એકમ–8	આંકડાશાસ્ત્રીય પદ્ધતિઓ (Statistical Techniques)

: રૂપરેખા :

8.0 ઉદ્દેશો (OBJECTIVES)

- **આ એકમના અભ્યાસ બાદ તમે નીચેની પરિસ્થિતિમાં હશો.**
- અંકગણિત સરેરાશ મધ્યક, મધ્યસ્થ અને બહુલક, મધ્યમવર્તી સ્થિતિના માપદંડો સમજશો.
- વિસ્તાર, ફરેફાર, પ્રમાણિત વિચલન અને ફેરફાર સહગુણાંકો તરીકે પ્રસારમાનના જુદાં જુદા માપદંડો સમજશો.
- સબસંબંધ અને નિયતસંબંધની પધ્ધતિઓ વિગતવાર જાણશો
- સમયશ્રેણીની વિગતોનું વિશ્લેષણ(analyse time series data)

8.1 પરિચય (INTRODUCTION)

આપણે આગળના એકમમાં વર્તમાન માહિતીને (ડેટા) કોષ્ટક અને આલેખના સ્વરૂપમાં રજુ કરવાની પધ્ધતિઓ સમજાવી. છતાં ઘણીવાર આપણને ઓછામાં ઓછા એક સારાંશમૂલ્યની જરૂર પડશે કે જે શ્રેણી દર્શાવતું હોય. ધારો કે આપણી પાસે ગ્રંથાલયની દરરોજ મુલાકાત લેનાર વ્યક્તિની માહિતી છે. આવી માહિતીને આપણે કોષ્ટકના સ્વરૂપમાં કે રેખા ગ્રાફના સ્વરૂપમાં દર્શાવી શકીએ છીએ. પરંતુ જો આપણે એક સારાંશની આકૃતિ જોઈતી હોય તો અંકગણિત સરેરાશ, આપણને સરેરાશ ગ્રંથાલયની મુલાકાત લેનાર વ્યક્તિઓની વ્યાખ્યા આપશે.

ગુણધર્મની દષ્ટિએ ચોક્કસ પધ્ધતિઓ હોવા છતાં આંકડાશાસ્ત્રીય પધ્ધતિઓ ગુણાત્મક માહિતી માટે વધુ અનુકૂળ છે. ગુણાત્મક માહિતી બે પ્રકારની હોય છે, અસતત અને સતત. આપણે બંને પ્રકારના ચલોની વિવિધ પધ્ધતિઓ સમજાવીશું. આપણે મધ્યવર્તી સ્થિતિના માપદંડોથી શરૂઆત કરીશું.

8.2 મધ્યવર્તી સ્થિતિના માપો (MEASURES OF CENTRAL TENDENCY)

મધ્યવર્તી સ્થિતિના માપદંડો આપણને સારાંશ આપે છે કે જે માહિતીનું મધ્યકેન્દ્ર દર્શાવે છે. મધ્યવર્તી સ્થિતિના મુખ્ય પાંચ માપદંડો છે. (1) અંકગણિત સરેરાશ, મધ્યક (2) મધ્યસ્થ (3) બહુલક (4) ભૌમિતિક સરેરાશ (5) સંવાદિત સરેરાશ.

આ પાંચમાંથી અંતિમ બે માપદંડો (4) ભૌમિતિક સરેરાશ અને (5) સંવાદિત સરેરાશના વિશિષ્ટ ઉપયોગ છે અને તેથી ઓછી વખત વપરાય છે. તેથી આપણે આ એકમમાં પ્રથમ ત્રણ માપદંડો અંગે ચર્ચા વિચારણા કરીશું.

આ માપદંડો સમજતા પહેલા આપણે ચોક્કસ સંકેતોને ઓળખીએ કે જેનો આપણે ઉપયોગ કરવાના છીએ.

પ્રમાણભૂત સંકેત x છે એ ચલ છે કે જેની કિંમતો x_1' x_2, x_3'....., x_N છે. એકમ 7 ના કોષ્ટક નંબર 7.3 વાચકોને આપેલ પુસ્તકોની સંખ્યાની માહિતી વિશે વિચારીએ. આપણે જાણીએ છીએ કે તે અસતત ચલ છે અને વાચકોને આપેલ પુસ્તકોની સંખ્યા 0 અને 5 ની વચ્ચે વિભાજિત થાય જેમ કે 0,1,2,3,4 અને 5 છે. દરેક અવલોકન માટે અનુક્રમ આવૃતિ 10,23,25,17,15, અને 10 છે.

