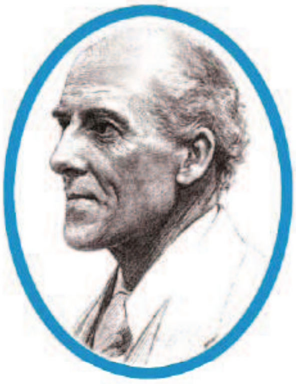
પ્રકરણ 15

આંકડાશાસ્ત્ર

❖ *"Statistics may be rightly called the science of averages and their estimates." – A. L. BOWLEY and A. L. BODDINGTON* ❖

15.1 પ્રાસ્તાવિક



Karl Pearson
(1857-1936)

આપણે જાણીએ છીએ કે આંકડાશાસ્ત્રનો વ્યવહાર કોઈ વિશેષ હેતુને લઈને એકત્રિત કરેલી માહિતી સાથે છે. આપણે માહિતીનું વિશ્લેષણ અને અર્થઘટન કરીને તેમના વિશે નિર્ણય લઈએ છીએ. આપણે આગળનાં ધોરણોમાં માહિતીને આલેખ અને કોષ્ટક સ્વરૂપમાં દર્શાવવાની રીતોનો અભ્યાસ કર્યો છે. આ નિરુપણ માહિતીનાં મહત્ત્વપૂર્ણ લક્ષણો અથવા વિશેષતાઓને દર્શાવે છે. આપણે આપેલ માહિતીનું પ્રતિનિધિત્વ રજૂ કરતાં મૂલ્યો શોધવાની રીતો વિશે અભ્યાસ કર્યો છે. આ મૂલ્યોને મધ્યવર્તી સ્થિતિમાનનાં માપ કહે છે. યાદ કરો કે ***મધ્યક*** *(સમાંતર મધ્યક),* ***મધ્યસ્થ*** *અને* ***બહુલક*** *(mean, median and mode)* એ મધ્યવર્તી સ્થિતિમાનનાં ત્રણ માપ છે. મધ્યવર્તી સ્થિતિમાનનું માપ આપણને એ વાતનો આભાસી ખ્યાલ આપે છે કે માહિતી ક્યાં કેન્દ્રિત થઈ છે. પરંતુ માહિતી પરથી વધુ સચોટ અર્થઘટન કરવા માટે, આપણને એ ખ્યાલ પણ હોવો જોઈએ કે પ્રાપ્તાંકો(માહિતી) કેટલા વિખેરાયેલા છે અથવા તો મધ્યવર્તી સ્થિતિમાનનાં માપની ચારે તરફ કઈ રીતે એકત્રિત થયેલા છે.

બે બૅટ્સમૅનો દ્વારા છેલ્લી દશ મેચમાં બનાવેલા રન પર વિચાર કરીએ.

બૅટ્સમૅન A : 30, 91, 0, 64, 42, 80, 30, 5, 117, 71

બૅટ્સમૅન B : 53, 46, 48, 50, 53, 53, 58, 60, 57, 52

સ્પષ્ટપણે માહિતીનો મધ્યક અને મધ્યસ્થ દર્શાવેલ છે :

	બૅટ્સમૅન A	બૅટ્સમૅન B
મધ્યક	53	53
મધ્યસ્થ	53	53

યાદ કરો કે આપણે માહિતીનો મધ્યક ($\bar{x}$ વડે દર્શાવીએ છીએ) અવલોકનોના સરવાળાને અવલોકનોની કુલ સંખ્યા વડે ભાગીને મેળવીએ છીએ. એટલે કે,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

મધ્યસ્થની ગણતરી માટે પ્રાપ્તાંકો પહેલાં ચઢતા કે ઊતરતા ક્રમમાં ગોઠવવામાં આવે છે અને પછી નીચે દર્શાવેલ નિયમનો ઉપયોગ કરવામાં આવે છે.

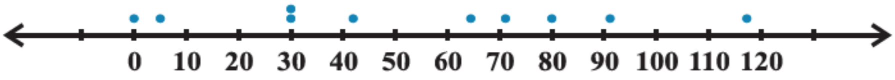
જો આપેલાં અવલોકનોની સંખ્યા અયુગ્મ હોય, તો મધ્યસ્થ એ $\left(\frac{n+1}{2}\right)$ મું અવલોકન છે.

જો અવલોકનોની સંખ્યા યુગ્મ હોય તો મધ્યસ્થ $\left(\frac{n}{2}\right)$ માં અને $\left(\frac{n}{2}+1\right)$ માં અવલોકનોની સરેરાશ છે.

આપણે જોઈ શકીએ છીએ કે બંને ખેલાડી A અને B દ્વારા બનાવેલા રનનો મધ્યક અને મધ્યસ્થ સરખા છે અને તે 53 છે. શું આપણે કહી શકીએ કે બન્ને ખેલાડીઓનું પ્રદર્શન સમાન છે ? સ્પષ્ટ છે કે નથી જ. કારણ કે A ના રનમાં ચલન 0 (ન્યૂનતમ) થી 117 (મહત્તમ) સુધી છે, જ્યારે B ના રનનો વિસ્તાર 46 થી 60 સુધી છે.

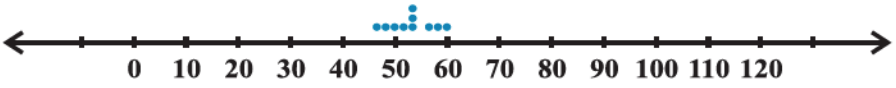
ચાલો, હવે ઉપર્યુક્ત રનની સંખ્યાઓને એક સંખ્યારેખા પર દર્શાવીએ. આપણને નીચે દર્શાવેલ આકૃતિઓ મળે છે :

બૅટ્સમૅન A માટે



આકૃતિ 15.1

બૅટ્સમૅન B માટે



આકૃતિ 15.2

આપણે જોઈ શકીએ છીએ કે બૅટ્સમૅન B ને અનુરૂપ બિંદુઓ એકબીજાની નજીક નજીક છે અને મધ્યવર્તી સ્થિતિમાનનાં માપ (મધ્યક અને મધ્યસ્થ) ની આસપાસ એકત્રિત થાય છે, જ્યારે બૅટ્સમૅન A ને અનુરૂપ બિંદુઓ ફેલાયેલાં છે અથવા વધુ વિખેરાયેલાં છે.

આમ આપેલ માહિતી વિશે સંપૂર્ણ જાણકારી આપવા માટે મધ્યવર્તી સ્થિતિમાનનાં માપ એકલાં પર્યાપ્ત નથી. જેનો અભ્યાસ આંકડાશાસ્ત્રના અંતર્ગત કરવો જોઈએ તેવું એક અન્ય પરિબળ પરિવર્તનશીલતા છે.

મધ્યવર્તી સ્થિતિમાનના માપની જેમ જ પરિવર્તનશીલતાના વર્ણન માટે પણ એક સંખ્યા જરૂરી છે. તે સંખ્યાને ***પ્રસારનું માપ*** (*measure of dispersion*) કહે છે. આ પ્રકરણમાં આપણે પ્રસારનાં માપનું મહત્ત્વ અને તેમની વર્ગીકૃત અને અવર્ગીકૃત માહિતી માટે ગણતરીની રીતો વિશે અભ્યાસ કરીશું.

15.2 પ્રસારનાં માપ

સંખ્યાઓમાં પ્રસારનું માપ અવલોકનો અને ત્યાં ઉપયોગમાં લેવાયેલ મધ્યવર્તી સ્થિતિમાનનાં માપના આધારે કરવામાં આવે છે. પ્રસારનાં માપ નીચે દર્શાવ્યા છે :

(i) *વિસ્તાર* (*Range*) (ii) *ચતુર્થક વિચલન* (*Quartile deviation*) (iii) *સરેરાશ વિચલન* (*Mean deviation*) (iv) *પ્રમાણિત વિચલન* (*Standard deviation*).

આ પ્રકરણમાં આપણે ચતુર્થક વિચલન સિવાયના અન્ય તમામ માપોનો અભ્યાસ કરીશું.

15.3 વિસ્તાર (Range)

યાદ કરો કે બે બૅટ્સમૅન A અને B દ્વારા બનાવેલા રનના ઉદાહરણમાં આપણને પ્રત્યેક શ્રેણીના મહત્તમ અને ન્યૂનતમ રનના આધાર પરથી રનની સંખ્યાઓમાં પરિવર્તનશીલતાનો ખ્યાલ આવે છે. આમાં એકલ સંખ્યા જાણવા માટે આપણે શ્રેણીની મહત્તમ અને ન્યૂનતમ સંખ્યાઓ વચ્ચેનો તફાવત(અંતર) મેળવીએ છીએ. આ તફાવતને *વિસ્તાર* કહેવામાં આવે છે.

બૅટ્સમૅન A નો વિસ્તાર $= 117 - 0 = 117$ અને બૅટ્સમૅન B નો વિસ્તાર$= 60 - 46 = 14$.

સ્પષ્ટ છે કે A નો વિસ્તાર > B નો વિસ્તાર. તેથી A ના રનની સંખ્યાઓમાં વિચલન અથવા પ્રસાર વધુ છે, પરંતુ B ના રનની સંખ્યાઓ એકબીજાની વધુ નજીક છે.

આમ, એક શ્રેણીનો વિસ્તાર =પ્રાપ્તાંકોનું મહત્તમ મૂલ્ય – પ્રાપ્તાંકોનું ન્યૂનતમ મૂલ્ય

માહિતીનો વિસ્તાર આપણને વિખેરાવ અથવા ચલનીયતાનો સ્થૂળ ખ્યાલ આપે છે, પરંતુ મધ્યવર્તી સ્થિતિનું માપ *માહિતીના પ્રસાર (dispersion)* વિશે કશું જ જણાવતું નથી. આ હેતુ માટે આપણને પરિવર્તનશીલતાનાં બીજાં કેટલાંક માપોની પણ જરૂર પડે છે. સ્પષ્ટ છે કે આ પ્રકારનાં માપ અવલોકનોના મધ્યવર્તી સ્થિતિમાનથી અંતર (અથવા વિચલન) પર આધારિત હોવા જોઈએ.

મધ્યવર્તી સ્થિતિમાનથી અવલોકનોના અંતરના આધાર પર શોધવામાં આવેલ પ્રસારનાં મહત્ત્વપૂર્ણ માપ એ સરેરાશ વિચલન અને પ્રમાણિત વિચલન છે. ચાલો આના ઉપર વિસ્તૃત ચર્ચા કરીએ.

15.4 સરેરાશ વિચલન (Mean Deviation)

યાદ કરો કે અવલોકન x નું અચળ મૂલ્ય 'a' થી અંતર $(x - a)$ એ અવલોકન x નું a થી વિચલન કહેવાય છે. 'x' ની કિંમતોનો મધ્યવર્તી કિંમત 'a' થી *પ્રસાર* શોધવા માટે આપણે 'a' થી વિચલનો શોધીએ છીએ. આ વિચલનોનો મધ્યક એ પ્રસારનું નિરપેક્ષ માપ હોય છે. મધ્યક શોધવા માટે આપણે વિચલનોનો સરવાળો મેળવીએ છીએ, પરંતુ આપણે જાણીએ છીએ કે, મધ્યવર્તી સ્થિતિમાનનું માપ એ અવલોકનોના ગણની મહત્તમ અને ન્યૂનતમ કિંમતોની મધ્યમાં હોય છે. તેથી કેટલાંક વિચલન ઋણ તથા કેટલાંક ધન હશે. આમ, વિચલનોનો સરવાળો શૂન્ય હોઈ શકે છે. આ ઉપરાંત મધ્યક ($\bar{x}$) થી વિચલનોનો સરવાળો શૂન્ય હોય છે જ. આ સાથે જ

$$\text{વિચલનોનો મધ્યક} = \frac{\text{મધ્યકથી વિચલનોનો સરવાળો}}{\text{અવલોકનોની સંખ્યા}} = \frac{0}{n} = 0$$

આમ, જ્યાં સુધી પ્રસારના માપને લાગેવળગે છે, મધ્યકની સાપેક્ષ વિચલનોનો મધ્યક શોધવાનું કોઈ ઔચિત્ય રહેતું નથી.

