Type of Events

Complementary Event : The event E^C is called complementary event for the event E. It consists of all outcomes not in E, but in S. For example, in a dice throw, if E = {Even nos} = {2, 4, 6} then E^C = {Odd nos} = {1, 3, 5}.

Equally Likely Events :
Two events E and F are equally likely iff p(E) = p(F)

For example, E = {1, 2, 3}
F = {4, 5, 6}
are equally likely, since p(E) = p(F) = 1/2.

Mutually Exclusive Events :
Two events E and F are mutually exclusive, if $E \cap F = \phi$ i.e., $p(E \cap F) = 0$. In other words, if E occurs, F cannot occur and if F occurs, then E cannot occur (i.e., both cannot occur together).

Collectively Exhaustive Events :
Two events E and F are collectively exhaustive, if $E \cup F = S$. i.e., together E and F include all possible outcomes, $p(E \cup F) = p(S) = 1$.

Independent Events :
Two events E and F are independent iff

$$p(E \cap F = p(E) * p(F)$$

Also $p(E | F) = p(E)$ and $p(F | E) = p(F)$.

Whenever E and F are independent. i.e., when two events E and F are independent , the conditional probability becomes same as marginal probability i.e., probability E is not affected by whether F has happened or not, and vice-versa. i.e., when E is independent of F, then F is also independent of E.

DeMorgan's Law :

1. $\left(\bigcup_{i=1}^{n} E_i \right)^C = \bigcap_{i=1}^{n} E_i^C$

Example : $(E_1 \cup E_2)^C = E_1^C \cap E_2^C$

$(E_1 \cap E_2)^C = E_1^C \cup E_2^C$

Note that $E_1^C \cap E_2^C$ is the event neither E_1 nor E_2

$E_1 \cup E_2$ is the event either E_1 or E_2 (or both).

Demorgan's law is often used to find the probability of neither E_1 not E_2.

i.e., $p(E_1^C \cap E_2^C) = p(E_1 \cup E_2)^C] = 1 - p(E_1 \cup E_2)$

Approaches to Probability

There are 2 approaches to quantifying probability of an Event E.

1. Classical Approach :

$$P(E) = \frac{n(E)}{n(S)} = \frac{|E|}{|S|}$$

i.e., the ratio of number of ways an event can happen to the number of ways sample space can happen, is the probability of the event. Classical approach assumes that all outcomes are equally likely.

Example 1 :

If out all possible jumbles of the word "BIRD", a random word is picked, what is the probability, that this word will start with a "B".

Solution :

$$p(E) = \frac{n(E)}{n(S)}$$

In this problem : n(S) = all possible jumbles of BIRD = 4!
n(E) = those jumbles starting with "B" = 3!

So, $$p(E) = \frac{n(E)}{n(S)} = \frac{3!}{4!} = \frac{1}{4}$$

Example 2 :

From the following table find the probability of obtaining "A" grade in this exam.

Grade	A	B	C	D
Number of Students	10	20	30	40

Solution :

N = total number of students = 100

By frequency approach,

$$p(\text{A grade}) = \frac{n(\text{A grade})}{N} = \frac{10}{100} = 0.1$$

Axioms of Probability

Consider an experiment whose sample space is S. For each event E of the sample space S we assume that a number P(E) is defined and satisfies the following three axioms.

Axiom - 1 : $0 \leq P(E) \leq 1$

Axiom - 2 : $P(S) = 1$

Axiom - 3 : For any sequence of mutually exclusive events $E_1, E_2, \ldots\ldots$ (that is events for which $E \cap E = \phi$ when $i \neq j$.

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

Example : $P(E_1 \cup E_2) = P(E_1) + P(E_2)$ where (E_1, E_2 are mutually exclusive).

Rules of Probability

There are six rules of probability using which probability of any compound event involving arbitrary event A and B, can be computed.

Rule 1 :

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

This rule is also called the inclusion-exclusion principle of probability.

This formula reduces to $p(A \cup B) = p(A) + p(B)$

If A and B are mutually exclusive, since $p(A \cap B) = 0$ in such a case.

Rule 2 :

$$p(A \cap B) = p(A) * p(B/A) = p(B) * p(A/B)$$

where p(A/B) represents the conditional probability of A given B and p(B/A) represents the condition probability of B given A.

