of a random experiment so that

(a) For each trial, there are only two possible outcomes, say, success and failure;

(b) The probabilities of the occurrence of these outcomes remain the same throughout the trials; and

(c) The trials are carried out independently.

Trials performed under these conditions are called ***Bernoulli*** *trials*. Despite of the simplicity of the situation, mathematical models arising from this basic random experiment have wide applicability.

If we denote ‘success’ by *S*, and event ‘failure’ by *F*. Also, let  *p*, denote probability of success P(S) and P(F) is the probability of failure in any trial. where *p+q=*1.

Possible outcomes resulting from performing a sequence of Bernoulli trials can be symbolically represented by

P (SSFSFFFS. . .FSF) = P(S)P(S) P(F) P(S) P(F) P(F) P(F) F(S)  ... P(F) P(S) P(F)

= p.p. q.p. q.q.q.p...q.p.q = p.p.p …...p . q.q.q q

Its probability mass function is given by

****

*A random experiment which has only two outcomes success(S) and failure (F) is called Bernoulli experiment*

1.1.4 Mean and Variance

****

**Now **

**Hence**

****

1.2 Binomial distribution

This distribution was discovered by James Bernoulli (1654-1705) in the year 1700 and was first published posthumously in 1713, eight years after his death. The probability distribution of a random variable *X* representing the number of successes in a sequence of *n* Bernoulli trials, regardless of the order in which they occur, is frequently of considerable interest in a variety of fields of application.

**Derivation:** Consider a set of n independent Bernoullian trials (n being finite) in which the probability ‘p’ of success in any trial is constant for each trial, and then is the probability of failure in any trial.

The probability of x successes and consequently (n - x) failures in n independent trials, in a specified order (say) SSFSFFFS. . .FSF (where S represents success and F represents failure) is given by the compound probability theorem by the expression:

P (SSFSFFFS. . .FSF) = P(S)P(S) P(F) P(S) P(F) P(F) P(F) F(S) x ... x P(F) P(S) P(F)

=p.p. q.p. q.q.q.p...q.p.q = p.p.p …...p . q.q.q q

{ x factors} {(n - x) factors}

But x successes in n trials can occur in ways and the probability for each of these ways is same, viz.,. Hence the probability of x successes in n trials in any order is given by the addition theorem of probability by the expression

The probability distribution of the number of successes, so obtained is called the Binomial probability distribution.

A random variable X is said to follow binomial distribution if it assumes only non-negative values and its probability mass function is given by



The two independent constants n and p in the distribution are known as the parameters of the distribution.

Symbolically, it is represented by i.e., random variable X follows binomial distribution with parameters n and p.

1.2.1 Physical conditions for Binomial Distribution

We get the binomial distribution under the following experimental conditions:

(i) Each trial results in two exhaustive and mutually disjoint outcomes, termed as success and failure.

(ii) The number of trials ‘n’ is finite.

(iii) The trials are independent of each other.

(iv) The probability of success ‘p’ is constant for each trial.

The trials satisfying the conditions (i), (iii) and (iv) are also called **Bernoulli trials.**

1.2.2 Moments of Binomial Distribution

Let a random variable X follows binomial distribution with parameters n and p i.e., .Then moments about origin of binomial distribution are obtained as follows:









Thus mean of the binomial distribution is **np**.

Now second moment about origin of binomial distribution is given by













*Hence variance of binomial distribution is given by*





C.T.M

* If X ~ B(n, p), then and 
* Variance = = Mean (as 0 <q <1)

***Hence, for the binomial distribution, variance is less than mean***

1.2.3 Recurrence Relation for the moments of Binomial Distribution

(Renovsky Formula)

Let X~B(n,p) then by definition rth central moment is given by



  (1)

Differentiating (1) w.r. to p, we get







 (2)

Substituting r=1,2,3 successively in (2) , we get





















Hence





, 

When p = q =1/2 the distribution is symmetric and leptokurtic for all values of n

1.2.4 Mode of Binomial Distribution

Mode is the value of x for which probability p(x) is maximum

We have

 so that





= 

 (1.1)

**Case I.** When (n + 1) p is not an integer. Let (n + 1) p = m +f, where m is an integer and f is fractional such that 0 <f< 1. Substituting in (1.1), we get

 (1.2)

From (2), we see that









It means that p(x) is maximum at x=m.

Thus, in this case there exists unique modal value for binomial distribution and it is m, the integral part of.

**Case II:** When (n + 1) p is an integer Let (n + 1) p = m (an integer). Substituting in (1.1), we get

 (1.3)

From (1.3), it is obvious that:



Now proceeding as in case I, we have



Thus, in this case the binomial distribution is bimodal and the two modal values are m and (m -1).

EXAMPLE : Determine the binomial distribution for which the mean is 8 and variance 4 and find its mode.

Solution. Let X ~ B (n, p), then we are given that

E (X) = np =8 (\*)

and Var (X) = npq = 4 (\*\*)

Dividing (\*\*) by (\*), we get

 Hence from (\*), we obtain 

Thus the given binomial distribution has parameters *n =* 16 and 

**Mode.** We have (n + l)p = 8.5, which is not an integer. Hence the unique mode of the binomial distribution is 8, the integral part of (n + 1) p.

Moment Generating Function of Binomial Distribution

Let X ~ B (n, p),then





Now













Additive Property of Binomial Distribution.

Let X - B (n1, p1) and Y~ B (n2, p2) be independent random variables. Then

, 

We have

 { since X and Y are independent}

=+ (1.4)

Since (1.4) cannot be expressed in the form , from uniqueness theorem of m.g.f’s it follows that X +Y is not a binomial variate. Hence, in general the sum of two independent binomial variates is not a binomial variate.

*In other words, binomial distribution does not possess the additive or reproductive property.*

**However**, if we take, (say), then from (1.4)

Mx+y (t) = 

which is the m.g.f. of a binomial variate with parameters (n1 + n2, p). Hence by uniqueness theorem of m.g.f.’s X + Y ~ B (n1 + n2, p). Thus the binomial distribution possesses the additive or reproductive property if p1 = p2

POISSON DISTRIBUTION

It was discovered by the French mathematician and physicist Simeon Denis Poisson (1781-1840). Poisson distribution is a limiting case of the binomial distributionunder the following conditions:

(i) n, the number of trials is indefinitely large, i.e., n →.