યાદ કરો કે પ્રસારનું યોગ્ય માપ શોધવા માટે આપણને પ્રત્યેક મૂલ્યના મધ્યવર્તી સ્થિતિમાનનું માપ અથવા કોઈ અચલ સંખ્યા 'a' થી અંતર મેળવવાનું હોય છે. યાદ કરો કે કોઈ બે સંખ્યાઓના તફાવતના માનાંકનું માપ, એ બે સંખ્યાઓ દ્વારા સંખ્યારેખા પર રજુ થતા બિંદુઓ વચ્ચેનું અંતર દર્શાવે છે. આમ, અચળ સંખ્યા 'a' થી પ્રસારનું માપ શોધવા માટે આપણે મધ્યવર્તી માપથી વિચલનોનાં નિરપેક્ષ મૂલ્યોનો મધ્યક લઈ શકીએ. આ મધ્યકને ***સરેરાશ વિચલન*** કહે છે. આમ, મધ્યવર્તી માપ 'a' ને સાપેક્ષ સરેરાશ વિચલન એ 'a' થી અવલોકનોનાં વિચલનોના નિરપેક્ષ મૂલ્યોનો મધ્યક છે. 'a' થી સરેરાશ વિચલનને M.D.(a) વડે દર્શાવવામાં આવે છે. આથી,

$$\text{M.D.}(a) = \frac{\text{'}a\text{' થી વિચલનોનાં નિરપેક્ષ મૂલ્યોનો સરવાળો}}{\text{અવલોકનની સંખ્યા}}$$

ટિપ્પણી : સરેરાશ વિચલન મધ્યવર્તી સ્થિતિમાનના કોઈપણ માપથી શોધી શકાય છે. પરંતુ આંકડાશાસ્ત્રના અભ્યાસમાં સામાન્ય રીતે મધ્યક અને મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલનનો ઉપયોગ થાય છે.

ચાલો, આપણે મધ્યકને સાપેક્ષ સરેરાશ વિચલન અને મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલનની ગણતરી કઈ રીતે કરવી તેનો અભ્યાસ કરીએ.

15.4.1 અવર્ગીકૃત માહિતી માટે સરેરાશ વિચલન (Mean deviation for ungrouped data)

ધારો કે n અવલોકનોના પ્રાપ્તાંકો $x_1, x_2, x_3, ..., x_n$ છે. મધ્યક અથવા મધ્યસ્થની સાપેક્ષે સરેરાશ વિચલનની ગણતરી નીચે દર્શાવ્યા પ્રમાણે થાય છે :

પગલું 1 : જેની સાપેક્ષે સરેરાશ વિચલન શોધવાનું છે એ મધ્યવર્તી સ્થિતિમાનના માપની ગણતરી કરો. ધારો કે તે 'a' છે.

પગલું 2 : પ્રત્યેક અવલોકન x_i થી a નું વિચલન શોધો, એટલે કે, $x_1 - a, x_2 - a, x_3 - a, . . . , x_n - a$.

પગલું 3 : વિચલનોનાં નિરપેક્ષ મૂલ્યો શોધો, અર્થાત્ જો ઋણ સંજ્ઞા હોય તો, (–) દૂર કરો એટલે કે,

$$|x_1 - a|, |x_2 - a|, |x_3 - a|, ..., |x_n - a| \text{ મેળવો.}$$

પગલું 4 : વિચલનોનાં નિરપેક્ષ મૂલ્યોનો મધ્યક શોધો. આ મધ્યક એ a ને સાપેક્ષ સરેરાશ વિચલન છે. એટલે કે,

$$\text{M.D.}(a) = \frac{\sum_{i=1}^{n} |x_i - a|}{n}$$

આમ, $$\text{M.D.}(\bar{x}) = \frac{1}{n} \sum_{i=1}^{n} |x_i - \bar{x}|, \text{ જ્યાં } \bar{x} = \text{મધ્યક}$$

અને $$\text{M.D. (M)} = \frac{1}{n} \sum_{i=1}^{n} |x_i - \text{M}|, \text{ જ્યાં } \text{M} = \text{મધ્યસ્થ}$$

નોંધ : આ પ્રકરણમાં જ્યાં સુધી અન્ય સૂચન ન હોય ત્યાં સુધી સંકેત M એ મધ્યસ્થ દર્શાવે છે . ચાલો હવે ઉપર વર્ણવેલ પદો સમજવા માટે નીચે આપેલ ઉદાહરણો જોઈએ.

ઉદાહરણ 1 : નીચે આપેલ માહિતી માટે મધ્યકને સાપેક્ષ સરેરાશ વિચલન શોધો :

6, 7, 10, 12, 13, 4, 8, 12

ઉકેલ : આપણે પદવાર આગળ વધીએ અને નીચે દર્શાવેલ વિગતો મેળવીએ :

પગલું 1 : આપેલ સંખ્યાઓનો મધ્યક

$$\bar{x} = \frac{6+7+10+12+13+4+8+12}{8} = \frac{72}{8} = 9$$

પગલું 2 : ક્રમશઃ અવલોકનોનું મધ્યક $\bar{x}$ થી વિચલન $x_i - \bar{x}$ અર્થાત્

$$6 - 9, 7 - 9, 10 - 9, 12 - 9, 13 - 9, 4 - 9, 8 - 9, 12 - 9$$

અથવા –3, –2, 1, 3, 4, –5, –1, 3 છે.

પગલું 3 : વિચલનોનાં માનાંકનાં મૂલ્યો, એટલે કે $|x_i - \bar{x}|$, 3, 2, 1, 3, 4, 5, 1, 3 છે.

પગલું 4 : મધ્યકને સાપેક્ષ માંગેલ સરેરાશ વિચલન

$$\text{M.D.}(\bar{x}) = \frac{\sum_{i=1}^{8} |x_i - \bar{x}|}{8} = \frac{3+2+1+3+4+5+1+3}{8} = \frac{22}{8} = 2.75$$

નોંધ દરેક વખતે બધાં જ પદોની ગણતરી કરવાને બદલે, આપણે પદોને અવગણીને પદવાર ગણતરી કરી શકીશું.

ઉદાહરણ 2 : નીચે આપેલ માહિતી માટે મધ્યકને સાપેક્ષ સરેરાશ વિચલન શોધો :

12, 3, 18, 17, 4, 9, 17, 19, 20, 15, 8, 17, 2, 3, 16, 11, 3, 1, 0, 5

ઉકેલ : સૌપ્રથમ આપણે આપેલ માહિતીનો મધ્યક ($\bar{x}$) શોધીશું.

$$\bar{x} = \frac{1}{20}\sum_{i=1}^{20} x_i = \frac{200}{20} = 10$$

ક્રમશઃ અવલોકનોના મધ્યક ($\bar{x}$) થી વિચલનનો માનાંક $|x_i - \bar{x}|$; એટલે કે,

2, 7, 8, 7, 6, 1, 7, 9, 10, 5, 2, 7, 8, 7, 6, 1, 7, 9, 10, 5

તેથી $$\sum_{i=1}^{20} |x_i - \bar{x}| = 124$$

અને $$\text{M.D.}\ (\bar{x}) = \frac{124}{20} = 6.2$$

ઉદાહરણ 3 : આપેલ માહિતી માટે મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલન શોધો :

3, 9, 5, 3, 12, 10, 18, 4, 7, 19, 21.

ઉકેલ : અહીં, અવલોકનોની સંખ્યા 11 અયુગ્મ છે. આપેલ સંખ્યાઓને ચઢતા ક્રમમાં ગોઠવતાં,

3, 3, 4, 5, 7, 9, 10, 12, 18, 19, 21 છે.

હવે મધ્યસ્થ $= \left(\frac{11 + 1}{2}\right)$ મું અથવા 6ઠ્ઠું અવલોકન $= 9$

મધ્યસ્થ M થી અવલોકનોનાં વિચલનોનાં નિરપેક્ષ મૂલ્યો, એટલે કે, $|x_i - \text{M}|$ એ

6, 6, 5, 4, 2, 0, 1, 3, 9, 10, 12 છે.

તેથી $$\sum_{i=1}^{11} |x_i - \text{M}| = 58$$

અને $$\text{M.D.}\ (\text{M}) = \frac{1}{11}\sum_{i=1}^{11} |x_i - \text{M}| = \frac{1}{11} \times 58 = 5.27$$

15.4.2 વર્ગીકૃત માહિતી માટે સરેરાશ વિચલન (Mean deviation for grouped data)

આપણે જાણીએ છીએ કે માહિતીનું બે પ્રકારે વર્ગીકરણ કરવામાં આવે છે :

(a) ***અસતત આવૃત્તિ-વિતરણ*** *(Discrete frequency distribution)*

(b) ***સતત આવૃત્તિ-વિતરણ*** *(Continuous frequency distribution)*

ચાલો, આ બંને પ્રકારની માહિતી માટે સરેરાશ વિચલન શોધવાની રીતો વિશે ચર્ચા કરીએ.

(a) અસતત આવૃત્તિ-વિતરણ : ધારો કે આપેલ માહિતીનાં n ભિન્ન અવલોકનો $x_1, x_2, ..., x_n$ છે અને તેમની આવૃત્તિઓ અનુક્રમે $f_1, f_2, ..., f_n$ છે. આ માહિતીને કોષ્ટક સ્વરૂપે નીચે પ્રમાણે દર્શાવી શકાય છે. તેને ***અસતત આવૃત્તિ-વિતરણ*** કહે છે.

$$x : x_1 \quad x_2 \quad x_3 \dots\dots\dots\; x_n$$

$$f : f_1 \quad f_2 \quad f_3 \dots\dots\dots\; f_n$$

(i) મધ્યકની સાપેક્ષ સરેરાશ વિચલન :

સૌપ્રથમ આપણે આપેલ માહિતીનો મધ્યક $\bar{x}$ શોધીશું

અહીં,
$$\bar{x} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i} = \frac{1}{N} \sum_{i=1}^{n} x_i f_i$$

સૂત્રનો ઉપયોગ કરીશું.

$\sum_{i=1}^{n} x_i f_i$ એ અવલોકનો x_i ના તેમને અનુરૂપ આવૃત્તિઓ f_i સાથેના ગુણાકારોનો સરવાળો દર્શાવે છે અને $N = \sum_{i=1}^{n} f_i$ એ આવૃત્તિઓનો સરવાળો છે.

પછી, આપણે અવલોકનો x_i ના મધ્યક $\bar{x}$ પરથી વિચલન શોધીએ છીએ અને તેમનાં નિરપેક્ષ મૂલ્ય મેળવીએ છીએ, એટલે કે પ્રત્યેક $i = 1, 2, ..., n$ માટે $|x_i - \bar{x}|$ શોધવામાં આવે છે.

તેનાં પછી વિચલનોનાં મધ્યકની સાપેક્ષ અપેક્ષિત સરેરાશ વિચલનની ગણતરી કરવામાં આવે છે.

આમ,
$$\text{M.D.}(\bar{x}) = \frac{\sum_{i=1}^{n} f_i |x_i - \bar{x}|}{\sum_{i=1}^{n} f_i} = \frac{1}{N} \sum_{i=1}^{n} f_i |x_i - \bar{x}|$$

(ii) મધ્યસ્થની સાપેક્ષ સરેરાશ વિચલન :

મધ્યસ્થની સાપેક્ષ સરેરાશ વિચલન શોધવા માટે આપેલ અસતત આવૃત્તિ-વિતરણનો મધ્યસ્થ શોધીશું. આના માટે અવલોકનોને ચઢતાં ક્રમમાં ગોઠવીશું. તેના પછી સંચયી આવૃત્તિ મેળવીશું. અહીં, આવૃત્તિઓનો સરવાળો N વડે દર્શાવ્યો છે. જેની સંચયી આવૃત્તિ $\frac{N}{2}$ ને સમાન અથવા એના કરતાં તરત જ વધારે હોય એ અવલોકન હવે નિર્ધારિત કરીશું. અવલોકનોનું આ મૂલ્ય સંખ્યાઓની મધ્યમાં સ્થાયી હોય છે, તેથી આ જરૂરી મધ્યસ્થ છે. મધ્યસ્થ શોધી લીધા પછી, આપણે મધ્યસ્થથી વિચલનોનાં નિરપેક્ષ મૂલ્યોનો મધ્યક શોધીએ છીએ. આ રીતે,

$$\text{M.D.(M)} = \frac{1}{N} \sum_{i=1}^{n} f_i |x_i - M|$$

ઉદાહરણ 4 : નીચે આપેલ માહિતી પરથી મધ્યકને સાપેક્ષ સરેરાશ વિચલન શોધો :

x_i	2	5	6	8	10	12
f_i	2	8	10	7	8	5

ઉકેલ : ચાલો, આપેલ માહિતીને કોષ્ટક 15.1 માં વધારાના સ્તંભો ગણતરી કરીને આગળ દર્શાવ્યા પ્રમાણે તૈયાર કરીએ.