(a) p(A) and p(B) are called the marginal probabilities of A and B respectively. This rule is also called the multiplication rule of probability.

(b) $p(A \cap B)$ is called the joint probability of A and B.

(c) If A and B are independent events, this formula reduces to $p(A \cap B) = p(A) * p(B)$
since when A and B are independent
$p(A/B) = p(A)$ and $p(B/A) = p(B)$
i.e., the conditional probabilities become same as the marginal (unconditional) probabilities.

(d) If A and B are independent, then so are A and B^C ; A^C and B and A^C and B^C.

(e) Condition for three events to independent :
Events A, B and C are independent iff

$$p(ABC) = p(A)\,p(B)\,p(C)$$

and $p(AB) = p(A)\,p(B)$

and $p(AC) = p(A)\,p(C)$

A, B, C arepairwise independent.

and $p(BC) = p(B)\,p(C)$

Note : If A, B, C are independent, then A will be independent of any event formed from B and C.

For instance, A is independent of $B \cup C$.

Rule 3 : Complementary Probability.

$$p(A) = 1 - p(A^C)$$

$p(A^C)$ is called the complementary probability of A and $p(A^C)$ represents the probability that the event A was not happen.

$$\therefore \quad p(A) = 1 - p(A^C)$$

$p(A^C)$ is also written as $p(A')$

Notice that $p(A) + p(A') = 1$

i.e., A and A′ are mutually exclusion as well as collectively exhaustive.

Also notice that by Demorgan's law since $A^C \cap B^C = (A \cup B)^C$

$$p(A^C \cap B^C) = p(A \cup B)^C = 1 - p(A \cup B)$$

i.e., p(neither A nor B) = 1 - p (either A or B)

Rule 4 : Conditional Probability Rule

Starting from the multiplication rule.

$$p(A \cap B) = p(B) * p(A/B)$$

by cross multiplying we get the conditional probability formula.

$$p(A/B) = \frac{p(A \cap B)}{p(B)}$$

By interchanging A and B in this formula we get

$$p(B/A) = \frac{p(A \cap B)}{p(A)}$$

Probability Distributions

Random Variables

It is frequently the case when an experiment is performed that we are mainly interested in some function of the outcome as opposed to the actual outcome itself.

For instances, in tossing dice we are often interested in the sum of two dice and are not really concerned about the separate value of each die. That is, we may be interested in knowing that the sum is 7 and not be concerned over whether the actual outcome was (1, 6) or (2, 5) or (3, 4) or (4, 3) or (5, 2) or (6, 1).

Also, in coin flipping we may be interested in the total number of heads that occur and not care at all about the actual head tail sequence that results. These quantities of interest, or more formally, these real valued functions defined on the sample space, are known as random variables.

Because the value of a random variable is determined by the outcome of the experiment, we may assign probabilities to the possible values of the random variable.

Types of Random Variable : Random variable may be discrete or continuous.

Discrete Random Variable : A variable that can take one value from a discrete set of values.

Example : Let x denotes sum of 2 dice. Now x is a discrete random variable as it can take one value from the set {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}, since the sum of 2 dice can only be one of these values.

Continuous Random Variable : A variable that can take one value form a continuous range of values.

Example : x denotes the volume of Pepsi in a 500 ml cup. Now x may be a number from 0 to 500, any of which value, x may take.

Probability Density Function (PDF)

Let x be continuous random variable then its PDF F(x) is defined such that

1. $R(x) \geq 0$ 2. $\int_{-\infty}^{\infty} F(x)dx = 1$

3. $P(a < x < b) = \int_{a}^{b} F(x)\, dx$

Probability Mass Function (PMF)

Let x be discrete random variable then its PMF p(x) is defined such that

1. $p(x) = P[X = x]$ 2. $p(x) \geq 0$ 3. $\sum p(x) = 1$

Distributions

Based on this we can divide distributions also into **discrete distribution** (based ob a discrete random variable) or continuous distribution (based on a continous random variable).

Examples of discrete distribution are binomal, Poisson and hypergeometric distributions.

Examples of continuous distribution are uniform, normal and exponential distributions.