ઉપરના ઉદાહરણમાં આપણે ઈસ્યુ થયેલા પુસ્તકોની સંખ્યાને x વડે અને તેની ધારેલી કિંમતોને $x_1, x_2, x_3, x_4, x_5, x_6$ વડે દર્શાવી શું અનુક્રમ આવૃતિ F_1, F_2, ... F_6 દર્શાવીશું. આપણે અવલોકનને લાક્ષણિક અવલોકન કહીશુ અને આવૃતિ fi સાથે xi વડે દર્શાવીશું. ઉદાહરણમાં ઈસ્યુ થયેલા પુસ્તકોની સંખ્યાનો વિસ્તાર 0 અને 5 ની વચ્ચે છે.

સતત ચલનાં કિસ્સામાં આપણે X_1, X_2, X_3... Xn વિભાગની મધ્યકિંમત અને અનુક્રમ આવૃતિ F_1, F_2... F_N લઈશુ.

8.2.1 અંકગણિત સરેરાશ / મધ્યક (Arithmetic Mean)

અંકગણિત સરેરાશને 'સરેરાશ' અથવા 'સરાસરી' પણ કહેવાય છે. તેને સરેરાશ ગણેલા ચલ ઉપર આડી લાઈન (બાર) વડે દર્શાવવામાં આવે છે. તેને અવલોકનોનો સરવાળો ભાગ્યા અવલોકનની સંખ્યા દવારા વ્યાખ્યાન્વિત કરવામાં આવે છે.

હવે, આવૃતિ વિતરણમાં આવેલા અવલોકનો માટે અંકગણિત સરેરાશ ગણીએ. જો $x_1, x_2, x_3 \dots x_n$ અવલોકનો હોય અને F_1, F_2, ... F_n અનુક્રમ આવૃતિ હોય તો

અંકગણિત સરેરાશ $\overline{X}$ (વચાય $\overline{x}$ બાર) અને $\overline{X} = \frac{1}{N}$ $(F_1 \times 1, F_2 \times 2 + \dots\dots$

$F_n \times n)$ વડે વ્યાખ્યાન્વિત થાય છે. તેને સંક્ષિપ્તમાં નીચે મુજબ લખી શકાય. $\overline{X}$

$= \frac{1}{N} \sum FiXi \dots (8.1)$

જ્યાં N= અવલોકનનોની કુલ સંખ્યા અને તે $\sum Fi$ ને બરાબર છે. સંજ્ઞા

$\sum$ Fi(વચાય 'સિગ્મા') ચલોનો સરવાળો દર્શાવે છે.

સતત મહિતીના કિસ્સામાં જ્યારે અવલોકનોને વિભાગીય ગાળામાં વર્ગીકૃત કરવામાં આવે છે ત્યારે અલગ અવલોકનો વિભાગીય ગાળામાં અલગ પારખી/ ઓળખી શકાતાં નથી. આ મુશ્કેલી ટાળવા માટે એવું ધારવામાં આવે છે કે દરેક અવલોકન કે જે વિભાગીય ગાળામાં આવે છે. તેની કિંમત વિભાગીય ગાળાની મધ્યકિંમત જેટલી છે.

ઉદાહરણ 8.1

નીચે દર્શાવેલી વિગતને આધારે વાચકોને ઈશ્યુ કરેલા કુલ પુસ્તકોની સરેરાશ શોધો.

પુસ્તકોની સંખ્યા	વાચકોની સંખ્યા
0	10
1	23
2	25
3	17
4	15
5	10
કુલ સંખ્યા	**100**

આપણે અંકગણિત સરેરાશ શોધવા માટે કોષ્ટક બનાવીશું

ઈસ્યુ પુસ્તકોની સંખ્યા (Xi)	ઈસ્યુ પુસ્તકોની સંખ્યા (Fi)	FiXi
0	10	0
1	23	23
2	25	50
3	17	51
4	15	60
5	10	50
કુલ	$\sum$ Fi= 100	$\sum$ FiXi= 234

બીજીરીતે, આપણે કોષ્ટક દ્વારા સમીકરણ (8.1) $X^{-} = \frac{1}{N} \sum Fi\,Xi$ માં કિંમત મૂકીશું.

અહીંયા N=100 and $\sum Fi\,Xi = 234$

તેથી $\overline{x} = \frac{1}{100} \times 234 = 2.34$

તેથી, ઈસ્યુ પુસ્તકોની સરેરાશ સંખ્યા 2.34

ઉદાહરણ 8.2

નીચે 100 વ્યક્તિ દ્વારા ખરીદાયેલા પુસ્તકો માટે માસિક ખર્ચ દર્શાવેલ છે. મહિનાનો સરેરાશ ખર્ચ શોધો.

વિભાગીય ગાળો /વર્ગ	આવૃતિ
100-200	12
200-300	18
300-400	28
400-500	19
500-600	13
600-700	07
700-800	03
કુલ	100

સતત માહિતીના કિસ્સામાં આપણે વર્ગની મધ્યકિંમત શોધવી પડે અને સમીકરણ 8.1 વાપરવું પડે.