કોષ્ટક 15.1

x_i	f_i	$f_i x_i$	$\lvert x_i - \bar{x} \rvert$	$f_i \lvert x_i - \bar{x} \rvert$
2	2	4	5.5	11
5	8	40	2.5	20
6	10	60	1.5	15
8	7	56	0.5	3.5
10	8	80	2.5	20
12	5	60	4.5	22.5
	40	300		92

અહીં $N = \sum_{i=1}^{6} f_i = 40, \quad \sum_{i=1}^{6} f_i x_i = 300,$

તેથી, $\bar{x} = \frac{1}{N} \sum_{i=1}^{6} f_i x_i = \frac{1}{40} \times 300 = 7.5$

$$\sum_{i=1}^{6} f_i \lvert x_i - \bar{x} \rvert = 92$$

અને $\text{M.D.}\ (\bar{x}) = \frac{1}{N} \sum_{i=1}^{6} f_i \lvert x_i - \bar{x} \rvert = \frac{1}{40} \times 92 = 2.3$

ઉદાહરણ 5 : આપેલ માહિતી માટે મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલન શોધો :

x_i	3	6	9	12	13	15	21	22
f_i	3	4	5	2	4	5	4	3

ઉકેલ : આપેલ અવલોકનો ચઢતા ક્રમમાં જ છે. આ માહિતીમાં સંચયી આવૃત્તિની એક હાર ઉમેરતાં આપણને (કોષ્ટક 15.2) મળે.

કોષ્ટક 15.2

x_i	3	6	9	12	13	15	21	22
f_i	3	4	5	2	4	5	4	3
સંચયી આવૃત્તિ	3	7	12	14	18	23	27	30

હવે, $N = 30$ યુગ્મ સંખ્યા છે.

મધ્યસ્થ એ 15 માં અને 16 માં અવલોકનોની સરેરાશ છે. આ બંને અવલોકનો સંચયી આવૃત્તિ 18 ને સંગત છે. તેને અનુરૂપ અવલોકન 13 છે.

માટે, મધ્યસ્થ $M = \dfrac{15 \text{ મું અવલોકન} + 16 \text{ મું અવલોકન}}{2} = \dfrac{13+13}{2} = 13$

હવે, મધ્યસ્થથી વિચલનોનાં નિરપેક્ષ મૂલ્યો એટલે કે, $\lvert x_i - M \rvert$ કોષ્ટક 15.3 માં દર્શાવ્યાં છે.

કોષ્ટક 15.3

$\|x_i - M\|$	10	7	4	1	0	2	8	9
f_i	3	4	5	2	4	5	4	3
$f_i\|x_i - M\|$	30	28	20	2	0	10	32	27

આપણને $\sum_{i=1}^{8} f_i = 30$ અને $\sum_{i=1}^{8} f_i \left|x_i - M\right| = 149$ મળે છે.

તેથી $\text{M.D.}(M) = \frac{1}{N} \sum_{i=1}^{8} f_i \left|x_i - M\right|$

$$= \frac{1}{30} \times 149 = 4.97$$

(b) સતત આવૃત્તિ-વિતરણ : સતત આવૃત્તિ-વિતરણમાં માહિતીનું, વચ્ચે અંતર ન હોય એવા વર્ગોમાં વર્ગીકરણ કરવામાં આવે છે. એવા પ્રકારની શ્રેણી અને તેમની આવૃત્તિ ક્રમાનુસાર લખવામાં આવે છે.

ઉદાહરણ તરીકે 100 વિદ્યાર્થીઓ દ્વારા પ્રાપ્ત કરેલા ગુણોને સતત આવૃત્તિ-વિતરણમાં નીચે પ્રમાણે દર્શાવી શકાય :

મેળવેલા ગુણ	0-10	10-20	20-30	30-40	40-50	50-60
વિદ્યાર્થીઓની સંખ્યા	12	18	27	20	17	6

(i) મધ્યકને સાપેક્ષ સરેરાશ વિચલન : એક સતત આવૃત્તિ-વિતરણના મધ્યકની ગણતરી કરતાં સમયે આપણે એ ધારી લીધું હતું કે, પ્રત્યેક વર્ગની આવૃત્તિ વર્ગની મધ્યકિંમત પર કેન્દ્રિત હોય છે. અહીં આપણે દરેક વર્ગની મધ્યકિંમત લખીએ છીએ અને અસતત આવૃત્તિ વિતરણની માફક સરેરાશ વિચલન શોધીએ છીએ. ચાલો નીચે આપેલ ઉદાહરણ જોઈએ.

ઉદાહરણ 6 : આપેલ માહિતી માટે મધ્યકને સાપેક્ષ સરેરાશ વિચલન શોધો :

મેળવેલા ગુણ	10-20	20-30	30-40	40-50	50-60	60-70	70-80
વિદ્યાર્થીઓની સંખ્યા	2	3	8	14	8	3	2

ઉકેલ : આપેલ માહિતી પરથી કોષ્ટક 15.4 તૈયાર કરીશું :

કોષ્ટક 15.4

મેળવેલા ગુણ	વિદ્યાર્થીઓની સંખ્યા f_i	મધ્યકિંમત x_i	f_ix_i	$\lvert x_i-\bar{x}\rvert$	$f_i\lvert x_i-\bar{x}\rvert$
10-20	2	15	30	30	60
20-30	3	25	75	20	60
30-40	8	35	280	10	80
40-50	14	45	630	0	0
50-60	8	55	440	10	80
60-70	3	65	195	20	60
70-80	2	75	150	30	60
	40		1800		400

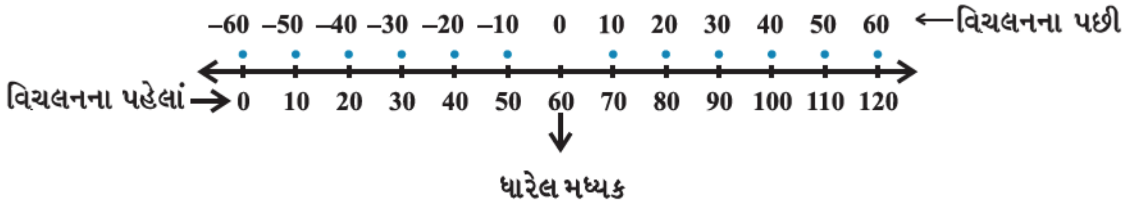
અહીં,
$$N = \sum_{i=1}^{7} f_i = 40, \quad \sum_{i=1}^{7} f_i x_i = 1800$$

તેથી,
$$\bar{x} = \frac{1}{N}\sum_{i=1}^{7} f_i x_i = \frac{1800}{40} = 45, \quad \sum_{i=1}^{7} f_i\lvert x_i-\bar{x}\rvert = 400$$

અને
$$\text{M.D.}(\bar{x}) = \frac{1}{N}\sum_{i=1}^{7} f_i\lvert x_i-\bar{x}\rvert = \frac{1}{40}\times 400 = 10$$

મધ્યકની સાપેક્ષ સરેરાશ વિચલન શોધવાની ટૂંકી રીત :

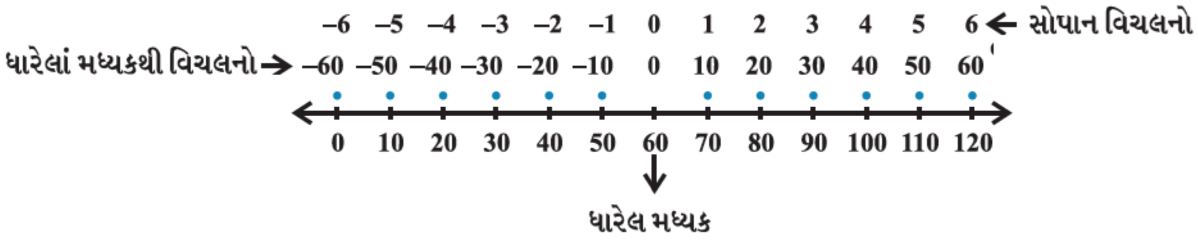
આપણે સોપાન *વિચલન રીત* (*step-deviation method*) નો ઉપયોગ કરીને $\bar{x}$ શોધવાની ગણતરીની કઠિનતા દૂર કરી શકીએ. યાદ કરો કે, આ રીતમાં આપણે માહિતીની મધ્યે અથવા તેની તદ્દન નજીક કોઈ અવલોકનને મધ્યક તરીકે કલ્પી લઈએ છીએ. પછી અવલોકનો (અથવા જુદા જુદા વર્ગની મધ્યકિંમતો) નું આ ધારેલ મધ્યકથી વિચલન મેળવીએ છીએ. આ વિચલન સંખ્યારેખા પર ઊગમબિંદુને શૂન્યથી પ્રતિસ્થાપિત કરીને ધારેલાં મધ્યક સુધી લઈ જવું એ જ છે આકૃતિ 15.3 માં આ દર્શાવ્યું છે.



આકૃતિ 15.3

જો બધાં વિચલનોનો કોઈ સામાન્ય અવયવ હોય તો વિચલનોને સરળ બનાવવા માટે આપણે તેમને આ સામાન્ય અવયવ વડે ભાગીએ છીએ. આ નવાં વિચલનોને *સોપાન-વિચલન* (*step-deviation*) તરીકે ઓળખવામાં આવે છે. સોપાન વિચલન લેવાની

પ્રક્રિયા એ સંખ્યારેખા પર માપ-પદ્ધતિ બદલવાની ક્રિયા છે. તે આકૃતિ 15.4 માં દર્શાવેલ છે.



આકૃતિ 15.4

વિચલનો અને સોપાન-વિચલનો અવલોકનોનાં કદ નાના કરે છે, તેથી ગુણાકાર જેવી ગણતરીઓ સરળ થઈ જાય છે. ધારો કે નવો ચલ $d_i = \frac{x_i - a}{h}$ વડે દર્શાવ્યો છે. અહીં, 'a' ધારેલ મધ્યક છે અને h એ સામાન્ય અવયવ છે. ત્યાર બાદ સોપાન વિચલન રીતે મધ્યક $\bar{x}$ નીચે આપેલાં સૂત્ર દ્વારા શોધી શકાય છે :

$$\bar{x} = a + \frac{\sum_{i=1}^{n} f_i d_i}{N} \cdot h$$

ચાલો ઉદાહરણ 6 ની માહિતી લઈએ અને સોપાન-વિચલન રીતનો ઉપયોગ કરીએ. આપણે ધારેલ મધ્યક $a = 45$ અને $h = 10$ લઈએ અને નીચે આપેલ કોષ્ટક 15.5 તૈયાર કરીએ :

કોષ્ટક 15.5

મેળવેલા ગુણ	વિદ્યાર્થીઓની સંખ્યા f_i	મધ્યબિંદુઓ x_i	$d_i = \frac{x_i - 45}{10}$	$f_i d_i$	$\|x_i - \bar{x}\|$	$f_i \|x_i - \bar{x}\|$
10-20	2	15	– 3	– 6	30	60
20-30	3	25	– 2	– 6	20	60
30-40	8	35	– 1	– 8	10	80
40-50	14	$45 = a$	0	0	0	0
50-60	8	55	1	8	10	80
60-70	3	65	2	6	20	60
70-80	2	75	3	6	30	60
	40			0		400

તેથી
$$\bar{x} = a + \frac{\sum_{i=1}^{7} f_i d_i}{N} \times h$$

$$= 45 + \frac{0}{40} \times 10 = 45$$

અને
$$\text{M.D. } (\bar{x}) = \frac{1}{N} \sum_{i=1}^{7} f_i |x_i - \bar{x}| = \frac{400}{40} = 10$$

ટિપ્પણી: સોપાનવિચલન રીતનો ઉપયોગ $\bar{x}$ મેળવવા માટે કરવામાં આવે છે. બાકીની પ્રક્રિયા એ જ પ્રમાણે છે.