Properties of Discrete Distribution

$$\sum P(x) = 1$$
$$E(x) = \sum x\, P(x)$$
$$V(x) = E(x^2) - (E(x))^2 = \sum x^2\, P(x) - [\sum x\, P(x)]^2$$

E(x) denotes expected value or average value of the random variable x, while V(x) denotes the variance of the random variable x.

Properties of Continous Distribution

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

$$F(x) = \int_{-\infty}^{\infty} f(x)dx \text{ (cumulative distribution function)}$$

$$E(x) = \int_{-\infty}^{\infty} xf(x)\,dx$$

$$V(x) = E(x^2) - [E(x)]^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \left[\int_{-\infty}^{\infty} xf(x)dx\right]^2$$

$$p(a < x < b) = p(a \le x < b) = P(a < x \le b) = P(a \le x \le b)$$

$$= \int_a^b f(x)dx$$

Types of Distributions

Discrete Distributions :

1. General Discrete Distribution
2. Binomial Distribution
3. Hypergeometric Distribution
4. Geometric Distribution
5. Poisson Distribution

General Discrete Distribution

Let X be a discrete random variable.

A table of possible values of x versus corresponding probability values p(x) is called as its probability distribution table.

Expectation E(x)

The mean value of the probability distributiobn of a variety is commonly known as its expectation.

$$\mu_X \ E(X) = \sum_{i=1}^{n} x_i \, x_i \, f(x_i \ x_i \ x_i) \qquad \text{(Discrete case)}$$

$$\mu_X \ E(X) = \int_{-\infty}^{\infty} x_i \, x_i \, f(x_i \ x_i \ x_i) \, dxdx \quad \text{(Continuous case)}$$

Variance Var(X)

$$\text{Var } X = E[(x - \mu\mu_x)^2]$$

$$\text{Var } X = \sum(x_i\, x - \mu\mu)^2 f(xx_i) \qquad \text{(Discrete case)}$$

$$\text{Var } X = \int_{-\infty}^{\infty}(xx - \mu\mu)^2 f(x)dx \qquad \text{(Continuous case)}$$

It can be proved that $\text{Var } X = E(XX^2) - [E(X)]^2$

Properties of Expectation and Variance.

If x_1 and x_2 are two random variance and a and b are constants,

$$E(ax_1 + b) = a\,E(x_1) + b$$

$$V(ax_1 + b) = a^2 V(x_1)$$

$$E(ax_1 + bx_2) = a\,E(x_1) + b\,E(x_2)$$

$$V(ax_1 + bx_2) = a^2\,V(x_1) + b^2 V(x_2) + 2ab\,\text{cov}(x_1, x_2)$$

Where cov (x_1, x_2) represents the covariance between x_1 and x_2

If x_1 and x_2 are independent, then $\text{cov}(x_1, x_2) = 0$ and the above formula reduces to

$$V(ax_1 + bx_2) = a^2 V(x_1) + b^2 V(x_2)$$

For example, from above formula we can say

$$E(x_1 + x_2) = E(x_1) + E(x_2)$$

$$E(x_1 - x_2) = E(x_1) - E(x_2)$$

$$V(x_1 + x_2) = V(x_1 - x_2) = V(x_1) + V(x_2)$$

Formula for calculating covariance between X and Y

$$\text{Cov }(x, Y) = E(XY) - E(X)\,E(Y)$$

$\therefore$ If X, Y are independent $E(XY) = E(X)\,E(Y)$

and hence $\text{Cov}(X, Y) = 0$

Binomial Distribution

The probability of obtaining x-successes from n trials is given by the binomial distribution formula.

$$P(X = x) = nC_x\; p^x(1 - p)^{n-x}$$

Where p is the probability of success in any trial and (1 - p) = q is the probability of failure.

Geometric Distribution

Consider repeated trials of a Bernoulli experiment $\in$ with probability P of success and q = 1 - P of fail.
Let x denote the number of times $\in$ must be repeated unit finally obtaining a success. The distribution of random variable x is given as follows.

k	1	2	3	4	5
P(k)	P	qP	q^2P	q^3P	q^4P

The experiment $\in$ will be repeated k times only in the case that there is a sequence of k - 1 failures follow by a success.

$$P(k) = P(x - k) = q \quad P$$

The geometric distribution is characterized by a single parameter P.