વર્ગ	મધ્યકિંમત	આવૃતિ	FiXi
100-200	150	12	1800
200-300	250	18	4500
300-400	350	28	9800
400-500	450	19	8550
500-600	550	13	7150
600-700	650	07	4550
700-800	750	03	2250
કુલ		**$\sum$ Fi 100**	**$\sum$ FiXi=38600**

કોષ્ટકમાંથી ક્રમાનુંસાર કિંમત સમીકરણ 8.1 માં મૂકતાં

$$\bar{X} = \frac{1}{N} \sum FiXi$$

અહીંયા N=100 and $\sum$ FiXi = 38600

$$\text{તેથી } \bar{X} = \frac{1}{100} \times 38600 = 386$$

તેથી 100 વ્યક્તિગત ગ્રુપ દ્વારા ખરીદાયેલા પુસ્તકો માટે સરેરાશ માસિક ખર્ચ 386 રૂપીયા.

8.2 મધ્યસ્થ (Median)

મધ્યસ્થ આપણને શ્રેણીમાં એકદમ વચ્ચેનું અવલોકન આપે છે કે જેથી તેની બંને બાજુ અડધા અવલોકનો રહે ઉદાહરમો તરીકે, જો તમારી પાસે 2,5,9, 14 અને 20 એમ પાંચ અવલોકનો હોય, તો 9 એ વચ્ચેનું અવલોકન કહેવાય અને તેની દરેક બાજુ બે અવલોકનો આવેલ હોય તેથી ઉપરની શ્રેણીનો મધ્યસ્થ 9 છે. ધારો કે કોઈ શ્રેણીમાં છ અવલોકનો 3,8,15,25,35, અને 43 આવેલ છે. આવા સંજોગોમાં મધ્યસ્થ 15 અને 25 વચ્ચેની કોઈ સંખ્યા હશે કારણ કે બે અવલોકનો બંને બાજુ બાકી રહે છે. અહીં આપણે શ્રેણીની વચ્ચે રહેલા બે અવલોકનોનો સરેરાશ શોધીશું.

અહીંયા,તે $\frac{15+25}{2} = 20$ થશે. આ કિસ્સામાં મધ્યસ્થ 20 થશે.

જ્યારે અવલોકનની સંખ્યા ઘણી મોટી હોય અથવા માહિતીને અનુક્રમે વહેંચવામાં આવી હોય ત્યારે મધ્યસ્થ શોધવો સરળ નથી. જો આપણી પાસે N અવલોકનો હોય તો ત્યારે મધ્યસ્થ અવલોકન $\frac{N}{2}$ અવલોકનને અનુરૂપ હશે. આપણે સૌ પ્રથમ વિતરણમાં ઉત્તરોત્ર વધતી જતી આવૃતિ શોધીશું અને જેમાં $\frac{N}{2}$ અવલોકનો રહેલા હોય તેવા વર્ગને શોધીશું. આવા વર્ગ એ મધ્યસ્થવર્ગ કહેવાય. મધ્યકિંમત શોધવા માટે આપણે નીચેના સૂત્રનો ઉપયોગ કરીશું.

$$Md = lm + \frac{\frac{N}{2} - c}{fm} \times h \quad ...(8.2)$$

જ્યાં

Im = મધ્યવર્ગની ઉર્ધ્વસીમા

N = કુલ આવૃતિ

C = અગાઉના મધ્યસ્થવર્ગની વર્ગપ્રક્રિયા વધતી જતી આવૃતિ

Fm = મધ્યસ્થ વર્ગની આવૃતિ

h = મધ્યસ્થ વર્ગની પહોળાઈ

ઉદાહરણ 8.3

ઉદાહરણ 8.2 ની માહિતી ઉપરથી પુસ્તકો માટે થયેલા માસિક ખર્ચનો મધ્યસ્થ શોધો.

આ ઉદાહરણ ઉકેલવા માટે નીચે મુજબ આગળ વધીશું.

(1) વધતી જતી આવૃતિ વિતરણ શોધીશું

(2) મધ્યસ્થ વર્ગ શોધીશું

(3) સમીકરણ (8.2) નો ઉપયોગ કરીશું.

વર્ગલંબાઈ	**આવૃતિ**	**વધતી આવૃતિ**
100-200	12	12
200-300	18	30
300-400	28	58
400-500	19	77
500-600	13	90
600-700	07	97
700-800	03	100
કુલ	100	---

અહીંયા 100 અવલોકનો છે. તેથી 50 માં અવલોકનને અનુરૂપ મધ્યકિંમત કે જે વર્ગલંબાઈ 300-400 માં આવે છે. તેથી મધ્યસ્થ વર્ગલંબાઈ 300-400 ની વચ્ચે કોઈ કિંમતે મળશે. તેથી મધ્યવર્ગ 300-400 છે ની કિસ્સામાં lm=300, C=30, N=100, fm=28, h=100, સમીકરણ (8.2)નો ઉપયોગ કરતા મધ્યસ્થ કિંમત

$$Md = 300 + \frac{100/2 - 30}{28} \times 100 = Rs.371.43$$

8.2.3 માનાંક/બહુલક (Mode)