(ii) મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલન :

આપણે સતત આવૃત્તિ-વિતરણ માટે મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલન શોધીશું. જે રીત આપણે મધ્યકને સાપેક્ષ સરેરાશ વિચલન મેળવવા માટે ઉપયોગમાં લીધી હતી એવી જ રીતે આ કાર્ય સંપન્ન કરીશું. કેવળ તફાવત એટલો જ છે કે અહીં જ્યારે વિચલનો લઈએ છીએ ત્યારે મધ્યકને બદલે મધ્યસ્થનો ઉપયોગ કરીએ છીએ.

ચાલો સતત આવૃત્તિ-વિતરણ માટે મધ્યસ્થ શોધવાની પ્રક્રિયાને યાદ કરીએ. સૌપ્રથમ સંખ્યાઓને ચઢતાં ક્રમમાં ગોઠવીએ છીએ. પછી સતત આવૃત્તિ-વિતરણનો મધ્યસ્થ શોધવા માટે પહેલાં જેમાં મધ્યસ્થ સ્થિત હોય છે એ વર્ગ નક્કી કરીએ છીએ. (આ વર્ગને મધ્યસ્થ વર્ગ કહે છે) પછી નીચે દર્શાવેલાં સૂત્રનો ઉપયોગ કરીએ છીએ :

$$\text{મધ્યસ્થ} = l + \frac{\frac{N}{2} - C}{f} \cdot h$$

અહીં, મધ્યસ્થ વર્ગ એ એવો વર્ગ છે કે જેની સંચયી આવૃત્તિ $\frac{N}{2}$ ને બરાબર અથવા તેનાથી તરત જ વધારે હોય. N એ આવૃત્તિઓનો સરવાળો, l, f, h અને C એ અનુક્રમે મધ્યસ્થ વર્ગની અધ:સીમા, આવૃત્તિ, વર્ગલંબાઈ, મધ્યસ્થ વર્ગની તરત આગળના વર્ગની સંચયી આવૃત્તિ છે. મધ્યસ્થ શોધ્યા પછી આપણે મધ્યસ્થથી પ્રત્યેક વર્ગની મધ્યકિંમત સાથેનાં વિચલનોનાં નિરપેક્ષ મૂલ્ય મેળવીએ છીએ. એટલે કે પ્રત્યેક x_i માટે $|x_i - M|$ પ્રાપ્ત કરીએ છીએ.

પછી
$$\text{M.D. (M)} = \frac{1}{N} \sum_{i=1}^{n} f_i |x_i - M|$$

આ પ્રક્રિયાને નીચે આપેલાં ઉદાહરણથી સ્પષ્ટ કરેલ છે :

ઉદાહરણ 7 : નીચે આપેલ માહિતી માટે મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલન શોધો :

વર્ગ	0-10	10-20	20-30	30-40	40-50	50-60
આવૃત્તિ	6	7	15	16	4	2

ઉકેલ : આપેલ માહિતી માટે નીચે દર્શાવેલ કોષ્ટક 15.6 તૈયાર કરો :

કોષ્ટક 15.6

વર્ગ	આવૃત્તિ f_i	સંચયી આવૃત્તિ (*c.f.*)	મધ્યકિંમત x_i	$\|x_i - \text{મધ્યસ્થ}\|$	$f_i\|x_i - \text{મધ્યસ્થ}\|$
0-10	6	6	5	23	138
10-20	7	13	15	13	91
20-30	15	28	25	3	45
30-40	16	44	35	7	112
40-50	4	48	45	17	68
50-60	2	50	55	27	54
	50				508

અહીં, N = 50 છે. તેથી $\frac{N}{2}$ મું એટલે કે 25મું અવલોકન એ વર્ગ 20-30 માં આવશે. તેથી, 20–30 એ મધ્યસ્થ વર્ગ છે.

આપણે જાણીએ છીએ કે,

$$\text{મધ્યસ્થ} = l + \frac{\frac{N}{2} - C}{f} \cdot h$$

અહીં $l = 20$, C = 13, $f = 15$, $h = 10$ અને N = 50 છે.

માટે, $\text{મધ્યસ્થ} = 20 + \frac{25-13}{15} \times 10 = 20 + 8 = 28$

આમ, મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલન,

$$\text{M.D. (M)} = \frac{1}{N}\sum_{i=1}^{6} f_i |x_i - M| = \frac{1}{50} \times 508 = 10.16$$

સ્વાધ્યાય 15.1

પ્રશ્ન 1 અને 2 માં આપેલ માહિતી માટે મધ્યકને સાપેક્ષ સરેરાશ વિચલન શોધો :

1. 4, 7, 8, 9, 10, 12, 13, 17

2. 38, 70, 48, 40, 42, 55, 63, 46, 54, 44

પ્રશ્ન 3 અને 4 માં આપેલ માહિતી માટે મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલન શોધો :

3. 13, 17, 16, 14, 11, 13, 10, 16, 11, 18, 12, 17

4. 36, 72, 46, 42, 60, 45, 53, 46, 51, 49

પ્રશ્ન 5 અને 6 માં આપેલ માહિતી માટે મધ્યકને સાપેક્ષ સરેરાશ વિચલન શોધો :

5.

x_i	5	10	15	20	25
f_i	7	4	6	3	5

6.

x_i	10	30	50	70	90
f_i	4	24	28	16	8

પ્રશ્ન 7 અને 8 માં આપેલ માહિતી માટે મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલન શોધો :

7.

x_i	5	7	9	10	12	15
f_i	8	6	2	2	2	6

8.

x_i	15	21	27	30	35
f_i	3	5	6	7	8

પ્રશ્ન 9 અને 10 માં આપેલ માહિતી માટે મધ્યકને સાપેક્ષ સરેરાશ વિચલન શોધો :

9.

એક દિવસની આવક	0-100	100-200	200-300	300-400	400-500	500-600	600-700	700-800
વ્યક્તિઓની સંખ્યા	4	8	9	10	7	5	4	3

10.

ઊંચાઈ સેમીમાં	95-105	105-115	115-125	125-135	135-145	145-155
કુમારોની સંખ્યા	9	13	26	30	12	10

11. આપેલ માહિતી માટે મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલન શોધો :

ગુણ	0-10	10-20	20-30	30-40	40-50	50-60
કુમારીઓની સંખ્યા	6	8	14	16	4	2

12. 100 વ્યક્તિઓનું વય વિતરણ નીચે આપેલ છે. મધ્યસ્થ વયની સાપેક્ષે સરેરાશ વિચલનની ગણતરી કરો.

વય(વર્ષમાં)	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55
સંખ્યા	5	6	12	14	26	12	16	9

[**સૂચન :** પ્રત્યેક વર્ગની અધઃસીમામાંથી 0.5 ઘટાડીને તેની ઊર્ધ્વસીમામાં 0.5 ઉમેરો અને આપેલ માહિતીને સતત આવૃત્તિ-વિતરણમાં ફેરવો.]

15.4.3 સરેરાશ વિચલનની મર્યાદાઓ :

જે શ્રેણીમાં ચલનની કક્ષા ખૂબ જ ઊંચી હોય, તેમાં મધ્યસ્થ એ મધ્યવર્તી સ્થિતિમાનનું ઉપયોગી માપ નથી હોતું. આમ, આ પરિસ્થિતિમાં મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલન ઉપર સંપૂર્ણ વિશ્વાસ કરી શકાય નહિ.

મધ્યકથી વિચલનોનો સરવાળો (ઋણ સંજ્ઞાને અવગણીને) એ મધ્યસ્થથી વિચલનોનાં સરવાળા કરતાં વધારે હોય છે. માટે, મધ્યકને સાપેક્ષ સરેરાશ વિચલન અધિક વૈજ્ઞાનિક નથી. આમ, ઘણી પરિસ્થિતિઓમાં સરેરાશ વિચલન સંતોષકારક પરિણામ નથી આપતું. સાથે જ સરેરાશ વિચલનને વિચલનોનાં નિરપેક્ષ મૂલ્યોને આધારે મેળવવામાં આવે છે અને તેથી તે વધુ બૈજિક ગણતરીઓ માટે યોગ્ય નથી હતું. આ સૂચવે છે કે આપણને પ્રસારના અન્ય માપની આવશ્યકતા છે. પ્રમાણિત વિચલન એ પ્રસારનું એવું જ એક માપ છે.

15.5 વિચરણ અને પ્રમાણિત વિચલન

યાદ કરો કે જ્યારે આપણે મધ્યક અથવા મધ્યસ્થને સાપેક્ષ સરેરાશ વિચલનની ગણતરી કરતા હતા ત્યારે આપણે વિચલનોના નિરપેક્ષ મૂલ્યો લીધા હતા. આ કરવા પાછળનું કારણ સરેરાશ વિચલનને સાર્થક બનાવવા માટેનું હતું, નહિ તો વિચલનોનો સરવાળો શૂન્ય થઈ જાત(ધન અને ઋણ સંજ્ઞાઓવાળા વિચલનોનો સરવાળો શૂન્ય થાય).

વિચલનોની સંજ્ઞાને કારણે ઊભી થયેલી આ સમસ્યાને વિચલનોનો વર્ગ લઈને પણ દૂર કરી શકાય છે. સ્પષ્ટ છે કે વિચલનોના વર્ગ હંમેશાં અનૃણ હોય છે.

ધારો કે $x_1, x_2, x_3, ..., x_n$ એ n અવલોકનો છે તથા તેમનો મધ્યક $\bar{x}$ છે.

$$(x_1-\bar{x})^2 + (x_2 - \bar{x})^2 + ... + (x_n-\bar{x})^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2.$$

જો આ સરવાળો શૂન્ય હોય તો પ્રત્યેક $(x_i - \bar{x})$ પણ શૂન્ય જ થશે. આનો અર્થ એ થયો કે કોઈ પણ માત્રામાં પ્રસાર નથી કારણ કે બધાં જ અવલોકનો $\bar{x}$ ની બરાબર થાય છે.

જો $\sum_{i=1}^{n}(x_i-\bar{x})^2$ નાની સંખ્યા હોય તો એ નિર્દેશ કરે છે કે અવલોકનો $x_1, x_2, x_3,...,x_n$ એ મધ્યક $\bar{x}$ ની નજીક છે અને તેથી અવલોકનોનો મધ્યક $\bar{x}$ ની સાપેક્ષ પ્રસાર નિમ્ન કક્ષાનો છે. આનાથી વિપરીત જો આ સરવાળો મોટો હોય, તો અવલોકનોનો પ્રસાર મધ્યક $\bar{x}$ થી ઉચ્ચ કક્ષાનો છે. આમ, શું આપણે કહી શકીએ કે સરવાળો $\sum_{i=1}^{n}(x_i-\bar{x})^2$ એ તમામ અવલોકનોના મધ્યક $\bar{x}$ ને સાપેક્ષ પ્રસાર અથવા ફેલાવાનાં માપનું એક સંતોષકારક પ્રતિક છે ?

ચાલો આના માટે આપણે છ અવલોકનો 5, 15, 25, 35, 45, 55 નો એક સમૂહ A લઈએ. આ અવલોકનોનો મધ્યક 30 છે. આ ગણમાં $\bar{x}$ થી વિચલનોના વર્ગોનો સરવાળો નીચે દર્શાવેલ છે :

$$\sum_{i=1}^{6}(x_i-\bar{x})^2 = (5-30)^2 + (15-30)^2 + (25-30)^2 + (35-30)^2 + (45-30)^2 + (55-30)^2$$

$$= 625 + 225 + 25 + 25 + 225 + 625 = 1750$$

એક બીજો સમૂહ B લઈએ. તેનાં 31 અવલોકનો નીચે આપેલ છે :

15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45.

આ અવલોકનોનો મધ્યક $\bar{y} = 30$ છે.

બંને સમૂહ A તથા B નો મધ્યક 30 છે.