Points to Remember :

Let x be a geometric random variable with distribution GEO(P). Then

1. $E(x) = \frac{1}{P}$
2. $Var(x) = \frac{q}{P^2}$
3. Cumulative distribution $F(k) = 1 - q^k$
4. $P(x > r) = q^r$

Geometric distribution possesses "no-memory" or "lack of memory" property which can be stated as

$$P(x > a + r \mid x > a) = P(x > r)$$

Poisson Distribution

A random variable X, taking on one of the values 0, 1, 2, ... is said to be a Poisson random variable with parameter λ if for some $\lambda > 0$.

ENTRI

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

For Poisson Distribution :

$$\text{Mean} = E(x) = \lambda$$

$$\text{Variance} = V(x) = \lambda$$

Therefore, expected value and variance of a Poisson random variable are both equal to its parameters λ.

Here λ is average number of occurrences of event in an observation period Δt. So, $\lambda = \alpha\Delta t$ where α is number of occurrences of event per unit time.

PROBABILITY

Prob = $\sum P(x) = 1 \rightarrow$ Discrete

$\int P(x) = 1$

Distribution

(i) Binomial Distribution **(ii) Poisson Distribution**

Binomial Distribution = $^nC_r \, p^r q^{n-r}$

$n \rightarrow$ lot number of variable

$r \rightarrow$ event $\quad p + q = 1 \quad q = 1 - p$

* Position cannot be found
* To find only 2 success or 2 failure is other words specific number of success.

Mean or expectation, $\mu = n p$

Variance $= npq$

Standard deviation $= \sqrt{npq}$

Poisson Distribution

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

$e \rightarrow 0, 1, 2, ...$ random variable

$\lambda \rightarrow$ mean $= xp$

Variance $= xp$.

$SD = \sqrt{xp}$

Continuous

i) $\int_{-\infty}^{\infty} p(x) = 1 \qquad E(x) = \int n\, p(n)$

$$Var(x) = \int x^2 p(x) - [\int x\, p(x)]^2$$

Properties

- $E(\text{constant} = \text{constant}$
- $E(ax + by) = a\,E(x)\,b\,E(y)$
- $E(ax - by) = a\,E(x) - b\,E(y)$
- $E(xy) = E(x)\,.\,E(y)$ If x & y are independent
- Variance (constant) $= 0$
- $v(ax \pm by) = a^2\,v(n) + b^2\,v(y)$
- Co-Variance $(xy) = E(x\,.\,y) - E(x)\,.\,E(y)$
- If x & y are independent, the covariance $(x\,.\,y) = 0$

Uniform Distribution

$[a, b]$ $\qquad p(x) = \frac{1}{b - a}$ $\qquad \int_a^b \frac{1}{b - a} = 1$

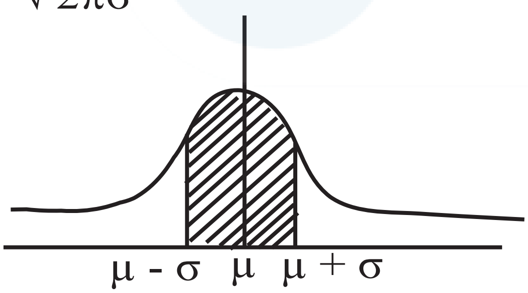
$\text{Mean} = \frac{a + b}{2}$ $\qquad \text{Variance} = \frac{(b - a)^2}{12}$

Normal Distribution or Gaussian

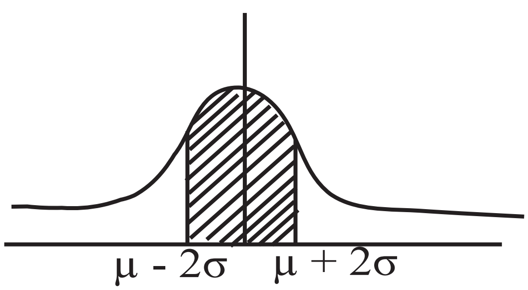
$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{\frac{-(x - \mu)^2}{2\sigma^2}}$$