(ii) p. the constant probability of success for each trial is indefinitely small, i.e., p→0.

(iii) np = λ, (say) is finite.

Thus , , where λ is a positive real number. The probability of x successes in a series of n independent trials is:









as ****

**Definition.** A random variable X is said to follow a Poisson distribution if it assumes only non-negative values and its probability mass function is given by



Here λ is known as the parameter of the distribution. It is denoted by X~P(λ ) i.e., X is a Poisson variate with parameter .

Moments of the Poisson Distribution**.**

It X ~ P(λ ) i.e., X is a Poisson variate with parameter λ .Then first moment about origin is given by

****





****

As ****

Second moment about origin is given by

****



****

****

****

Hence



Thus Poisson distribution 

Recurrence Relation for Moments of the Poisson Distribution

The rth order central moment is given by

   (1.5)

Differentiating (1.5) w.r. to, we get









 (1.6)

Now putting r=1, 2 and 3 successively in (1.6), we get

,

,



Co-efficients of skewness and kurtosis are given by:

, 

, 

Hence the Poisson distribution is always a skewed distribution.

Moment Generating Function of the Poisson Distribution.

****





Additive or Reproductive Property of Independent Poisson Variates

“Sum of independent Poisson variates is also a Poisson variate”. More elaborately, If Xi, (i= 1, 2, n) are independent Poisson variates with parameters λi , i = 1, 2 ..n respectively, then is also a Poisson variate with parameter

**Proof:** If 

then

****

 {as are independent}



which is the m.g.f of a Poisson variate with parameter λ1+λ2 + ….. +λn.

Hence, by uniqueness theorem of m.g.f’s,  is also a Poisson variate with parameter

* *the converse of the above result is also true*,
* *The difference of two independent Poisson variates is not a Poisson variate.*

GEOMETRIC DISTRIBUTION

Let us consider sequence of Bernoullian trial with p be the probability of success and suppose X be the number of failures preceding the first success .Then X follows a geometric distribution with parameter p. Its density function is



It is denoted by X~Geo(p)

Since the first x trials are the failure and the last trial must be the successes.

Moment generating function of Geometric distribution

By definition













Thus the variance of the geometric distribution is greater than its mean.

Lack of memory in geometric distribution.

This distribution is said to lack memory in a sense .let us consider that an event E can occur at any of the times and occurrence time i.e., waiting time has geometric distribution with parameter p. Thus



Again suppose that the event E has not occurred before k , i.e., .Again suppose that is the additional time required for the event E to occur. Here we can show that



It means that additional time to wait has same distribution as initial time to wait.

Proof: If X~Geo(p) then we have



 (1.7)

Now

 (1.8)

Using (1.7)



 Using (1.8)

Since it does not depends on k, it, in a sense *lacks memory* how much we have shifted the time origin.

Hyper-geometric distribution

Let us consider an urn containing N balls out of which M are white and N-M are red. Suppose we wish to draw a random sample (SRSWOR) of n balls form the urn. Then the probability of getting k white balls out of n (k<n) is given by



Since k white balls out of M white balls can be drawn in ways and the remaining

(n-k) red balls out of (N-M) red balls can in  ways. So the total number of favourable cases will be and exhaustive number of cases will be , since n balls out of N ball can be drawn in.

**Definition**: A random variable X is said to follow hyper-geometric distribution with parameters N, M and n, if it assumes only non-negative values and its probability mass function is given by



Where N and M are positive integers, M must not exceed N and n is also positive integer which is at most N.

Mean and variance of hyper geometric distribution

By definition













Now









Hence





Limiting case of hyper-geometric distribution

*“Hyper geometric distribution tends to Binomial distribution as and *

**Proof**: Probability mass function of hyper geometric distribution with parameters N,M and n is given by









Proceeding to the limits and , we get







Which is binomial distribution with parameters n and p

Negative binomial distribution

Sometimes the size of the sample is not fixed in advance but it is determined as the sample size required to achieve rth success ( This process is called inverse sampling). A random variable X ,the number of failures before rth success result either

Let us consider n Bernoulli trials with constant probability of success p and the trials are independent.

Suppose denote the probability of x failures preceding rth success in x+r trials, now the rth trial must be success with probability p. In the remaining (x+r-1) trial we must have (r-1) successes whose probability is given by binomial law



Therefore, probability of x failures preceding rth success in x+r trials by compound probability law is the product of two probabilities



A random variable X is said to follow negative binomial distribution if its probability mass function is given by

 (\*)

Remark 1: We have So that





Substituting in above relation in probability mass function of negative binomial distribution we have

** (\*\*)

Expression given above is the (x+1)th term in the expansion of , which is a binomial expansion with negative index. Hence named as negative binomial distribution.

Remark 2:If we substitute and  so that ,then from (\*\*) we have

** (\*\*\*)

We see that (\*\*\*) is the general term in the negative binomial expansion of 

Remark 3: If we take r=1 in (\*) we get ,



Which is the probability mass function of geometric distribution. Hence negative binomial distribution may be regarded as generalization of geometric distribution.

1.6.1 moment generating function of negative binomial distribution

**

 (1.8)

Now first moment about origin





Hence mean of negative binomial is rP

Second moment about origin of negative binomial distribution is given by





Now variance is given by





Remark: Since Q>1 ,  which means that mean of negative binomial distribution is less than its variance.

Poisson distribution as a limiting case of negative binomial distribution

Negative binomial distribution tends to Poisson distribution as ,such that

is finite. We have

.

Proceeding to the limits, we get







Which is probability function of Poisson distribution with parameter 

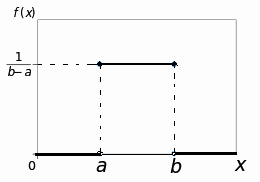
Unit-IV Semester 2

RECTANGULAR (OR UNIFORM) DISTRIBUTION

**Uniform distribution:** A random variable X is said to follow uniform rectangular distribution if assumes only its probability mass function is given by



* Here a and b, (a <b) are the two parameters of the distribution. A uniform or rectangular variate X on the interval (a, b) is written as : X ~ U [a, b] or X~R[a,b].