હવે, સમૂહ B નાં અવલોકનોના મધ્યક $\bar{y}$ થી વિચલનોના વર્ગોનો સરવાળો નીચે આપેલ છે :

$$\sum_{i=1}^{31}(y_i-\bar{y})^2 = (15-30)^2 + (16-30)^2 + (17-30)^2 + ... + (44-30)^2 + (45-30)^2$$

$$= (-15)^2 + (-14)^2 + ... + (-1)^2 + 0^2 + 1^2 + 2^2 + 3^2 + ... + 14^2 + 15^2$$

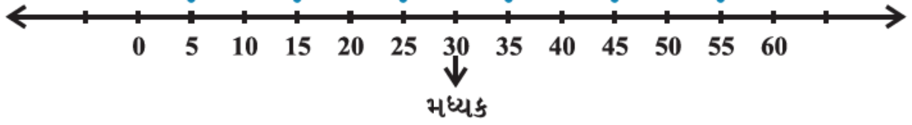
$$= 2\,[15^2 + 14^2 + ... + 1^2]$$

$$= 2\times\frac{15\times(15+1)\,(30+1)}{6} = 5 \times 16 \times 31 = 2480$$

(કારણ કે પ્રથમ n પ્રાકૃતિક સંખ્યાઓના વર્ગોનો સરવાળો = $\frac{n\,(n+1)\,(2n+1)}{6}$. અહીં, $n = 15$)

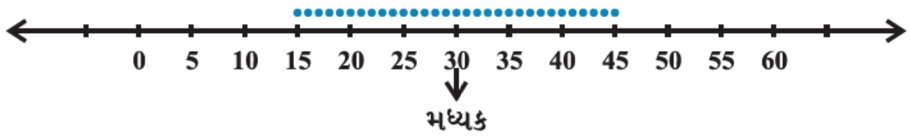
જો $\sum_{i=1}^{n}(x_i-\bar{x})^2$ જ મધ્યકને સાપેક્ષ પ્રસાર માપ હોય, તો આપણે એ કહેવા માટે પ્રેરિત થઈશું કે 31 અવલોકનો ધરાવતાં ગણ B નો 6 અવલોકનોવાળા ગણ A ની તુલનાએ મધ્યકની સાપેક્ષ પ્રસાર વધારે છે. ભલે ને A માં 6 અવલોકનોના મધ્યક $\bar{x}$ ને સાપેક્ષ પ્રસાર(વિચલનોનો વિસ્તાર –25 થી 25) ગણ B ની સરખામણીએ (જ્યાં, વિચલનોનો વિસ્તાર –15 થી 15) વધારે છે. આ હકીકત નીચે આપેલ આકૃતિ પરથી સ્પષ્ટ થાય છે :

ગણ A માટે આકૃતિ 15.5 છે.



આકૃતિ 15.5

ગણ B માટે આકૃતિ 15.6 છે.



આકૃતિ 15.6

આમ, આપણે કહી શકીએ કે મધ્યકથી વિચલનોના વર્ગોનો સરવાળો, એ પ્રસારનું ઉપયોગી માપ નથી. આ મુશ્કેલીને દૂર કરવા માટે આપણે વિચલનોના વર્ગોનો મધ્યક લઈએ, એટલે કે આપણે $\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ લઈએ. ગણ A માટે આપણને મળે છે.

મધ્યક $= \frac{1}{6} \times 1750 = 291.67$ અને ગણ B માટે મધ્યક $\frac{1}{31} \times 2480 = 80$.

આ દર્શાવે છે કે ગણ A માં પ્રસાર ગણ B ની સરખામણીએ વધારે છે. તે બંને ગણોના અપેક્ષાનુસાર પરિણામ અને ભૌમિતિક નિરૂપણ સાથે સુસંગત છે.

આમ, આપણે $\frac{1}{n}\sum(x_i-\bar{x})^2$ સૂત્રને પ્રસારનાં યોગ્ય માપ તરીકે લઈ શકીએ. આ સંખ્યા એટલે કે મધ્યકથી વિચલનોના વર્ગોના મધ્યકને *વિચરણ* (*variance*) કહે છે અને તેને σ^2 (સિગ્માનો વર્ગ એમ વંચાય છે) વડે દર્શાવાય છે.

આમ, n અવલોકનો $x_1, x_2, ..., x_n$ નું વિચરણ

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 \text{ છે.}$$

15.5.1 પ્રમાણિત વિચલન (Standard Deviation)

વિચરણ (*variance*)ની ગણતરીમાં આપણે જોયું કે સ્વતંત્ર અવલોકનો x_i તથા તેમના મધ્યક $\bar{x}$ ના ચલનમાં $(x_i-\bar{x})$ ના વર્ગોનો સમાવેશ થાય છે. આ કારણે વિચરણના ધન વર્ગમૂળને અવલોકનોના મધ્યકને સાપેક્ષ ચલનના પ્રમાણિત માપના સ્વરૂપે દર્શાવવામાં આવે છે અને તેને *પ્રમાણિત વિચલન (standard deviation)* કહે છે. પ્રમાણિત વિચલનને સામાન્ય રીતે σ વડે દર્શાવવામાં આવે છે અને નીચે પ્રમાણે સૂત્ર સ્વરૂપે લખાય છે :

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad \text{... (1)}$$

ચાલો, અવર્ગીકૃત માહિતીનાં ચલન અને પ્રમાણિત વિચલન શોધવાની ગણતરી દર્શાવતાં કેટલાંક ઉદાહરણો જોઈએ:

ઉદાહરણ 8 : નીચે આપેલ માહિતી માટે વિચરણ શોધો.

6, 8, 10, 12, 14, 16, 18, 20, 22, 24

ઉકેલ : આપેલ માહિતી પરથી આપણે નીચેનું કોષ્ટક 15.7 તૈયાર કરીએ. મધ્યકની ગણતરી સોપાન-વિચલન પદ્ધતિ અનુસાર કરી છે અને 14 ને મધ્યક તરીકે ધારી લીધો છે. અવલોકનોની સંખ્યા $n = 10$ છે.

કોષ્ટક 15.7

x_i	$d_i = \frac{x_i - 14}{2}$	મધ્યકથી વિચલનો $(x_i - \bar{x})$	$(x_i - \bar{x})^2$
6	–4	–9	81
8	–3	–7	49
10	–2	–5	25
12	–1	–3	9
14 = a	0	–1	1
16	1	1	1
18	2	3	9
20	3	5	25
22	4	7	49
24	5	9	81
	5		330

તેથી, મધ્યક $\bar{x}$ = ધારેલો મધ્યક + $\frac{\sum_{i=1}^{n} d_i}{n} \times h = 14 + \frac{5}{10} \times 2 = 15$

અને વિચરણ $\sigma^2 = \frac{1}{n}\sum_{i=1}^{10}(x_i - \bar{x})^2 = \frac{1}{10} \times 330 = 33$

આમ, પ્રમાણિત વિચલન $\sigma = \sqrt{33} = 5.74$

15.5.2 અસતત આવૃત્તિ-વિતરણનું પ્રમાણિત વિચલન

આપેલ અસતત આવૃત્તિ-વિતરણ નીચે મુજબ છે :

$x: \quad x_1, \; x_2, \; x_3, \ldots, x_n$

$f: \quad f_1, \; f_2, \; f_3, \ldots, f_n$

આ સંજોગોમાં પ્રમાણિત વિચલન $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i(x_i - \bar{x})^2}$ જ્યાં, $N = \sum_{i=1}^{n} f_i$ ... (2)

ચાલો, નીચે આપેલ ઉદાહરણ લઈએ.

ઉદાહરણ 9 : નીચે આપેલ માહિતી માટે વિચરણ અને પ્રમાણિત વિચલન શોધો :

x_i	4	8	11	17	20	24	32
f_i	3	5	9	5	4	3	1

ઉકેલ : આપેલ માહિતીને કોષ્ટક 15.8 માં દર્શાવેલ છે અને આ કોષ્ટકની રચના કરેલ છે.

કોષ્ટક 15.8

x_i	f_i	$f_i x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
4	3	12	–10	100	300
8	5	40	–6	36	180
11	9	99	–3	9	81
17	5	85	3	9	45
20	4	80	6	36	144
24	3	72	10	100	300
32	1	32	18	324	324
	30	420			1374

અહીં $N = 30$, $\sum_{i=1}^{7} f_i x_i = 420$. તેથી $\bar{x} = \frac{\sum_{i=1}^{7} f_i x_i}{N} = \frac{1}{30} \times 420 = 14$

$$\sum_{i=1}^{7} f_i (x_i - \bar{x})^2 = 1374$$

અને તે પરથી વિચરણ $\sigma^2 = \frac{1}{N} \sum_{i=1}^{7} f_i (x_i - \bar{x})^2$

$$= \frac{1}{30} \times 1374 = 45.8$$

અને પ્રમાણિત વિચલન $\sigma = \sqrt{45.8} = 6.77$

15.5.3 સતત આવૃત્તિ-વિતરણનું પ્રમાણિત વિચલન :

આપેલ સતત આવૃત્તિ-વિતરણના બધા વર્ગોની મધ્યકિંમતો લઈને તેને અસતત આવૃત્તિ-વિતરણમાં રૂપાંતરિત કરી શકાય છે. તે પછી અસતત આવૃત્તિ-વિતરણ માટે પ્રમાણિત વિચલન શોધવાની રીતનો ઉપયોગ કરીશું.

જેનો પ્રત્યેક વર્ગ તેની મધ્યકિંમત x_i તથા આવૃત્તિ f_i દ્વારા વ્યાખ્યાયિત હોય તેવું n વર્ગોવાળું આવૃત્તિ-વિતરણ આપેલ હોય તો તેનું પ્રમાણિત વિચલન નીચે દર્શાવેલ સૂત્ર દ્વારા મેળવી શકાય :

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{n} f_i (x_i - \bar{x})^2}$$

અહીં $\bar{x}$ એ આવૃત્તિ-વિતરણનો મધ્યક છે અને $N = \sum_{i=1}^{n} f_i$.

પ્રમાણિત વિચલન માટેનું બીજું સૂત્ર :

આપણે જાણીએ છીએ કે,

$$\text{વિચરણ } \sigma^2 = \frac{1}{N} \sum_{i=1}^{n} f_i (x_i - \bar{x})^2 = \frac{1}{N} \sum_{i=1}^{n} f_i (x_i^2 + \bar{x}^2 - 2\bar{x}\, x_i)$$

$$= \frac{1}{N}\left[\sum_{i=1}^{n} f_i x_i^2 + \sum_{i=1}^{n} \bar{x}^2 f_i - \sum_{i=1}^{n} 2\bar{x}\, f_i x_i\right]$$

$$= \frac{1}{N}\left[\sum_{i=1}^{n} f_i x_i^2 + \bar{x}^2 \sum_{i=1}^{n} f_i - 2\bar{x}\sum_{i=1}^{n} x_i f_i\right]$$

$$= \frac{1}{N}\left[\sum_{i=1}^{n} f_i x_i^2 + \bar{x}^2\, N - 2\bar{x}\,.\,N\,\bar{x}\right] \qquad \left[\text{અહીં } \frac{1}{N}\sum_{i=1}^{n} x_i f_i = \bar{x} \text{ અથવા } \sum_{i=1}^{n} x_i f_i = N\bar{x}\right]$$

$$= \frac{1}{N}\sum_{i=1}^{n} f_i x_i^2 + \bar{x}^2 - 2\bar{x}^2$$

$$= \frac{1}{N}\sum_{i=1}^{n} f_i x_i^2 - \bar{x}^2$$

અથવા $\sigma^2 = \frac{1}{N}\sum_{i=1}^{n} f_i x_i^2 - \left(\frac{\sum_{i=1}^{n} f_i x_i}{N}\right)^2 = \frac{1}{N^2}\left[N\sum_{i=1}^{n} f_i x_i^2 - \left(\sum_{i=1}^{n} f_i x_i\right)^2\right]$

આમ, પ્રમાણિત વિચલન $\sigma = \frac{1}{N}\sqrt{N\sum_{i=1}^{n} f_i x_i^2 - \left(\sum_{i=1}^{n} f_i x_i\right)^2}$ ... (3)

ઉદાહરણ 10 : નીચે આપેલ વિતરણ માટે મધ્યક, વિચરણ અને પ્રમાણિત વિચલનની ગણતરી કરો :

વર્ગ	30-40	40-50	50-60	60-70	70-80	80-90	90-100
આવૃત્તિ	3	7	12	15	8	3	2

ઉકેલ : આપેલ માહિતી પરથી આપણે નીચે દર્શાવેલ કોષ્ટક 15.9 તૈયાર કરીએ.