બહુલક એ ઉચ્ચ આવૃતિ સાથેનું અવલોકન છે. વિભાગીય માહિતી માટે બહુલક શોધવો સરળ છે. પરંતુ સતત માહિતીના કિસ્સામાં આપણે 'ઉચ્ચ વર્ગ' શોધવો પડે કે જે વર્ગલંબાઈ ઉચ્ચ આવૃતિ ધરાવતો હોય. આપણે જોઈ શકીએ છીએ કે વર્ગની પહોળાઈ સરખી છે નહીંતર ઉચ્ચ વર્ગલંબાઈ વધુ સંખ્યામાં અવલોકનો ધરાવશે અને નાની વર્ગલંબાઈમાં થોડોક જ અવલોકનો હશે. બહુલક સોધવા માટે નીચેના સૂત્રનો ઉપયોગ થાય છે.

$$M = lm + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h.....(8.3)$$

lm = ઉચ્ચવર્ગની અધઃસીમા

Δ_1 (=fm-fm-1)= વધતી જતી વર્ગની આવૃતિ અને ઉચ્ચવર્ગની આવૃતિ વચ્ચેનો તફાવત

Δ_2 (=fm-fm+1)= ઉચ્ચવર્ગની આવૃતિ અને પછી આવતી આવૃતિ વચ્ચેનો તફાવત.

h= ઉચ્ચ વર્ગની પહોળાઈ

ઉદાહરણ 8.4

ઉદાહરણ 8.2 માં આપેલ માહિતી પરથી બહુલક શોધો.

(1) ઉચ્ચવર્ગ શોધો વર્ગલંબાઈ ઉચ્ચ આવૃતિ ધરાવે છે. ઉચ્ચ વર્ગના કિસ્સામાં વર્ગલંબાઈ 30-400 છે.

(2) $lm, \Delta_1(= fm - fm - 1), \Delta_2(= fm - fm + 1)$ અને h

(3) સમીકરણ (8.3) નો ઉપયોગ કરો.

આપણે જોઈ શકીએ છીએ કે બહુલક

$$M = lm + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h = 300 + \frac{28-18}{(28-18)+(28-19)} \times 100$$

$= 352.63$

નોંધો કે એક જ માહિતી માટે મધ્યક, મધ્યસ્થ અને બહુલક અલગ-અલગ કિંમત મળે છે. ઉદાહરણ 8.2 માં દર્શાવેલ પુસ્તકો માટે મહિનાના ખર્ચ માટે આપણને મધ્યક, મધ્યસ્થ અને બહુલક. અનુક્રમ 386 રૂા.; 371.43 રૂા. અને 352 રૂા. મળે છે.

8.3 વિસ્તરણના માપદંડો/પ્રસારમાનનાં માપો (MEASURES OF DISPERSION)

મધ્યવર્તી સ્થિતિના માપો આપણને ગણમાહિતી માટે સારાંશનું મહત્વ પૂરુ પાડે છે. તેમનામાં ઘણી પરિસ્થિતિમાં આ માપદંડો માહિતીનું વિસ્તરણ દર્શાવતા નથી. ઉદાહરણ તરીકે, નીચે આપેલ માહિતીના ગણ તપાસો.

1) ગણ A : 2,5,17,17, and 44

2) ગણ B : 17,17,17,17, and 17

3) ગણ C : 13,14,17,17 and 24

આ ત્રણેય ગણ માટે મધ્યક, મધ્યસ્થ અને બહુલક ગણો. તમને દરેક માટે એકસરખી કિંમત મળશે. 17. હજુ આ ત્રણેય ગણ અલગ અલગ છે. જ્યારે ગણ B માં દરેક અવલોકનો સરખા છે. ગણ A માં બધા અવલોકનો અલગ અલગ છે. ચોક્કસપણે આપણને બીજા માપદંડની જરૂર પડશે કે જે અલગ માહિતીના વર્ણન માટે ગણતરી કરે.

વેર વિખેર શબ્દ આપણને માહિતીમાં જુદી-જુદી જાત કે પ્રકારની પ્રમાણ દર્શાવે છે તેની મહત્વની લાક્ષણિકતાએ છે કે જે સૌથી અંદરના અવલોકનન પ્રમાણ દર્શાવે છે. અવલોકનોના ગણનું વિસ્તરણ શૂન્ય હશે. જ્યારે તે ગણ B ની જેમ બધા અવલોકન સરખા હોય. એક અવલોકનથી બીજા અવલોકન વચ્ચેના મોટા તફાવતમાં મોટું અવલોકન જુદુ જુદુ હશે. ગણ A નું વિસ્તરણ ગણ C કરતા મોટુ હોવું જોઈએ. વિસ્તરણના માપદંડમાં આવી માહિતીમાં પરિવર્તન ક્ષમતા સમાવવી જોઇએ.