Properties

* The distribution is called uniform distribution on (a, b) since it assumes a constant (uniform) value for all x in (a, b).
* The distribution is also known as rectangular distribution, since the curve **y =f(x)** describes a rectangle over the x-axis and between the ordinates at x = a and x = b.

Since F(x) is not continuous at x= a and x = b, it is not differentiable at these points. Thus exist everywhere except at the points except at the points x = a and x = b . Its distribution function is given by



****

Mean and variance of uniform rectangular distribution

Let X ~ U [a, b] then by definition





Now second moment about origin is given by





Hence



=

Moment generating function of variance of uniform rectangular

distribution

If X~ U (a, b), then by definition m.g.f (about origin) is given by





distribution function of uniform rectangular distribution



EXAMPLE**:** The average amount of weight that a person can lose by a slimming therapy over the period of four months is uniformly distributed from 0 to 20 kgs.

(i) Find the probability a person will lose between 5 and 10 kgs of weight by this therapy.

(ii) Also find probability that weight loss is atleast 10 kgs.

**Solution**: Let random variable X denote amount of weight lost due to slimming therapy , then X is uniformly distributed [0,20] with p,d,f given by



Probability a person will lose between 5 and 10 kgs by



(ii) Probability that weight loss is atleast 10 kgs



NORMAL DISTRIBUTION

It was first discovered in 1733 by English mathematician De-Moivre, who obtained this continuous distribution as a limiting case of the binomial distribution and applied it to problems arising in the game of chance. It was also known to Laplace, no later than 1774 but through a historical error it was credited to Gauss, who first made reference to it in the beginning of 19th century (1809), as the distribution of errors in Astronomy. The normal model has, nevertheless, become the most important probability model in

statistical analysis.

**Definition**: A r.v. X is said to have a normal distribution with parameters ,(called ‘mean’) and  (called ‘variance’), if its p.d.f. is given by the probability law

 ; 

Here andare called its parameters and we write it as 

Standard Normal Variate

If then  is a standard normal variate with and  and we write 

Proof: If , and  then

E(Z) =

 and 

The p.d.f of standard normal variate Z is given by

 ; 

Chief Characteristics of the Normal Distribution and Normal

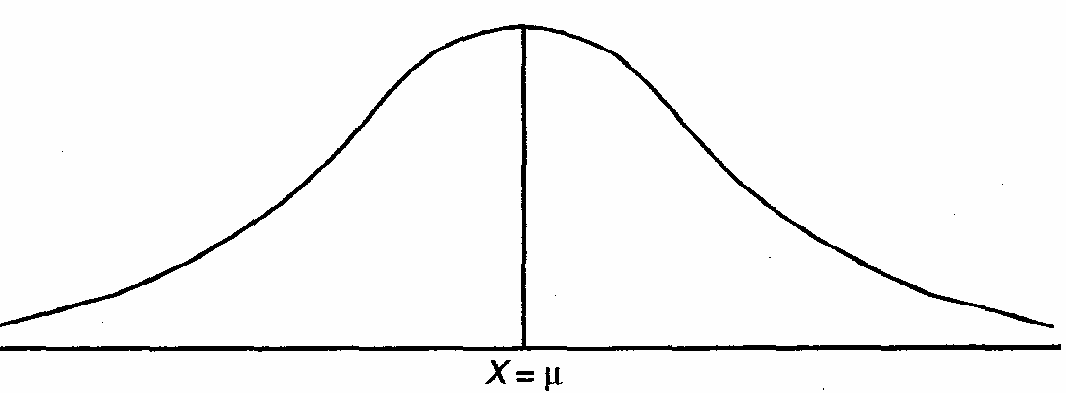
Probability Curve

The normal probability curve with mean  and standard deviationis given by the equation:



It has the following properties

(i) The curve is bell-shaped and symmetrical about the line , and is non-zero over the entire real line.



(ii) Mean, median and mode of the distribution coincide.

(iii) As x increases numerically, f(x) decreases rapidly, the maximum probability occurring at the point , and is given by:

[p (x)]max = 

(iv) and.

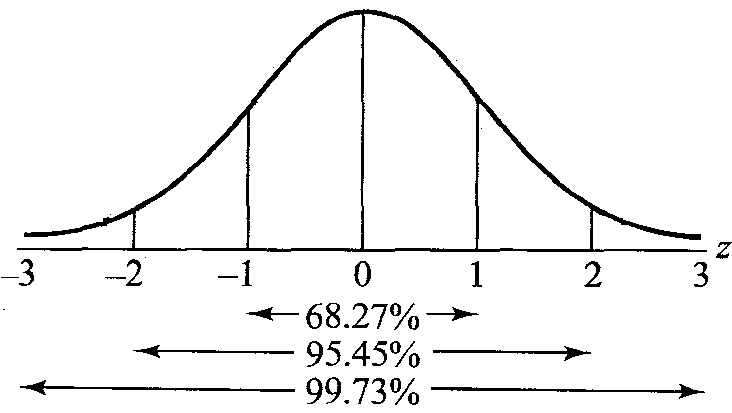
(v) , i.e.,odd order moments vanishes

(vi) Since f (x) being the probability, can never be negative, no portion of the curve lies

below the x-axis.

(vii) Linear combination of independent normal variates is also a normal variate.

(viii) Area property





 and



Mode of Normal Distribution

Mode is the value of x for which f(x) is maximum, i.e., mode is the solution of



For normal distribution with mean with mean  and standard deviation,

 so that



where c = log , is a constant. Differentiating w.r. to x, we get





 (2.1)

Now 

At the point  we have from (2.1)



Hence  is the mode of the normal distribution.

Median of Normal Distribution

If M is the median of the normal distribution, we have







But



So we have





Hence, for the normal distribution, Mean = Median.