કોષ્ટક 15.9

વર્ગ	આવૃત્તિ (f_i)	મધ્ય-કિંમત (x_i)	$f_i x_i$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
30-40	3	35	105	729	2187
40-50	7	45	315	289	2023
50-60	12	55	660	49	588
60-70	15	65	975	9	135
70-80	8	75	600	169	1352
80-90	3	85	255	529	1587
90-100	2	95	190	1089	2178
	50		3100		10,050

આમ, મધ્યક $\bar{x} = \frac{1}{N}\sum_{i=1}^{7} f_i x_i = \frac{3100}{50} = 62$

વિચરણ $\sigma^2 = \frac{1}{N}\sum_{i=1}^{7} f_i (x_i - \bar{x})^2$

$= \frac{1}{50} \times 10050 = 201$

અને પ્રમાણિત વિચલન $\sigma = \sqrt{201} = 14.18$

ઉદાહરણ 11 : નીચે આપેલ માહિતી માટે પ્રમાણિત વિચલન શોધો :

x_i	3	8	13	18	23
f_i	7	10	15	10	6

ઉકેલ : ચાલો નીચેનું કોષ્ટક 15.10 તૈયાર કરીએ :

કોષ્ટક 15.10

x_i	f_i	$f_i x_i$	x_i^2	$f_i x_i^2$
3	7	21	9	63
8	10	80	64	640
13	15	195	169	2535
18	10	180	324	3240
23	6	138	529	3174
	48	614		9652

હવે સૂત્ર (3) નો ઉપયોગ કરતાં,

$$\sigma = \frac{1}{N}\sqrt{N\sum f_i x_i^2 - \left(\sum f_i x_i\right)^2}$$

$$= \frac{1}{48}\sqrt{48 \times 9652 - (614)^2}$$

$$= \frac{1}{48}\sqrt{463296 - 376996}$$

$$= \frac{1}{48} \times 293.77 = 6.12$$

માટે, પ્રમાણિત વિચલન $\sigma = 6.12$

15.5.4 વિચરણ અને પ્રમાણિત વિચલન શોધવાની ટૂંકી રીત :

કેટલીક વાર અસતત વિતરણમાં x_i ની કિંમતો અથવા સતત વિતરણના જુદા જુદા વર્ગોની મધ્યકિંમતો x_i ની કિંમતો ઘણી મોટી હોય છે. તેથી મધ્યક અને ચલનની ગણતરી કંટાળાજનક હોય છે અને વધારે સમય લે છે. આવા આવૃત્તિ-વિતરણ કે

જેમાં વર્ગની લંબાઈ સમાન હોય તેમાં સોપાન-વિચલન રીત દ્વારા આ પ્રક્રિયાને સરળ બનાવી શકાય છે.

માની લો કે ધારેલ મધ્યક 'A' છે અને માપ પદ્ધતિને (*scale*) $\frac{1}{h}$ ગણી કરી છે. (h એ વર્ગ અંતરાલની લંબાઈ છે) ધારો કે પદ-વિચલનો અથવા નવી કિંમતો y_i છે.

એટલે કે $$y_i = \frac{x_i - A}{h} \quad \text{અથવા} \quad x_i = A + hy_i \qquad \ldots (1)$$

આપણે જાણીએ છીએ કે $$\overline{x} = \frac{\sum_{i=1}^{n} f_i x_i}{N} \qquad \ldots (2)$$

(1) માંથી x_i ની કિંમત (2) માં મૂકતાં, આપણી પાસે,

$$\overline{x} = \frac{\sum_{i=1}^{n} f_i (A + hy_i)}{N}$$

$$= \frac{1}{N}\left(\sum_{i=1}^{n} f_i A + \sum_{i=1}^{n} h\, f_i y_i\right)$$

$$= \frac{1}{N}\left(A\sum_{i=1}^{n} f_i + h\sum_{i=1}^{n} f_i y_i\right)$$

$$= A \cdot \frac{N}{N} + h\,\frac{\sum_{i=1}^{n} f_i y_i}{N} \qquad \left(\text{કારણ કે, } \sum_{i=1}^{n} f_i = N\right)$$

આમ, $$\overline{x} = A + h\,\overline{y} \qquad \ldots (3)$$

હવે, ચલ x નું વિચરણ $$\sigma_x^2 = \frac{1}{N}\sum_{i=1}^{n} f_i (x_i - \overline{x})^2$$

$$= \frac{1}{N}\sum_{i=1}^{n} f_i\,(A + hy_i - A - h\,\overline{y})^2 \qquad \text{((1) અને (3) નો ઉપયોગ કરતાં)}$$

$$= \frac{1}{N}\sum_{i=1}^{n} f_i\, h^2\,(y_i - \overline{y})^2$$

$$= \frac{h^2}{N}\sum_{i=1}^{n} f_i\,(y_i - \overline{y})^2 = h^2 \times \text{ચલ } y_i \text{ નું વિચરણ}$$

એટલે કે $$\sigma_x^2 = h^2 \sigma_y^2$$

અથવા $$\sigma_x = h\sigma_y \qquad \ldots (4)$$

(3) અને (4) પરથી આપણી પાસે,

$$\sigma_x = \frac{h}{N}\sqrt{N\sum_{i=1}^{n} f_i y_i^2 - \left(\sum_{i=1}^{n} f_i y_i\right)^2} \qquad \ldots (5)$$

ચાલો, ઉદાહરણ 11 ને સમીકરણ (5) નો ઉપયોગ કરીને ટૂંકી રીત દ્વારા ઉકેલીએ.

ઉદાહરણ 12 : નીચે આપેલ આવૃત્તિ-વિતરણ માટે મધ્યક, વિચરણ અને પ્રમાણિત વિચલન શોધો.

વર્ગ	30-40	40-50	50-60	60-70	70-80	80-90	90-100
આવૃત્તિ	3	7	12	15	8	3	2

ઉકેલ : ધારો કે ધારેલ મધ્યક A = 65 છે. અહીં $h = 10$

આપેલ માહિતી પરથી નીચે દર્શાવેલ કોષ્ટક 15.11 તૈયાર કરવામાં આવ્યુ છે.

કોષ્ટક 15.11

વર્ગ	આવૃત્તિ f_i	મધ્યકિંમત x_i	$y_i = \frac{x_i - 65}{10}$	y_i^2	$f_i y_i$	$f_i y_i^2$
30-40	3	35	– 3	9	– 9	27
40-50	7	45	– 2	4	– 14	28
50-60	12	55	– 1	1	– 12	12
60-70	15	65 = A	0	0	0	0
70-80	8	75	1	1	8	8
80-90	3	85	2	4	6	12
90-100	2	95	3	9	6	18
	N=50				– 15	105

તેથી $$\bar{x} = A + \frac{\sum f_i y_i}{50} \times h = 65 - \frac{15}{50} \times 10 = 62$$

વિચરણ $$\sigma^2 = \frac{h^2}{N^2}\left[N\Sigma f_i y_i^2 - (\Sigma f_i y_i)^2\right]$$

$$= \frac{(10)^2}{(50)^2}\left[50 \times 105 - (-15)^2\right]$$

$$= \frac{1}{25}[5250 - 225] = 201$$

અને પ્રમાણિત વિચલન $\sigma = \sqrt{201} = 14.18$

સ્વાધ્યાય 15.2

પ્રશ્ન 1 થી 5 માં આપેલ પ્રત્યેક માહિતી માટે મધ્યક અને વિચરણ શોધો :

1. 6, 7, 10, 12, 13, 4, 8, 12
2. પ્રથમ n-પ્રાકૃતિક સંખ્યાઓ

3. ત્રણના પ્રથમ 10 ગુણિત

4.

x_i	6	10	14	18	24	28	30
f_i	2	4	7	12	8	4	3

5.

x_i	92	93	97	98	102	104	109
f_i	3	2	3	2	6	3	3

6. ટૂંકી રીતનો ઉપયોગ કરીને મધ્યક અને પ્રમાણિત વિચલન શોધો.

x_i	60	61	62	63	64	65	66	67	68
f_i	2	1	12	29	25	12	10	4	5

પ્રશ્ન 7 અને 8 માં આપેલ આવૃત્તિ વિતરણ માટે મધ્યક અને વિચરણ શોધો.

7.

વર્ગ	0-30	30-60	60-90	90-120	120-150	150-180	180-210
આવૃત્તિ	2	3	5	10	3	5	2

8.

વર્ગ	0-10	10-20	20-30	30-40	40-50
આવૃત્તિ	5	8	15	16	6

9. ટૂંકી રીતનો ઉપયોગ કરીને મધ્યક, વિચરણ અને પ્રમાણિત વિચલન શોધો.

ઊંચાઈ સેમીમાં	70-75	75-80	80-85	85-90	90-95	95-100	100-105	105-110	110-115
બાળકોની સંખ્યા	3	4	7	7	15	9	6	6	3

10. એક ડિઝાઈનમાં બનાવેલ વર્તુળોના વ્યાસ (મિમીમાં) નીચે આપ્યા છે :

વ્યાસ	33-36	37-40	41-44	45-48	49-52
વર્તુળોની સંખ્યા	15	17	21	22	25

વર્તુળોના વ્યાસનું પ્રમાણિત વિચલન અને મધ્યક વ્યાસ શોધો.

[**સૂચન :** પ્રથમ આપેલ માહિતીને સતત બનાવો. તે માટે વર્ગોને 32.5-36.5, 36.5-40.5, 40.5-44.5, 44.5 - 48.5, 48.5 - 52.5 માં પરિવર્તિત કરો અને પછી આગળ વધો.]

15.6 આવૃત્તિ-વિતરણનું વિશ્લેષણ

આ પ્રકરણના આગળના ભાગોમાં આપણે પ્રસારનાં કેટલાંક માપ વિશે અભ્યાસ કર્યો છે. જે એકમોમાં માહિતી આપેલ હોય છે એ જ એકમો સરેરાશ વિચલન અને પ્રમાણિત વિચલનના પણ હોય છે. જ્યારે આપણે ભિન્ન એકમોનો ઉપયોગ કરી સમાન મધ્યકવાળી બે

શ્રેણીની તુલના, તેનાં માટે કરવા માંગીએ છીએ, ત્યારે કેવળ પ્રસારના માપની ગણતરી જ નથી કરતાં, પરંતુ આપણને એવાં માપની જરૂરત હોય છે કે જે એકમોથી સ્વતંત્ર હોય. એકમથી સ્વતંત્ર, ચલનના માપને *ચલનાંક* (*coefficient of variation*) કહે છે. તેને C.V. વડે દર્શાવવામાં આવે છે.

ચલનાંકને નીચે પ્રમાણે વ્યાખ્યાયિત કરવામાં આવે છે :

$$\text{C.V.} = \frac{\sigma}{\bar{x}} \times 100, \ \bar{x} \neq 0,$$

જ્યાં σ અને $\bar{x}$ અનુક્રમે આપેલ માહિતીના પ્રમાણિત વિચલન અને મધ્યક છે.

બે શ્રેણીઓમાં ચલન અથવા પ્રસારની સરખામણી કરવા માટે આપણે દરેક શ્રેણીનો ચલનાંક (C.V.) મેળવીએ છીએ. જે શ્રેણીનો ચલનાંક મોટો હોય તેને બીજી શ્રેણી કરતાં વધારે *ચલનશીલ શ્રેણી* કહે છે. નાના (C.V.) વાળી શ્રેણીને બીજી કરતાં વધારે સ્થિર કહે છે.

15.6.1 બે સમાન મધ્યકવાળા આવૃત્તિ-વિતરણોની સરખામણી

ધારો કે $\bar{x}_1$ અને σ_1 એ પ્રથમ આવૃત્તિ-વિતરણનાં મધ્યક અને પ્રમાણિત વિચલન છે તથા $\bar{x}_2$ અને σ_2 એ દ્વિતીય વિતરણના મધ્યક અને પ્રમાણિત વિચલન છે.