ત્યાં/થોડાં વિસ્તરણના માપદંડો છે. આપણે વિસ્તાર, મધ્યક વિચલન, તફાવત અને પ્રમાણિત વિચલન વિશે આ વિભાગમાં ચર્ચા કરીશું.

8.3.1 વિસ્તાર (Range)

વિસ્તારને મોટામાં મોટા અવલોકન અને નાનામાં નાના અવલોકન વચ્ચેના તફાવત તરીકે વ્યાખ્યાયિત કરવામાં આવે છે. તેથી ગણ A માં આપેલ માહિતી મુજબ, વિસ્તાર 44-2=42 તે જ રીતે ગણ B માં વિસ્તાર 17-17=0 અને ગણ c માં 11 છે. સમૂહિક (ગ્રુપ) માહિતીના કિસ્સામાં અલગ અવલોકન ઓળખી શકાતુ નથી. આવા કિસ્સામાં આપણે બે અધઃસીમા વચ્ચે તફાવત શોધીશું.

8.3.2 ફેરફાર / તફાવત (Variance)

ફેરફાર એ વિસ્તરણમાં બહોળા પ્રમાણમાં ઉપયોગ થતો માપદંડ છે. તેને સંજ્ઞા (વચાય સિગ્મા સ્કવર્ડ) વડે દર્શાવાય અને નીચે પ્રમાણે વ્યાખ્યાયિત થાય છે.

$$\text{ફેરફાર} = \sigma^2 = \frac{1}{N} \sum (X_1 - X^-)^2 \text{....(8.4)}$$

આવૃતિ વિતરણના કિસ્સામાં ફેરફાર નીચેના સૂત્ર વડે આપવામાં આવે છે.

$$\sigma^2 = = \frac{1}{\mathrm{N}} \sum fi (X_1 - X^-)^2 \text{.....(8.5)}$$

જ્યાં, $N = \sum fi$ અવલોકનની કુલ સંખ્યા

$i = 1$

ગણતરીને સરળ કરવા માટે આપણને નીચેના સૂત્રનો ઉપયોગ કરીશું.

$$\sigma^2 = = \frac{1}{\mathrm{N}} \sum fi (X_1 - X^-)^2 \text{.....(8.6)}$$

યાદ રાખો કે આપણે સમીકરણ (8.5) અથવા (8.6) નો ઉપયોગ કરીશું તો પણ સરખી જ કિંમત મળશે.

8.3.3 પ્રમાણિત વિચલન (Standard Deviation)

પ્રમાણિત વિચલન એ વિસ્તરણનો બહોળા પ્રમાણમાં ઉપયોગ થતો બીજો માપદંડ છે. તેને ફેરફારના ઘન વર્ગમૂળ તરીકે વ્યાખ્યાયિત કરવામાં આવે છે. અને 5 વડે દર્શાવવામાં આવે છે. યાદ રાખો કે પ્રમાણિત વિચલન ઋણ ન હોઈ શકે.

ઉદાહરણ 8.5

નીચેની માહિતીમાં વર્ગમાં વિદ્યાર્થીઓએ મેળવેલા ગુણનો વિસ્તાર છે. પ્રમાણિત વિચલન શોધો.

ગુણ	વિદ્યાર્થીઓની સંખ્યા
15-25	08
25-35	12
35-45	20
45-55	10
55-65	06
65-75	04
કુલ	**60**

પ્રમાણિત વિચલન ગણવા માટે આપણે નીચેના પગથિયાને અનુસરીશું.

1) વર્ગની મધ્ય કિંમતો શોધો
2) અંકગણિત મધ્યક શોધો
3) ફેરફાર શોધવા માટે સમીકરણ (8.5) અથવા (8.6)નો ઉપયોગ કરીશું.
4) જરૂરી સ્તંભો સાથે કોષ્ટક બનાવો. સમીકરણ (8.5) નો ઉપયોગ કરી આપણે એક કોષ્ટક બનાવ્યું છે.
5) તફાવત શોધો
6) તફાવતનું ઘન વર્ગમૂળ શોધો

ગુણ	વિદ્યાર્થીની સંખ્યા	મ.કિંમત xi	fixi	(xi–x)	$(xi–x)^2$	$fi(xi–x)^2$
1	2	3	4	5	6	7
15-25	8	20	160	-21	441	3528
25-35	12	30	360	-11	121	1452
35-45	20	40	800	-1	1	20
45-55	10	50	500	09	81	810
55-65	06	60	360	19	361	2166
65-75	04	70	280	29	841	3364
કુલ	60	--	2460	--	--	11340

સૌ પ્રથમ ઉપરના કોષ્ટકમાંથી આપણે એક ગણિત સરેરાશ/મધ્યક શોધીશું. પછી આપણે ફેરફાર શોધીશું પછી આપણે ફેરફાર શોધીશું.

$$\sigma^2 = \frac{1}{N}\sum fi(X_1 - X^{-})^2 = \frac{1}{60}\times 11340 = 189$$

તેથી પ્રમાણિત ફેરફાર..