M.G.F. of Normal Distribution

The m.g.f. (about origin) is given by

If X~ N (, ), then by definition m.g.f is given by





{ By substituting Z =  so that  and  }







Hence 

M.G.F. of Standard Normal Variate

If X~ N (u, ), then standard normal variate given by

Z = 





*Alternatively:* we know that ,taking and  in the m.g.f of the random variable X we get the desired result.

Mean of Normal Distribution

If X~ N (u, ), we have

E[X] = Mean= 

= (2.2)

Put Z =  so that  and  in (2.2) we get

Mean==



=+0 (2.3)

Since the integrand is an odd function of z

Put 

We get

==

==

Substituting in (2.3) we get

Mean== Hence mean is 

Variance of Normal Distribution

If X~ N (, ), we have

== (2.4)

Put Z =  so that  and  in (2.4) we get

= (2.5)

put  in (2.5) we get

Variance**=**

==

==

Hence Variance of Normal Distribution 

A linear combination of independent normal variates is also a

variate

Let, (i = 1,2, 3, ..., n) be n independent normal variates with mean  and variance respectively. Then

 (2.6)

The m.g.f. of their linear combination ,where a1, a2, ..., an are constant is given by

 (2.7)



From (2.6), we have



which is the m.g.f of a normal variate with mean  and variance .

Hence by uniqueness theorem of m.g.f,

 (\*)

*Important deductions*

* If we take ,then 
* If we take ,then 

Thus we see that the **sum as well as the difference** of two independent normal variates is also a normal variate.

* If we take ,then 

i.e., the sum of independent normal variates is also a normal variate, which establishes the **additive property** of the normal distribution.

Moments of Normal Distribution

Odd order moments about mean are given by

=

=





{Since the integrand is an odd function of z.}

Hence odd order moments of normal distribution vanishes

**Even order moments** about mean are given by:

==

Put Z =  so that  and  we get

Or 



 {by substituting  so that }



Changing n to, we get



 [as ]



Which gives the recurrence relation for the moments of normal distribution.

Exercise: show that 

Sol: By using the recurrence relation  and substituting we get





…………………….





Multiplying we get 

Cumulants and cumulative function

By definition cumulative function is given by



But 

Where are various Cumulants



Area Property (Normal Probability Integral)

If X~ N (u, ), then the probability that random value of X will lie between . and  is given by:

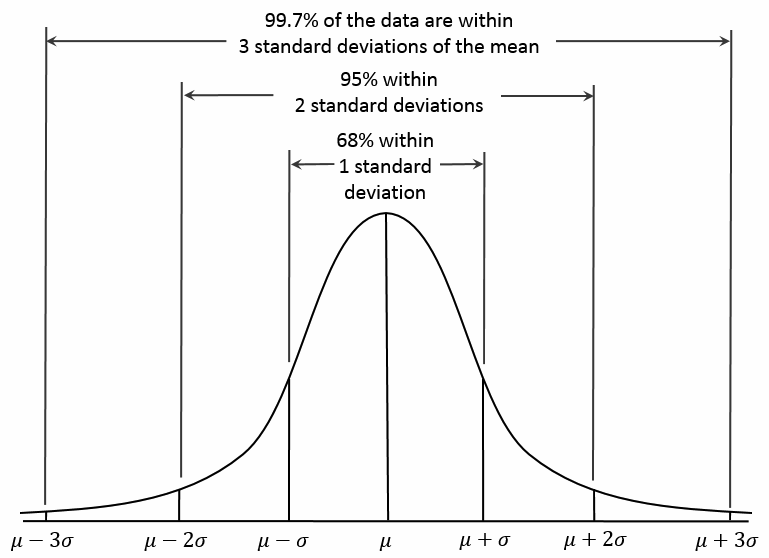


Put Z =  so that 

When 



Where , is the probability function of standard normal variate. The definite integral is known as normal probability integral and gives the area under standard normal variate between the ordinates Z=0. and Z = z1



Example: A mobile battery company says that an average battery lasts 1000 hours with a standard deviation of 100 hours. Assume that battery life is normally distributed. What is the probability that a randomly selected battery will last for 1200

Hours or less?

**Solution**: We are given that mean score is 1000 and standard deviation is 100.

Let X denote the battery life, here we want to find

 or or   
From the normal probability tables ,the cumulative probability is 0.977. Thus, there is a 97.7% probability that will last 1200 hours or less.

Importance of Normal Distribution

The normal distribution is important because it describes the statistical behavior of many real-world events. In fact, normal distribution plays a very important role in statistical theory. The shape of the normal distribution is completely described by the mean and the standard deviation. Thus, given the mean and standard deviation, we can use the properties of the normal distribution to quickly compute the cumulative probability for any value. It is important due to following main reasons.

(i) This distribution is important because of Central Limit theorem.  In simple terms, if we have many independent variables that may be generated by all kinds of distributions, the aggregate of those variables will tend toward a normal distribution.

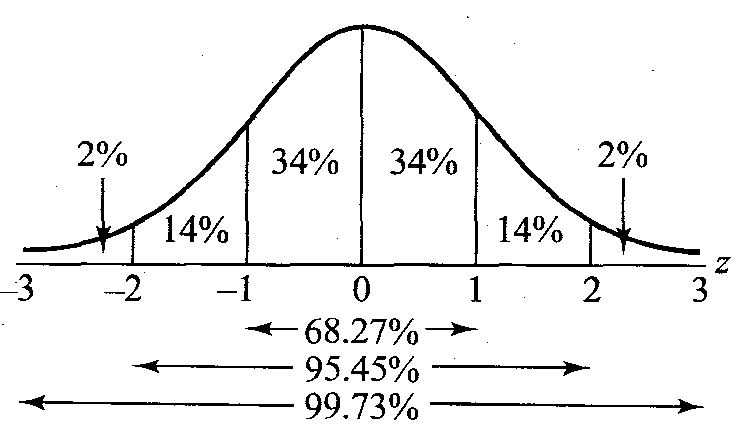
Most of the distributions occurring in practice, e.g., Binomial, Poisson, Hyper geometric distributions, etc., can be approximated by normal distribution. Moreover, many of the sampling distributions, e.g., Student’s t, Snedecor’s F, Chi-square distributions, etc., tend to normality for large samples.