$\mu \longrightarrow$ mean

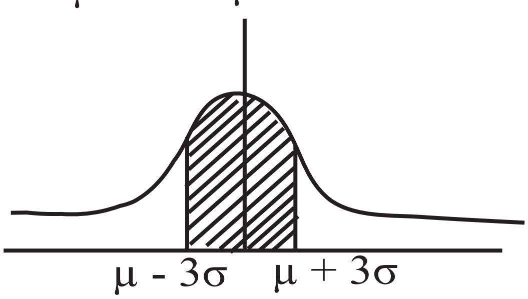
$\sigma \longrightarrow$ SD



$\mu - \sigma \longrightarrow \mu + \sigma \longrightarrow 68.34\%$ or 0.6834



$\mu - 2\sigma \longrightarrow \mu + 2\sigma \longrightarrow 95.5\%$ or 0.955



$\mu - 3\sigma \longrightarrow \mu + 3\sigma \longrightarrow 99.7\%$ or 0.997

Exponential Distribution

$$P(x) = \lambda e^{-\lambda n} \quad x \geq 0$$
$$= 0 \quad x < 0$$

$$\text{Mean} = \frac{1}{\lambda} \qquad \text{Var} = \frac{1}{\lambda^2}$$

Standard Normal Distribution

$$\mu = 0 \qquad \sigma = 1$$

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

Arithmetic Mean for Raw Data

The formula for calculating the arithmetic mean for raw data is $\overline{x} = \frac{\sum x}{n}$

$\overline{x}$: Arithmetic mean

x : Refers to the value an observation

n : Number of observations

Example :

The number of visits made by ten mothers to a clinic were ; 8 6 5 5 7 4 5 9 7 4

Calculate the average number of visits.

The Arithmetic Mean for Grouped Data (Frequency Distribution)

The formula for the arithmetic mean calculated from a frequency distribution has to be amended to include the frequency. It becomes

$$\overline{x} = \frac{\sum(fx)}{\sum f}$$

Example :

To show how we can calculate the arithmetic mean of a grouped frequency distribution, there is a example of weights of 75 pigs. The classes and frequencies as given in following table :

Weight (kg)	**Midpoint of class x**	**Number of pigs f(frequency)**	**fx**
0 & under 20	15	1	15
20 & under 30	25	7	175
30 & under 40	35	8	280
40 & under 50	45	11	495
50 & under 60	55	19	1045
60 & under 70	65	10	650
70 & under 80	75	7	525
80 & under 90	85	5	425
90 & under 100	95	4	380
100 & under 110	105	3	215
Total		**75**	**4305**

Median for Raw Data :

In general, if we have n values of x, they can be arranged in ascending order as : $x_1 < x_2 < \ldots\ldots < x_n$

Suppose n is odd, then Median $=$ the $\frac{(n+1)}{2}$ -th value

However, if n is even, we have two middle points

$$\text{Median} = \frac{\left(\frac{n}{2}\right)^{th} \text{value} + \left(\frac{n}{2}+1\right)^{th} \text{value}}{2}$$

Example :

The heights (in cm) of six students in class are 160, 157, 156, 161, 159, 162. What is median height ?

Median for Grouped Data

1. Identify the median class which contains the middle observation $\left(\frac{N+1}{2}\right)^{th}$ observation. This can be done by observing the first class in which the cumulation frequency is equal or more than $\frac{N+1}{2}$. Here, $N = \Sigma f =$ Total number of observations.
2. Calculate Median as follows :

$$\text{Median} = L + \left[\frac{\left(\frac{N+1}{2}\right) - (F+1)}{f_m}\right] \times h.$$

Where, l = Lower limit of median class

n = Total number of observations

cf = Cumulative frequency of the class preceding the median class.

f = Frequency of median class

C = class length

Example :

Consider the following table giving the marks obtained by students in an exam.

Mark Range	f No. of Students	Cumulative Frequency
0 - 20	2	2
20 - 40	3	5
40 - 60	10	15
60 - 80	15	30
80 - 100	20	50

Mode

Mode is defined as the value of the variable which occurs most frequently.

Mode for Raw Data

In raw data, the most frequently occurring observation is the mode. That is data with highest frequency mode. If there is more than one data with highest frequency, then each of them is a mode. Thus we have Unimodal(single mode), Bimodal (two modes) and Trimodal (three modes data sets.