$\sigma^2 = +\sqrt{189} = 13.75$ ગુણ

યાદ રાખો કે પ્રમાણિત વિચલન હંમેશા માપદંડના એકમ તરીકે પણ દર્શાવાય છે. તે વાસ્તવિક નંબર નથી. તફાવત સાથે મુશ્કેલી એ છે કે તેને માપદંડના એકમના વર્ગ તરીકે દર્શાવાય છે કે જેનો કોઈ અર્થ નથી.

8.3.4 તફાવતનો પ્રકીર્ણન ગુણાંક (Coefficient of Variation)

ઘણી વખત આપણે માહિતીની વિવિધ શ્રેણી વચ્ચે પરિવર્તનક્ષમતાને સરખાવવી પડે છે. જો એકમોને જુદાં જુદાં એકમોમાં માપવામાં આવે તો આપણે પ્રમાણિત વિચલન માટે જુદી જુદી કિંમતો મેળવીશું. ઉદાહરણ તરીકે ચાલો બે ગામડામાં ગૃહસ્થીની આર્થિક પરિસ્થિતિની સરખામણી કરીએ. નીચે બે ગામડાઓના ગૃહસ્થની માસિક કેલરી આયાતના માહિતીગત આંકડાઓ આપેલા છે.

	ગામ A	ગામ B
ગૃહસ્થીની સંખ્યા	817	561
સરેરાશ કેલરી આયાત	2417	2236
કેલરી આયાતનું (S.D)	418	232

પ્રમાણભૂત વિચલન

સમસ્યા એ છે કે આમાંથી એવું ગામ શોધો કે જેમાં કેલરી આયાત માટે સૌથી વધુ અસમાનતા હોય. આપણે જોઈ શકીએ કે ગામ A માં સૌથી વધારે સરેરાશ આયાત કેલરી છે પરંતુ તેમાં વધુ પ્રમાણમાં પ્રમાણિત વિટલન છે અને B ગામની

સરખામણીમાં ગૃહસ્થી વિચલન છે અને B ગામની સરખામણીમાં ગૃહસ્થીની સંખ્યા વધુ છે. તેથી ગામ B કરતા ગામ A માં ગરીબ ગૃહસ્થોની સંખ્યા વધારે છે. આવી પરિસ્થિતિની સરખામણી કરવા માટે આપણે ફેરફારના સહગુણક (C.V) નો ઉપયોગ કરીશું તેને આપણે સરેરાશ પ્રતિશત પ્રમાણિત વિચલન પ્રતિ એકમ તરીકે વ્યાખ્યાયિત કરીએ છીએ.

તફાવતનો સહગુણક (C.V) = $= \frac{6}{X} \times 100 \ldots (8.7)$

σ અને $\bar{X}$ પાસે માપદંડનો સરખો એકમ હોવાથી, ફેરફારનો સહગુણક એ વાસ્તવિક નંબર છે અને માપદંડના એકમની પસંદગીથી તે વિચલિત થતો નથી.

ગામ A માટે, C.V $= \frac{418}{2417} \times 100 = 17.29$

અને ગામ B માટે, C.V $= \frac{232}{2235} \times 100 = 10.38$

પરંતુ ગામ A માં તફાવતનો સહગુણક ગામ B માં તફાવતના સહગુણક કરતા વધુ હોવાથી ગામ B ની સરખામણીમાં ગામ A માં વધારે અસમાનતા છે.

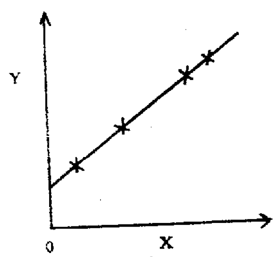
8.4 પારસ્પરિક સંબંધ / સહસંબંધ (CORRELATION)

અત્યાર સુધી આપણે માહિતીની એક જ લાક્ષણિકતા સાથે સંકળાયેલા હતા પરંતુ ત્યાં અમુક કિસ્સા એવા હોય છે. જ્યારે આપણે એક જ સમયે એક કરતાં વધારે લાક્ષણિકતાઓને પરીક્ષણ કરવામાં રસ ધરાવતા હતા. ઉદાહરણ તરીકે, તમને વાચકોએ વાંચેલ પુસ્તકોની સંખ્યા અને વાચકોની ઉંમર વચ્ચેના સંબંધનો અભ્યાસ કરવો ગમશે. આવી માહિતી કે જેમાં અભ્યાસ હેઠળની બે લાક્ષણિકતાઓ હોય તેને ચલ (bivarite) માહિતી કહેવાય છે. બે ચલો વચ્ચે માપ અથવા સહસંબંધની માત્રા શોધવા માટે માપદંડોમાંનો એક માપદંડ પારસ્પરિક સબંધ સહગુણક કહેવાય છે.