(ii) Even if a variable is not normally distributed, it can sometimes be brought to normal form by simple transformation of variable.

(iii) The entire theory of small sample tests, viz., t, F,  tests, etc., is based on the fundamental assumption that the parent populations from which the samples have been drawn follow normal distribution.

(iv) Many of the distributions of sample statistics (e.g., the distributions of sample mean, sample variance, etc.) tend to normality for large samples and as such they can best be studied with the help of the normal curves.

(v) Normal distribution finds large applications in Statistical Quality Control in industry for setting control limits.



)

This property of the normal distribution forms the basis of entire Large Sample

(vii) Normal distribution is important is that many psychological and educational variables are distributed approximately normally.

**Example**: Assuming that mean height ofvolleyball players is 68.22 inches with standard deviation 3.286.A training program for volleyball players is attended by 500 players, how many players you expect to be over six feet of height among the players in the training camp.

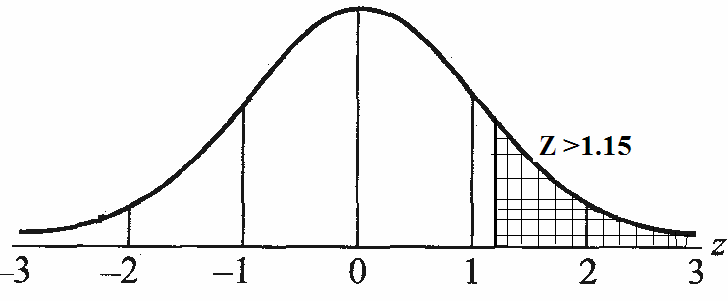
Solution: Let the height X of volleyball players follow normal distribution with mean and standard deviation . We are given to be  and 

Here we have to find the probability of players who are over 6 (72 inches) feet tall









Therefore total number of players in a batch of 500 whose height is more than 6 feet is given by



**Example**:The weight distribution of a group of 7,000 men is normal with mean weight 64.5” and s.d. 4.5”. Find the number of men whose weight is

(a) less than 69” but greater than 45.5”,

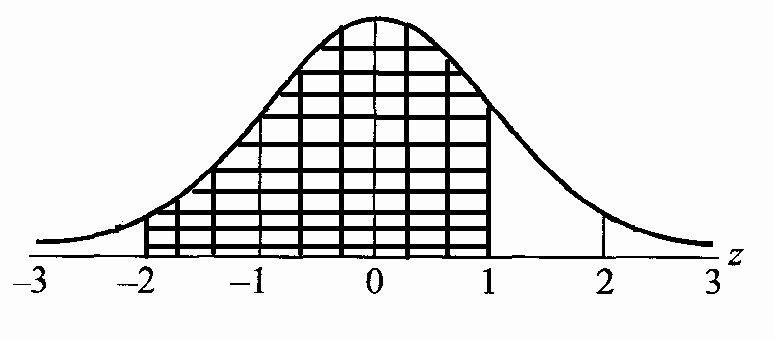
(b) less than 45.5”, and

(c) more than 75.5”.

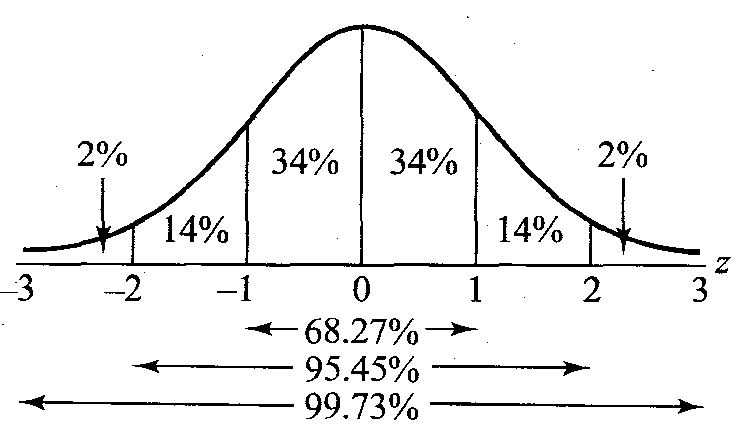
**Solution** The mean and standard deviation of the normal distribution are given to be  = 64.5” and 

(a) Percentage of men whose weight lies between 45.5” and 69”

Area under standard normal curve between the vertical lines at the corresponding standardized values, viz.



 and 



From the percentage distribution of area under the standard normal curve (Fig. above), it is found that the area between z = - 2 and z =1 is (14% ÷ 34% + 34%) = 82%. This means that 82% of the total number of 7,000 men are expected to have weights between 55.5” and 69”. Hence, the required number of men is 82% of 10,000



{b) Percentage of men whose weight is less than 55.5”

Area under standard normal curve to the left of the Standardized value



20% .The number of men is therefore, 20% of 7000 i.e. 1400.

(c) Percentage of men whose weight is more than 73.5”.= Area under standard normal curve to the right of the Standardized value



Hence, the number of men whose weight is more than 73.5” is 20% of 7,000. i.e. 1400.

**Example;** X is normal variate with mean 40 and S.D. 6. Find the probabilities

(i) (ii)(iii) 

**Solution:**

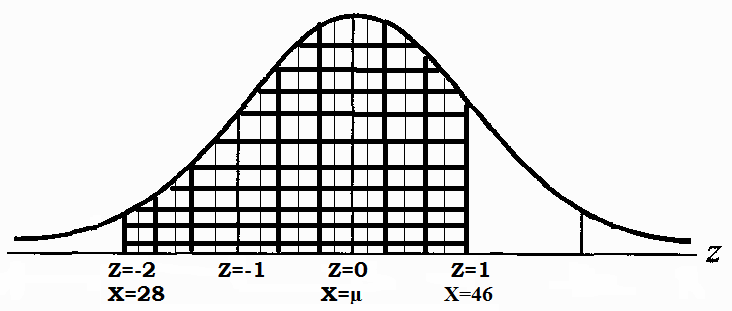
(i) When , and when X=46 ,



 ( By symmetry)

=0.4771+0.3413=0.8184

(From Normal Tables)