તેથી C.V. (પ્રથમ આવૃત્તિ-વિતરણ) = $\frac{\sigma_1}{\bar{x}_1} \times 100$

અને C.V. (દ્વિતિય આવૃત્તિ-વિતરણ) = $\frac{\sigma_2}{\bar{x}_2} \times 100$

જો $\bar{x}_1 = \bar{x}_2 = \bar{x}$ આપેલ હોય, તો

C.V. (પ્રથમ આવૃત્તિ-વિતરણ) = $\frac{\sigma_1}{\bar{x}} \times 100$ ... (1)

અને C.V. (દ્વિતિય આવૃત્તિ-વિતરણ) = $\frac{\sigma_2}{\bar{x}} \times 100$ ... (2)

(1) અને (2) પરથી સ્પષ્ટ છે કે બંને C.V. ની સરખામણી σ_1 અને σ_2 ના આધારે જ કરી શકાય છે.

આમ, આપણે કહી શકીએ કે સમાન મધ્યકવાળી બે શ્રેણીઓ પૈકી જે શ્રેણીમાં વધારે પ્રમાણિત વિચલન હોય તેને વધારે ચલિત અથવા ફેલાયેલી શ્રેણી કહે છે. તદ્ઉપરાંત પ્રમાણિત વિચલનનાં નાના(ઓછા) મૂલ્યવાળી શ્રેણીને પ્રમાણમાં બીજી શ્રેણી કરતાં વિશેષ સ્થિર શ્રેણી કહેવાય છે.

ચાલો આપણે નીચે આપેલાં ઉદાહરણો જોઈએ.

ઉદાહરણ 13 : એક કારખાનામાં બે એકમો A અને B માં કર્મીઓની સંખ્યા અને તેમને ચૂકવવામાં આવતાં વેતન નીચે આપ્યા છે :

	A	B
કર્મીઓની સંખ્યા	5000	6000
સરેરાશ માસિક વેતન	₹ 2500	₹ 2500
વેતનોની આવૃત્તિનું વિચરણ	81	100

વ્યક્તિગત વેતનોમાં A અથવા B એકમમાંથી કયા કારખાનામાં વધારે ચલનીયતા છે ?

ઉકેલ : એકમ A માં વેતનોના વિતરણનું વિચરણ $\sigma_1^2 = 81$

તેથી, એકમ A માં વેતનોના આવૃત્તિ-વિતરણનું પ્રમાણિત વિચલન $\sigma_1 = 9$

સાથે જ એકમ B માં વેતનોના આવૃત્તિ-વિતરણનું વિચરણ $\sigma_2^2 = 100$

તેથી, એકમ B માં વેતનોના આવૃત્તિ-વિતરણનું પ્રમાણિત વિચલન $\sigma_2 = 10$

બે એકમોમાં સરેરાશ વેતન સમાન એટલે ₹ 2500 છે. તેથી મોટા પ્રમાણિત વિચલનવાળા એકમમાં વધારે ચલન હશે.

આમ, એકમ B માં વ્યક્તિગત વેતનમાં વધારે ચલન છે.

ઉદાહરણ 14 : બે વિતરણોના ચલનાંક (C.V.) અનુક્રમે 60 અને 70 છે તથા એમનાં પ્રમાણિત વિચલનો અનુક્રમે 21 અને 16 છે. તેમના મધ્યક શું થશે ?

ઉકેલ : અહીં, C.V. (પ્રથમ આવૃત્તિ-વિતરણ) = 60, $\sigma_1 = 21$

C.V. (દ્વિતીય આવૃત્તિ-વિતરણ) = 70, $\sigma_2 = 16$ આપેલ છે.

ધારો કે $\bar{x}_1$ અને $\bar{x}_2$ એ અનુક્રમે પ્રથમ અને દ્વિતીય વિતરણનાં મધ્યકો છે.

હવે C.V. (પ્રથમ આવૃત્તિ-વિતરણ) = $\dfrac{\sigma_1}{\bar{x}_1} \times 100$

માટે $60 = \dfrac{21}{\bar{x}_1} \times 100$ અથવા $\bar{x}_1 = \dfrac{21}{60} \times 100 = 35$

અને C.V. (દ્વિતીય આવૃત્તિ-વિતરણ) = $\dfrac{\sigma_2}{\bar{x}_2} \times 100$

એટલે કે $70 = \dfrac{16}{\bar{x}_2} \times 100$ અથવા $\bar{x}_2 = \dfrac{16}{70} \times 100 = 22.85$

ઉદાહરણ 15 : ધોરણ 11 ના એક સેક્શનમાં વિદ્યાર્થીઓની ઊંચાઈ અને વજન માટે નીચે પ્રમાણે માહિતી મળી છે :

	ઊંચાઈ	વજન
મધ્યક	162.6 સેમી	52.36 કિગ્રા
વિચરણ	127.69 સેમી^{2}	23.1361 કિગ્રા^{2}

શું આપણે કહી શકીએ કે વજનમાં ઊંચાઈની સરખામણીએ વધારે ચલન છે ?

ઉકેલ : આપણે ચલનની સરખામણી માટે તેમના ચલનાંક (C.V.) ની ગણતરી કરીશું.

ઊંચાઈમાં વિચરણ = 127.69 સેમી^{2}

તેથી ઊંચાઈનું પ્રમાણિત વિચલન = $\sqrt{127.69}$ સેમી = 11.3 સેમી

હવે, વજનમાં વિચરણ = 23.1361 કિગ્રા^{2}

તેથી વજનનું પ્રમાણિત વિચલન = $\sqrt{23.1361}$ કિગ્રા = 4.81 કિગ્રા

હવે, ચલનાંક (C.V.) નીચે પ્રમાણે મેળવવામાં આવે છે :

ઊંચાઈનો ચલનાંક (C.V.) = $\dfrac{\text{પ્રમાણિત વિચલન}}{\text{મધ્યક}} \times 100$

$$= \frac{11.3}{162.6} \times 100 = 6.95$$

અને વજનનો ચલનાંક (C.V.) $= \frac{4.81}{52.36} \times 100 = 9.18$

સ્પષ્ટ છે કે વજનનો C.V. એ ઊંચાઈના C.V. કરતાં મોટો છે.

તેથી આપણે કહી શકીએ કે વજનમાં ઊંચાઈ કરતાં વધારે ચલન છે.

સ્વાધ્યાય 15.3

1. નીચે આપેલ માહિતી પરથી બતાવો કે A અને B માંથી કયા સમૂહમાં વધારે ચલન છે ?

ગુણ	10-20	20-30	30-40	40-50	50-60	60-70	70-80
સમૂહ A	9	17	32	33	40	10	9
સમૂહ B	10	20	30	25	43	15	7

2. X અને Y નાં નીચે આપેલાં શૅરનાં મૂલ્યો પરથી બતાવો કે ક્યા શેરનાં મૂલ્યોમાં વધારે સ્થિરતા છે ?

X	35	54	52	53	56	58	52	50	51	49
Y	108	107	105	105	106	107	104	103	104	101

3. એક કારખાનાની બે શાખાઓ A અને B ના કર્મીઓના આપેલાં માસિક વેતન નું વિશ્લેષણ નીચે પ્રમાણે છે :

	શાખા A	**શાખા B**
વેતન મેળવનારા કર્મીઓની સંખ્યા	586	648
માસિક વેતનોનો મધ્યક	₹ 5253	₹ 5253
વિતરણનું વિચરણ	100	121

(i) A અને B માંથી કઈ શાખા પોતાના કર્મીઓને વધારે રકમ માસિક વેતનના રૂપમાં ચૂકવે છે ?

(ii) વ્યક્તિગત વેતનોમાં કઈ શાખા A અથવા B માં વધારે ચલનીયતા છે ?

4. ટીમ A દ્વારા એક સત્રમાં રમેલી ફૂટબૉલ મેચના આંકડા નીચે આપ્યા છે :

નોંધાવેલ ગૉલની સંખ્યા	0	1	2	3	4
મૅચની સંખ્યા	1	9	7	5	3

ટીમ B દ્વારા રમવામાં આવેલી મૅચમાં બનાવેલ ગૉલની સંખ્યાનો મધ્યક પ્રતિ મૅચ 2 અને ગૉલની સંખ્યાનું પ્રમાણિત વિચલન 1.25 હતાં. કઈ ટીમને વધારે સુસંગત માની શકાય ?

5. 50 વનસ્પતિ ઉત્પાદનોની લંબાઈ x (સેમીમાં) અને વજન y (ગ્રામમાં) નો સરવાળો અને વર્ગોનો સરવાળો નીચે આપેલો છે :

$$\sum_{i=1}^{50} x_i = 212, \quad \sum_{i=1}^{50} x_i^2 = 902.8, \quad \sum_{i=1}^{50} y_i = 261 \text{ અને } \sum_{i=1}^{50} y_i^2 = 1457.6$$

શેમાં વધારે ચલન છે, લંબાઈ કે વજન?

પ્રકીર્ણ ઉદાહરણો

ઉદાહરણ 16 : 20 અવલોકનોનું વિચરણ 5 છે. જો પ્રત્યેક અવલોકનને 2 વડે ગુણવામાં આવે, તો પ્રાપ્ત થયેલ અવલોકનો માટે નવું વિચરણ શોધો.

ઉકેલ : ધારો કે અવલોકનો $x_1, x_2, ..., x_{20}$ છે અને તેમનો મધ્યક $\overline{x}$ છે. વિચરણ = 5 અને $n = 20$ આપેલ છે. આપણે જાણીએ છીએ કે,

$$\text{વિચરણ } \sigma^2 = \frac{1}{n}\sum_{i=1}^{20}(x_i - \overline{x})^2 \text{ એટલે કે, } 5 = \frac{1}{20}\sum_{i=1}^{20}(x_i - \overline{x})^2$$

અથવા $$\sum_{i=1}^{20}(x_i - \overline{x})^2 = 100 \qquad ... (1)$$

જો પ્રત્યેક અવલોકનોને 2 વડે ગુણવામાં આવે અને પરિણામે મળતા નવાં અવલોકનો y_i હોય, તો

$$y_i = 2x_i \text{ એટલે કે, } x_i = \frac{1}{2}y_i$$

માટે $$\overline{y} = \frac{1}{n}\sum_{i=1}^{20}y_i = \frac{1}{20}\sum_{i=1}^{20}2x_i = 2 . \frac{1}{20}\sum_{i=1}^{20}x_i$$

એટલે કે $$\overline{y} = 2\,\overline{x} \quad \text{અથવા} \quad \overline{x} = \frac{1}{2}\overline{y}$$

હવે x_i અને $\overline{x}$ ની કિંમતો (1) માં મૂકતાં, આપણને મળે છે.

$$\sum_{i=1}^{20}\left(\frac{1}{2}y_i - \frac{1}{2}\overline{y}\right)^2 = 100, \text{ એટલે કે, } \sum_{i=1}^{20}(y_i - \overline{y})^2 = 400$$

આમ, નવાં અવલોકનોનું વિચરણ $= \frac{1}{20} \times 400 = 20 = 2^2 \times 5$

નોંધ અહીં વાંચકે નોંધ કરવી જોઈએ કે જો પ્રત્યેક અવલોકનને અચળ સંખ્યા k વડે ગુણવામાં આવે તો પરિણામે મળતાં નવાં અવલોકનોનું વિચરણ એ મૂળ વિચરણના k^2 ગણું થાય છે.

ઉદાહરણ 17 : પાંચ અવલોકનોનો મધ્યક 4.4 છે તથા તેમનું વિચરણ 8.24 છે. જો ત્રણ અવલોકનો 1, 2 અને 6 હોય, તો બાકીનાં બે અવલોકનો શોધો.

ઉકેલ : ધારો કે અન્ય બે અવલોકનો x અને y છે.

માટે તે શ્રેણી 1, 2, 6, x, y છે.

હવે, મધ્યક $$\overline{x} = 4.4 = \frac{1+2+6+x+y}{5}$$

અથવા $$22 = 9 + x + y$$

માટે, $$x + y = 13 \qquad ... (1)$$

વળી, $$\text{વિચરણ} = 8.24 = \frac{1}{n}\sum_{i=1}^{5}(x_i - \overline{x})^2$$

એટલે કે $$8.24 = \frac{1}{5}\left[(3.4)^2 + (2.4)^2 + (1.6)^2 + x^2 + y^2 - 2\times 4.4\,(x+y) + 2\times(4.4)^2\right]$$

અથવા $$41.20 = 11.56 + 5.76 + 2.56 + x^2 + y^2 - 8.8 \times 13 + 38.72$$

માટે, $x^2 + y^2 = 97$ ... (2)

પરંતુ (1) પરથી, આપણી પાસે,

$$x^2 + y^2 + 2xy = 169 \quad \text{... (3)}$$

તથા (2) અને (3) પરથી, આપણી પાસે,

$$2xy = 72 \quad \text{... (4)}$$

હવે (4) ને (2) માંથી બાદ કરતાં આપણને મળે છે

$$x^2 + y^2 - 2xy = 97 - 72 \text{ i.e. } (x-y)^2 = 25$$

અથવા $x - y = \pm 5$ ... (5)

તેથી, (1) અને (5) પરથી આપણને મળે છે

જ્યારે $x - y = 5$ ત્યારે $x = 9, \quad y = 4$

અથવા જ્યારે $x - y = -5$ ત્યારે $x = 4, \quad y = 9$

આમ, બાકીનાં બે અવલોકનો 4 અને 9 છે.