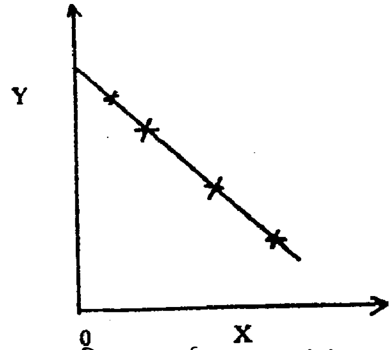
બે અથવા વધુ ચલોના સહ-તફાવતના વિશ્લેષણને સમાન્ય રીતે પારસ્પરિક સંબંધ કહેવાય છે જો બે લાક્ષણિકતાઓ અલગ પડે કે જેથી એક માપનું વિચલન બીજા વિચલન સાથે થાય છે. આવી લાક્ષણિકતાઓને પારસ્પરિક સંબંધ કહે છે. ઉદાહરણ તરીકે, પુરવઠો અને કિંમત, આવક અને ખર્ચ વગેરે વચ્ચે સબંધ છે. સહસંબંધના વિશ્લેષણની મદદ વડે આપણે બે ચલો વચ્ચેના સંબંધની કક્ષા એક આંકડામાં મળે છે.

8.4.1 પ્રકીર્ણન / વિકીર્ણ આલેખ (Scatter Diagram)

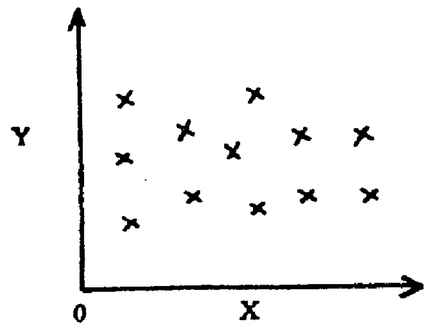
જો આપણે બે ચલો વચ્ચેના સંબંધ શોધવામાં રસ ધરાવતા હોય તો, સરળ રસ્તો એ છે કે કલ્પના દ્વારા (બિંદુ) ટપકાં આલેખ તૈયાર કરો કે જે પ્રકીર્ણન આલેખ કહેવાય છે. આ રીતનો ઉપયોગ કરી, આપેલ માહિતીને બિંદુ સ્વરૂપે આલેખ પેપર ઉપર દર્શાવો. ઉદાહરણ તરીકે x અને y ની કિંમતની દરેક જોડ માટે, આપણે બિંદુ દર્શાવીશું અને અવલોકનની સંખ્યા જેટલા જ બિંદુ મેળવીશું. હવે જુદા જુદા બિંદુઓના પ્રકિર્ણન જોતાં આપણે ખાતરી કરી શકીએ છીએ કેચલો જોડાયેલા છે કે નહી આલેખ પર દર્શાવાયેલા બિંદુઓ પ્રકીર્ણન મોટુ અને બે ચલો વચ્ચેનો સંબંધ ઓછો છે. જો બિંદુ નજીક હોય તો આલેખના બિંદુઓ રેખા પર આવે છે. ઉચ્ચ સંબંધોની માત્રા નીચે બે ચલોના સહસંબંધોના કેટલાક ઉદાહરણો આપેલા છે.



આકૃતિ 8.1 : પૂર્ણ ધન સહસંબધ r=1



આકૃતિ 8.2 પૂર્ણ ઋણ સહસંબંધ r=–1

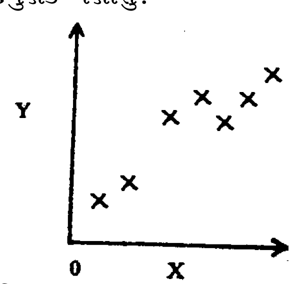


આકૃતિ 8.3 સહસબંધ નથી r=0

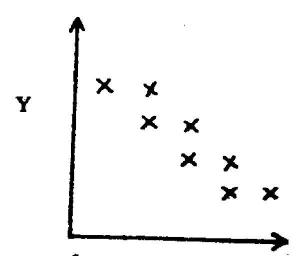
તમે જોઈ શકો છો કે આપણને બે ચલોના વચ્ચેના સંબંધોની માત્રાનો આછો ખ્યાલ આપે છે.