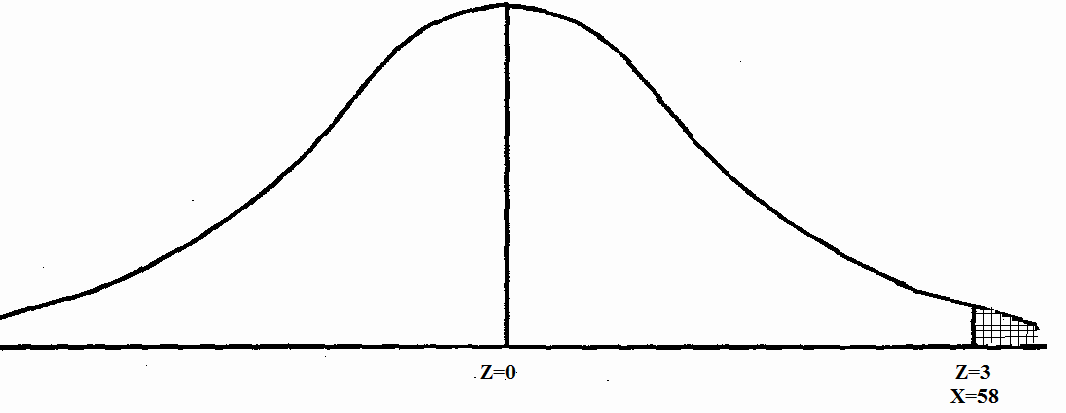
(ii)In order to compute first we convert X to Z

When X=58, 

Therefore,







(iii) 



 (By symmetry)

=0.4515+0.1293=0.5808 (From normal tables)



**Exponential Distribution**

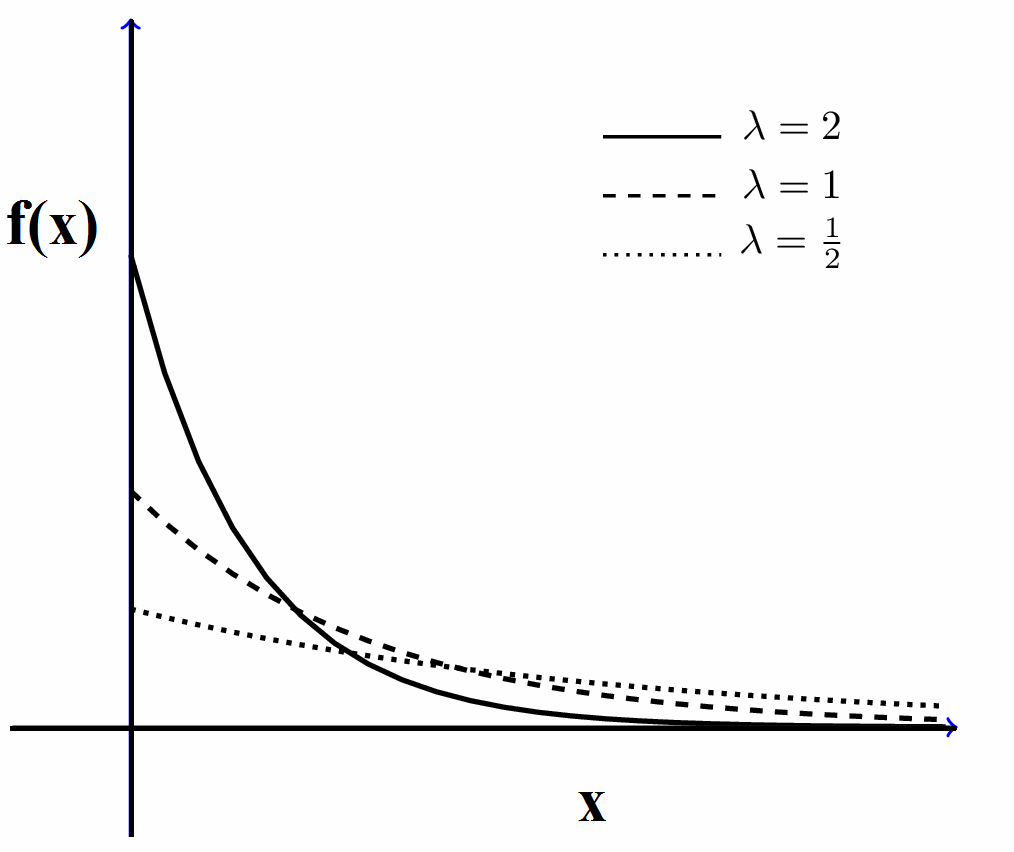
This distribution is one of the widely used continuous distributions. It is often used to model the time elapsed between events. The exponential distribution is that it can be viewed as a continuous analogue of the geometric distribution. The most important of these properties is that the exponential distribution is memoryless.

Definition: A continuous random variable X is said to follow Exponential distribution with parameter λ>0 is given by

****

It is denoted by X~Exponential (λ),

Figure given below shows the p.d.f of exponential distribution for different values of  λ.



The c.d.f of exponential distribution is given by:

****

Mean and variance of exponential distribution

Let X~Exponential(λ) than

****

Integrating by parts we get

****

****

****

Second moment about origin is given by

****

****

****

Now variance is given by

****

Memory less property of Exponential distribution

Suppose at time t=0 an alarm clock started which will ring after a time X that is exponentially distributed with rate ****. Let us consider X the lifetime of the clock. For any t > 0, we have

 (3.9)

Suppose that event has not happened before ‘S’ That is, we have observed the event {X > S}. If we let Y denote the remaining lifetime of the clock given that {X > S}, then

P(Y > t|X > S) = P(X > S + t|X > s)

= P(X > S + t,X > S)/P(X > S)

=P(X > S + t)/P(X > S)

 Using (3.9)

But this implies that the remaining lifetime after we observe the alarm has not yet gone off at time S has the same distribution as the original lifetime X. This implies that the distribution of the remaining lifetime does not depend on S. This property is called the *memoryless property* of the exponential distribution because we don’t need to remember when we started the clock. If the distribution of the lifetime X is Exponential(****), then if given that the clock has currently not yet gone off, We can forget the past and still know the distribution of the time from my current time to the time the alarm will go off.