ઉદાહરણ 18 : જો પ્રત્યેક અવલોકન $x_1, x_2, ..., x_n$માં કોઈ ધન કે ઋણ સંખ્યા 'a' ઉમેરવામાં આવે, તો સાબિત કરો કે વિચરણ બદલાતું નથી.

ઉકેલ : ધારો કે $\bar{x}$ એ અવલોકનો $x_1, x_2, ..., x_n$નો મધ્યક છે, તો વિચરણ નીચેના સૂત્રથી દર્શાવાય છે :

$$\sigma_1^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

જો પ્રત્યેક અવલોકનોમાં 'a' ઉમેરવામાં આવે તો નવાં અવલોકનો y_i થશે,

$$y_i = x_i + a \quad \text{... (1)}$$

ધારો કે નવાં અવલોકનોનો મધ્યક $\bar{y}$ છે અને

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n}(x_i + a) = \frac{1}{n}\left[\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} a\right] = \frac{1}{n}\sum_{i=1}^{n} x_i + \frac{na}{n} = \bar{x} + a$$

એટલે કે $\bar{y} = \bar{x} + a$ ... (2)

આમ, નવાં અવલોકનોનું વિચરણ,

$$\sigma_2^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i + a - \bar{x} - a)^2 \quad \text{[(1) અને (2) નો ઉપયોગ કરતાં]}$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sigma_1^2$$

તેથી $\sigma_1 = \sigma_2$

આમ,નવાં અવલોકનોનું વિચરણ મૂળ અવલોકનોનું હતું તે જ છે.

> ☛ **નોંધ** ધ્યાન રાખો કે અવલોકનોના કોઈ પણ સમૂહમાં પ્રત્યેક અવલોકનમાં કોઈ એક સંખ્યા ઉમેરવાથી કે બાદ કરવાથી વિચરણમાં કોઈ જ ફેરફાર થતો નથી.

ઉદાહરણ 19 : એક વિદ્યાર્થીએ 100 અવલોકનોનો મધ્યક 40 અને પ્રમાણિત વિચલન 5.1 મેળવ્યા છે, પરંતુ એણે ભૂલથી એક અવલોકન 40 ને બદલે 50 લઈ લીધું હતું, તો સાચો મધ્યક અને પ્રમાણિત વિચલન શું છે ?

ઉકેલ : આપેલ અવલોકનોની સંખ્યા $n = 100$ તથા ખોટો મધ્યક $\bar{x} = 40,$

અને ખોટું પ્રમાણિત વિચલન $\sigma = 5.1$

આપણે જાણીએ છીએ કે, $$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

એટલે કે, $$40 = \frac{1}{100}\sum_{i=1}^{100} x_i \quad \text{અથવા} \quad \sum_{i=1}^{100} x_i = 4000$$

આનો અર્થ એ છે કે ખોટાં અવલોકનોનો સરવાળો = 4000

આમ, સાચાં અવલોકનોનો સરવાળો = ખોટો સરવાળો – 50 + 40

$$= 4000 - 50 + 40 = 3990$$

$$\text{તેથી સાચો મધ્યક} = \frac{\text{સાચો સરવાળો}}{100} = \frac{3990}{100} = 39.9$$

$$\text{પ્રમાણિત વિચલન } \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2}$$

$$= \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - (\bar{x})^2}$$

એટલે કે, $$5.1 = \sqrt{\frac{1}{100} \times \text{ખોટો} \sum_{i=1}^{n} x_i^2 - (40)^2}$$

અથવા $$26.01 = \frac{1}{100} \times \text{ખોટો} \sum_{i=1}^{n} x_i^2 - 1600$$

માટે $$\text{ખોટો} \sum_{i=1}^{n} x_i^2 = 100\,(26.01 + 1600) = 162601$$

હવે $$\text{સાચો} \sum_{i=1}^{n} x_i^2 = \text{ખોટો} \sum_{i=1}^{n} x_i^2 - (50)^2 + (40)^2$$

$$= 162601 - 2500 + 1600 = 161701$$

$$\text{માટે સાચું પ્રમાણિત વિચલન} = \sqrt{\frac{\text{સાચો} \sum x_i^2}{n} - (\text{સાચો મધ્યક})^2}$$

$$= \sqrt{\frac{161701}{100} - (39.9)^2}$$

$$= \sqrt{1617.01 - 1592.01} = \sqrt{25} = 5$$

પ્રકીર્ણ સ્વાધ્યાય 15

1. આઠ અવલોકનોના મધ્યક અને વિચરણ અનુક્રમે 9 અને 9.25 છે, જો આમાંથી છ અવલોકનો 6, 7, 10, 12, 12 અને 13 હોય, તો બાકીનાં બે અવલોકનો શોધો.

2. સાત અવલોકનોના મધ્યક તથા વિચરણ અનુક્રમે 8 અને 16 છે. જો આમાંથી પાંચ અવલોકનો 2, 4, 10, 12, 14 હોય, તો બાકીનાં બે અવલોકનો શોધો.

3. 6 અવલોકનોના મધ્યક અને પ્રમાણિત વિચલન અનુક્રમે 8 અને 4 છે. જો પ્રત્યેક અવલોકનને 3 વડે ગુણવામાં આવે, તો પરિણામી અવલોકનોના મધ્યક અને પ્રમાણિત વિચલન શોધો.

4. જો n અવલોકનો $x_1, x_2, ..., x_n$ના મધ્યક $\bar{x}$ અને વિચરણ σ^2 હોય, તો સાબિત કરો કે અવલોકનો $ax_1, ax_2, ax_3, ..., ax_n$ના મધ્યક અને વિચરણ અનુક્રમે $a\bar{x}$ અને $a^2\sigma^2$ છે, $(a \neq 0)$.

5. વીસ અવલોકનોના મધ્યક અને પ્રમાણિત વિચલન અનુક્રમે 10 અને 2 છે. પુનઃતપાસ કરતાં માલૂમ પડ્યું કે અવલોકન 8 ખોટું છે. નીચે આપેલ પ્રત્યેક કિસ્સામાં સાચો મધ્યક અને સાચું પ્રમાણિત વિચલન શોધો.

   (i) ખોટા અવલોકનને દૂર કરવામાં આવે.

   (ii) તેને બદલે 12 મૂકવામાં આવે.

6. એક ધોરણના 50 વિદ્યાર્થીઓ દ્વારા ત્રણ વિષયો ગણિત, ભૌતિકશાસ્ત્ર અને રસાયણશાસ્ત્રમાં મેળવેલા ગુણનો મધ્યક અને પ્રમાણિત વિચલન નીચે પ્રમાણે છે :

વિષય	ગણિત	ભૌતિકશાસ્ત્ર	રસાયણશાસ્ત્ર
મધ્યક	42	32	40.9
પ્રમાણિત વિચલન	12	15	20

   કયા વિષયમાં સૌથી વધુ ચલન અને કયા વિષયમાં સૌથી ઓછું ચલન છે ?

7. 100 અવલોકનોના સમૂહનો મધ્યક અને પ્રમાણિત વિચલન અનુક્રમે 20 અને 3 છે. પછીથી જાણ થાય છે કે ત્રણ અવલોકનો 21, 21 અને 18 ખોટાં હતાં. આ ખોટાં અવલોકનોને દૂર કરવામાં આવે તો મધ્યક અને પ્રમાણિત વિચલન શોધો.

સારાંશ

◆ **પ્રસારનાં માપ :** વિસ્તાર, ચતુર્થક વિચલન, સરેરાશ વિચલન, વિચરણ, પ્રમાણિત વિચલન એ પ્રસારનાં માપ છે.

વિસ્તાર = મહત્તમ મૂલ્ય – ન્યૂનતમ મૂલ્ય

◆ **અવર્ગીકૃત માહિતી માટે સરેરાશ વિચલન**

$$\text{M.D.}(\bar{x}) = \frac{\sum |x_i - \bar{x}|}{n}, \quad \text{M.D. (M)} = \frac{\sum |x_i - \text{M}|}{n}$$

◆ **વર્ગીકૃત માહિતી માટે સરેરાશ વિચલન**

$$\text{M.D.}(\bar{x}) = \frac{\sum f_i |x_i - \bar{x}|}{\text{N}}, \quad \text{M.D.}(\text{M}) = \frac{\sum f_i |x_i - \text{M}|}{\text{N}}, \text{ જ્યાં, } \text{N} = \sum f_i$$

◆ **અવર્ગીકૃત માહિતી માટે વિચરણ અને પ્રમાણિત વિચલન**

$$\sigma^2 = \frac{1}{n} \sum (x_i - \bar{x})^2, \quad \sigma = \sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2}$$

◆ **અસતત આવૃત્તિ-વિતરણ માટે વિચરણ અને પ્રમાણિત વિચલન**

$$\sigma^2 = \frac{1}{\text{N}} \sum f_i (x_i - \bar{x})^2, \quad \sigma = \sqrt{\frac{1}{\text{N}} \sum f_i (x_i - \bar{x})^2}$$

◆ **સતત આવૃત્તિ-વિતરણ માટે વિચરણ અને પ્રમાણિત વિચલન**

$$\sigma^2 = \frac{1}{\text{N}} \sum f_i (x_i - \bar{x})^2, \quad \sigma = \frac{1}{\text{N}} \sqrt{\text{N} \sum f_i x_i^2 - \left(\sum f_i x_i\right)^2}$$

◆ **વિચરણ અને પ્રમાણિત વિચલન શોધવા માટેની ટૂંકી રીત**

$$\sigma^2 = \frac{h^2}{\text{N}^2} \left[\text{N} \sum f_i y_i^2 - \left(\sum f_i y_i\right)^2\right], \quad \sigma = \frac{h}{\text{N}} \sqrt{\text{N} \sum f_i y_i^2 - \left(\sum f_i y_i\right)^2},$$

જ્યાં, $y_i = \dfrac{x_i - \text{A}}{h}$

◆ **ચલનાંક (C.V.)** $= \dfrac{\sigma}{\bar{x}} \times 100, \bar{x} \neq 0.$

સમાન મધ્યકોવાળી શ્રેણીઓમાંથી જે શ્રેણીનું પ્રમાણિત વિચલન ઓછું હોય, તે વધારે સુસંગત અથવા ઓછી ફેલાયેલી હોય છે.

Historical Note

'Statistics' is derived from the Latin word 'status' which means a political state. This suggests that statistics is as old as human civilisation. In the year 3050 B.C., perhaps the first census was held in Egypt. In India also, about 2000 years ago, we had an efficient system of collecting administrative statistics, particularly, during the regime of Chandra Gupta Maurya (324-300 B.C.). The system of collecting data related to births and deaths is mentioned in Kautilya's *Arthshastra* (around 300 B.C.) A detailed account of administrative surveys conducted during Akbar's regime is given in *Ain-I-Akbari* written by Abul Fazl.

Captain John Graunt of London (1620-1674) is known as father of vital statistics due to his studies on statistics of births and deaths. Jacob Bernoulli (1654-1705) stated the Law of Large numbers in his book "Ars Conjectandi", published in 1713.

The theoretical development of statistics came during the mid seventeenth century and continued after that with the introduction of theory of games and chance (i.e., probability). Francis Galton (1822-1921), an Englishman, pioneered the use of statistical methods, in the field of Biometry. Karl Pearson (1857-1936) contributed a lot to the development of statistical studies with his discovery of *Chi square test* and foundation of *statistical laboratory* in England (1911). Sir Ronald A. Fisher (1890-1962), known as the Father of modern statistics, applied it to various diversified fields such as Genetics, Biometry, Education, Agriculture, etc.