Example :

Find the mode of the data set : 50, 50, 70, 50, 50, 70, 60.

Mode for Grouped Data

Mode is that value of x for which the frequency is maximum. If the values of x are grouped into the classes (such that they are uniformly distributed within any class) and we have a frequency distribution then :

1. Identify the class which has the largest frequency (modal class)
2. Calculate the mode as

$$\text{Mode} = L + \frac{f_0 - f_1}{2f_0 - f_1 - f_2} \times h$$

Where,

L = Lower limit of the modal class

f_0 = Largest frequency (frequency of Modal Class)

f_1 = Frequency in the class preceding the modal class.

f_2 = Frequency in the class next to the modal class

h = Width of the modal class

Example :

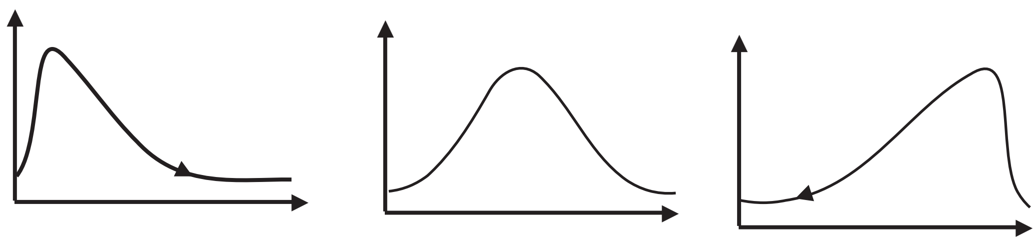
Data relating to the height of 352 school students are given in the following frequency distribution.

Calculate the modal height.

Height (in feet)	**Number of students**
3.0 - 3.5	12
3.5 - 4.0	37
4.0 - 4.5	79
4.5 - 5.0	152
5.0 - 5.5	65
5.5 - 6.0	7
Total	**352**

Properties Relating Mean, Median and Mode

1. Empirical mode = 3 median - 2 mean
   When an approximate value of mode is required above empirical formula for mode may be used.
2. There are three types of frequency of distributions. Positively skewed, symmetric and negatively skewed distribution.



(a) (Positively Skewed) (a) Symmetric (a) Negatively Skewed

(a) In positively skewed distribution.

Mode $\leq$ Median $\leq$ Mean

(b) In symmetric distribution

Mean = Median = Mode

(c) In negatively skewed distribution

Mean $\leq$ Median $\leq$ Mode

Standard Deviation

Standard Deviation is a measure of disperson or variation amongst data.

Insteady of taking absolute deviation from the arithmetic mean. We may square each deviation and obtain the arithmetic mean of squared deviations. This gives us the variance of the values.

The positive square root of the variance is called the **"Standard Deviation"** of the given values.

Standard Deviation for Raw Data

Suppose $x_1, x , ... x_n$ are n values of the x, their arithmetic mean is :

$\bar{x} = \frac{1}{N}\Sigma x_i$ and $x_1 - \bar{x}, x_2 - \bar{x}, ... x_n - \bar{x}$ are the deviations

of the values of x from $\bar{x}$. Then,

$\sigma^2 = \frac{1}{n^2}\Sigma(x_i - \bar{x})^2$ is the variance of x. It can be shown that

$$\sigma^2 = \frac{\Sigma(x_i - \bar{x})^2}{n} = \frac{1}{n}\Sigma x_i^2 - \bar{x}^2 = \frac{n\Sigma x_i^2 - (\Sigma x_i)^2}{n^2}$$

It is conventional to represents the variance by the symbol σ^2. Infact, σ is small sigma and Σ is capital sigma. Square root of the variance is the standard deviation.

$$\sigma = +\sqrt{\frac{1}{n}\Sigma(x_i - \bar{x})^2} = \sqrt{\frac{1}{n}\Sigma x_i^2 - \bar{x}^2} = \sqrt{\frac{n\Sigma x_i^2 - (\Sigma x_i)^2}{n^2}}$$

Example :

Consider three students in a class, and their marks in exam was 50, 60 and 70. What is the standard deviation of this data set ?

Example :

The frequency distribution for heights of 150 young ladies in a beauty contest is given below for which we have to calculate standard deviation.