Moment generating function

Let, moment generating function about origin is given by:







Hence

Mean=,****

**Remark**: (i) We can also find the moments about origin by using the relation



****

****



****

(ii) We see that ****

****

Therefore **,**

****

****

****

Hence, variance may be >,= or < than mean depending upon different values of ****

(ii) If are independent random variable following exponential

distribution with parameter ,respectively , then 

has an exponential distribution with parameter.

(iii) If are identically distributed following exponential distribution with parameter ,respectively , then is also exponentially distributed with parameter

UNIT-V

Gamma Distribution

Gamma distribution has special importance in probability statistics. This distribution is frequently used to model waiting times. For instance, in [life testing](https://en.wikipedia.org/wiki/Accelerated_life_testing), the waiting time until death is a [random variable](https://en.wikipedia.org/wiki/Random_variable) that is frequently modeled with a gamma distribution. In particular, the [arrival times](http://www.math.uah.edu/stat/poisson/Gamma.html) in the [Poisson process](http://www.math.uah.edu/stat/poisson/index.html) have gamma distributions, and the [chi-square distribution](http://www.math.uah.edu/stat/special/ChiSquare.html) in statistics is a special case of the gamma distribution. Also, the gamma distribution is widely used to model physical quantities that take positive value

Defintion:A random variable X is said to follow gamma distribution if its p.d.f is given by



X is known as a Gamma variate with parameter  and referred to as variate.

* The function f(x) defined above repress ents a probability function, since



Some important properties of Gamma Function

(i) 

(ii) 

(iii) 

(iv) 

M.G.F. of Gamma Distribution

Let, moment generating function about origin is given by:







Mean and variance of Gamma Distribution

If then its moment generating function is given by



Now first moment about origin is given by





Hence mean of gamma distribution is 

Second moment about origin of negative binomial distribution is given by





Now variance is given by





*Hence mean and variance of gamma distribution are equal.*

Limiting form of Gamma distribution

*“Gamma distribution tends to normal distribution as”.*

We know that if X ~ then E (X) = (say) and Var (X) = = , (say). Then standard gamma variate is given by: 

Now







Taking log both sides of the above equation we get





Where are the terms containing and higher powers of in the denominator



So that 

Which is the moment generating function of a standard normal variate. Hence by uniqueness theorem of m.g.f, standard gamma variate tends to standard normal variate as .

Additive property of gamma distribution

*Sum of independent Gamma variate is also a gamma variate .*If are independent Gamma variates with parameters  respectively then  is also a Gamma variate with parameters

Proof: If is , then 

Now the m.g.f. of sum  is given by





Which is m.g.f. of a Gamma variate with parameter . Hence the result follows by uniqueness theorem of m.g.f.

Exercise: Evaluate the following:

(i) 

(ii) 

Solution: (i)  





ii) We know that (\*)

Taking and in our problem and using (\*), we get





EXERCISE: Using the properties of the gamma function, show that show that for α>0,

λ>0



Solution: From property (iv) of Gamma function (refer,sec.3.1.1)



So that



beta distribution of First kind

In [probability theory](https://en.wikipedia.org/wiki/Probability_theory) the beta distribution is a family of continuous [probability distributions](https://en.wikipedia.org/wiki/Probability_distribution) defined on the interval [0, 1] [having](https://en.wikipedia.org/wiki/Parametrization) two parameters, denoted by , that appear as exponents of the random variable and control the shape of the distribution.

The beta distribution has been applied to model the behavior of [random variables](https://en.wikipedia.org/wiki/Random_variables) limited to intervals of finite length in a wide variety of disciplines. For example, it has been used as a statistical description of [allele frequencies](https://en.wikipedia.org/wiki/Allele_frequencies) in [population genetics](https://en.wikipedia.org/wiki/Population_genetics), time allocation in [project management](https://en.wikipedia.org/wiki/Project_management)/control systems,  variability of soil properties and heterogeneity in the probability of [HIV](https://en.wikipedia.org/wiki/HIV) transmission etc.

The beta distribution is a suitable model for the random behavior of percentages and proportions.

Definition: A random variable X is said to have a beta distribution of first kind with parameters  if its p.d.f is given by

 (3.1)

* Here random variable X is known as a beta variate of first kind with parameter  and referred to as .
* is the beta function.
* If we take in (3.1), we get



Which is p.d.f of uniform distribution on [0,1].

Mean and variance

The rth moment about origin is given by







 (3.2)

In particular ,for r=1 for (3.2) we get





Now





Variance is given by



.



beta distribution of second kind

Definition: A random variable X is said to have a beta distribution of second kind with parameters  if its p.d.f is given by

 (3.3)

* Here random variable X is known as a beta variate of second kind with parameter  and referred to as .
* is the beta function.

Mean and variance of beta distribution of second kind

The rth moment about origin is given by







 (3.4)

In particular ,for r=1 from (3.4) we get





Now





Variance is given by







Remarks:

1. If X and Y are impendent Gamma variates with parametersrespectively Let and ,then U and Z are independent and U is  and Z is variates respectively.

2. If X and Y are impendent Gamma variates with parameters respectively Let and ,then U and Z are independent and U is  and Z is variates respectively.

CHEBYSHEV’S INEQUALITY

This inequality guarantees that in any probability distribution, "nearly all" values are close to the mean — the precise statement being that no more than 1/k2 of the distribution's values can be more than k [standard deviations](https://en.wikipedia.org/wiki/Standard_deviations) away from the mean . The inequality has great utility because it can be applied to any probability distribution in which the mean and variance are defined.

In fact, the role of standard deviation as a parameter to characterize variance is precisely interpreted by means of this well-known *Chebychev’s inequality* discovered in 1853 was later on discussed in 1856 by *Bienayme*.

**Definition**: If X is a random variable with mean and variance, then for any number k, we have



Or  (3.5)

**Proof.** Case (i). X is a continuous random variable. Then by definition.,



 Where f(x) is p.d.f of X



 ….(\*)

We know that:

 and 

 ….(\*\*)

Substituting in (\*), we get

 [from (\*\*)]





 (\*\*\*)

Also since

  [using (\*\*\*)]

This establishes (3.5).

Case (ii). In case of discrete random variable, the proof follows exactly similarly on replacing integration by summation.

Generalised Form of Bienayme-Chebychev’s Inequality

Let g(X) be a non-negative function of a random variable X. Then for every k> 0, we have

 (3.6)

Proof: Here we shall prove the theorem for continuous random variable. The proof can be adapted to the case of discrete random variable on replacing integration by summation over the given range of the variable.

Let S be the set of all X, where 

i.e., 

 (3.7)

where F (x) is the distribution function of X.





If we take g(X) = { X -E (X)}2 = {X - }2 and replace k by k in (3.6), we get



Which is Chebychev’s inequality.

**Markov’s Inequality:** Taking g(X) = X in (3.6) {generalised form of Bienayme-Chebychev’s Inequality} we get, for any k> 0

 , which is Markov inequality.

CONVERGENCE IN PROBABLITY

Convergence in probability or stochastic convergence which is defined as follows:

A sequence of random variables X1, X2 ,…. Xn, ... is said to converge in probability to a constant a, if for any 



or its equilivant

 and we write 

If a sequence of constants , then regarding the constant as a r.v. having one-point distribution at that point, we can say that as 

Chebychev’s Theorem**:** As an immediate consequence of Chebychev’s inequality, we have the following theorem on convergence in probability.

is a sequence of random variables and if the mean and standard deviation  of exists for all n and if then 

Proof. By Chebychev’s inequality, for> 0,

Hence provided 

WEAK LAW OF LARGE NUMBERS (W.L.L.N.)

Statement: Let X1, X2, ..., Xn be a sequence of random variables and ,,.. , be their respective expectation and let

then ,

for all n > n0, where and are arbitrary small positive numbers, provided



Proof. Using Chebychev’s Inequality , to the r.v. (X1 + X2 + ... + Xn)/n, we get for any ,



Since 



So far, nothing is assumed about the behaviour of Bn for indefinitely increasing values of n. Since  is arbitrary, we assume , as n becomes indefinitely large.

Thus, having chosen two arbitrary small positive numbers  and , number n0 can be

found so that the inequality will hold for . Consequently, we shall have

, for all 

This conclusion leads to the following important result, known as the (Weak) Law of Large Numbers:

“With the probability approaching unity or certainty as near as we please, we may expect that the arithmetic mean of values actually assumed by n random variables will differ from the arithmetic mean of their expectations by less than any given number, however small, provided the number of variables can be taken sufficiently large and provided the condition:  is fulfilled

**For the existence of the law we assume the following conditions:**

(i) exists for all i,

(ii) exists and

(iii) 

Condition (i) is necessary, without it the law itself cannot be stated. But the conditions (ii) and (iii) are not necessary; (iii) is however a sufficient condition**.**

**WLLN for i.i.d. random variables**. If the variables X1, X2 ….Xn are independent and identically distributed,if and (say) for all i=1,2…n



the covariance terms vanish, since the variables are independent.

Hence 

Thus, the weak law of large number holds for the sequence {Xn} of i. i. d. r. v. ‘s and we get



It means that 

**Example**: Two unbiased dice are thrown. If X is the sum of the numbers shown up, prove that

 ,Compare this with the actual probability.

**Solution:** The probability distribution of the r.v. X (the sum of the numbers on the two dice) is as given in the following table

|  |  |  |
| --- | --- | --- |
| X | Favorable cases | Probability |
| 2  3  4  5  6  7  8  9  10  11  12 | (1,1)  (1,2) (2,1)  (1,3)(3,1)(2,2)  (1,4)(4,1)(2,3)(3,2)  (1,5)(5,1)(2,4)(4,2)(3,3)  (1,6)(6,1)(2,5)(5,2)(3,4)(4,3)  (2,6)(6,2)(3,5)(5,3)(4,4)  (3,6)(6,3)(4,5)(5,4)  (4,6)(6,4)(5,5)  (5,6)(6,5)  (6,6) | 1/36  2/36  3/36  4/36  5/36  6/36  5/36  4/36  3/36  2/36  1/36 |







So that variance of X



By Chebychev’s inequality, for k> 0, we have

Actual Probability





{Taking k=3}



**Example** If X is the number scored in a throw of a fair die, show that the Chebychev’s inequality gives 

where is the mean of X, while the actual probability is zero.

**Solution**. Here X is a random variable which takes the values 1, 2,3… 6, each with probability 1/6 . Hence







For k>0, Chebychev’s inequality gives



Choosing k=2.5 we get



**Example :** Examine whether the weak law of large numbers holds for the sequence (Xk) of independent random variables defined as follows:



Sol. We have









Now



Hence (Weak)Law of large numbers, holds for sequence of independent r.v.’s (Xk}.

**Example** The r.v.’s X1, X2, ..., Xn have equal expectations and finite variation. Is the weak law of large numbers applicable to this sequence if all the co-variances  are negative?

Sol. We have



if for all the convergences are negative



Hence WLLN holds.

**Example :** Let (Xk) be mutually independent and identically distributed random variables with mean and finite variance. If Sn = X1 + X2 + ... + Xn prove that the law of large numbers does not hold for the sequence {Sn}.

**Sol.** The variables now are S1. S2, ... Sn.









(Covariance terms vanish since variables are independent.)

Let for all i



Hence 

So that we cannot draw any conclusion whether WLLN holds or not.

**CENTRAL LIMIT THEOREM (C.LT.)**

This theorem was first stated by Laplace in 1812 and a rigorous proof under fairly general conditions was given by Liapounoff in 1901.

**Statement:** “If Xi, (i = 1, 2, ..., n) be independent random variables such that E(Xi) = and V(Xi) =, then under certain very general conditions, the random variable

Sn = X1 + X2 + ... + Xn is asymptotically normal with mean and standard deviation where

and 

In other words Sn =  said to satisfy C.L.T if 

**Jenson’s inequality on expectation.**

**Ans.** If is a continuous and convex function and X is a random variable having finite mean, i.e.,E (X) = , then

****

In case  is a continuous and concave function**,**

****