અહીં આપણે બે ચલો વચ્ચેના સબંધ વિશે વિચારીએ તો ત્યાં બે કરતાં વધુ ચલો વચ્ચે સંબંધ મળે. દસ ચલોના કિસ્સામાં આપણે દ્વિ-પરિમાણીય આલેખ પેપર ઉપર આપણે બિંદુ દર્શાવી શકીએ છીએ એટલે કે X- અક્ષ અને Y- અક્ષ સાથે. પરંતુ પ્રકીર્ણન આલેખની મર્યાદા છે કે જ્યારે બે કરતા વધુ ચલો હોય ત્યારે પ્રકીર્ણન આલેખ દોરી શકાતી નથી. બીજી વાત એ કે પ્રકીર્ણન આલેખ બે ચલો વચ્ચેનાં સંબંધની માત્રા વચ્ચેની ચોક્કસ આલેખ આપતો નથી.

ચલો વચ્ચેના સંબંધની ચોક્કસ વિસ્તારની આકૃતિ સુધી પહોચવા માટે અને એક જ સમયે બે કરતાં વધુ ચલોની વિચારણામાંથી બહાર આવવા માટે આપણે સહસંબંધ સહગુણક ગણીશું.



આકૃતિ 8.4 ધન સહસબંધની ઉચ્ચ પદવી



આકૃતિ 8.5 ઋણ સહસંબંધની ઉચ્ચ પદવી

8.4.2 પિયર્સનનો ગુણન પ્રધાત (પરિબળ ગુણાકાર) સહસંબધ (Pearson's Product Moment Correlation)

સહસંબંધ માપવાની કેટલીક રીતો પૈકી, વ્યવહારમાં (ગુણાકાર પધ્ધતિ) વધારે પિયર્સનનો ગુણાકાર સહસંબંધાક ઉપયોગી છે. તેને સંજ્ઞા વડે દર્શાવાય છે.

$$r = \frac{\frac{\Sigma}{i=1}(x_1 - x^-)(y_i - y^-)}{\sqrt{[\frac{\Sigma(x_1 - x^-)^-}{i=1}][\frac{z^n(y_1 - y^-)^2}{i=1}]}}$$

સહસંબધ ગુણક માટે બીજુ સૂત્ર :

$$r = \frac{\frac{\sum n}{i=1}(X_1 - X^-)(Y_i - Y^-)}{\frac{1}{n}6 \times^6 Y}$$

જ્યાં $x_1 = (x_i - x^-)$ *and* $Y_i = (y_i - y^-)$

સહ સંબંધના ગુણાંક ગણવા માટે નીચે પ્રમાણે પગથિયા અનુસરીશું.

(1) સરેરાશ x માંથી x નું વિચલન લો અને તેને n વડે દર્શાવો

(2) વિચલનનો વર્ગ કરો અને કુલ ટોટલ મેળવો.

(3) સરેરાશ y માંથી y નું વિચલન લો અને તેને n વડે દર્શાવો

(4) વિચલનનો વર્ગ કરો અને કુલ ટોટલ મેળવો.

(5) x અને y નો ગુણાકાર કરી ટોટલ મેળવો.

(6) ઉપરના સૂત્રમાં $\sum x_1 y_i . \sum x_1^2$ અને $\sum y_1^2$ ની કિંમતો મૂકો.

યાદ રાખો કે સહસંબંધ બે ચલો x અને y વચ્ચેના સુમેળની પદવી દર્શાવે છે. તે ફક્ત એટલું દર્શાવે છે કે એક ચલમાં થતો ફેરફાર બીજા ચલમાં થતા ફેરફારને આભારી છે. સહસંબંધ ગુણાંક ચલો વચ્ચેના સંબંધોમાં કારણો અને અસરો દર્શાવતા નથી. આપણે એમ ન કહી શકીએ કે x એ y ના કારણે છે અને y એ x ના કારણે છે. બીજી યાદ રાખવા જેવી વાત એ છે કે r=+0.6 એ r=0.6 કરતા વધારે દર્શાવતો નથી. ઋણ નિશાની કે ધન નિશાની સંબંધોની દિશા દર્શાવે છે જો r = 0.6 હોય તો, જાણીતું પરિણામ છે કે x માં વધારો એ y થતા વધારા બરાબર છે. બીજુ બાજુ r = -0.6 એ x અને y વચ્ચે વિરોધી સંબંધ દર્શાવે છે. ત્રીજી વાત સહસંબધ એ માપ અને ઉદ્ભવના તફાવતમાં તટસ્થ રહે છે. આનો અર્થ એ થયો કે 8 માં ઉદાહરણમાં ફેરફારમાં દા.ત. જો x ચલને કોઈ રકમ વડે ભાગવામાં આવે તેને 5 કહો. અને y ચલ માટે 2 કહો. આપણે r માટે એક સરખી કિંમત મેળવીશું. તેને માપમાં તફાવત કહેવામાં આવે છે. ક્ષણ માટે જો તમે ઊંચાઈને સે.મી. માં. ફેરવો તો તેની ઈંચના સહસંબંધના ગુણાંકમાં ફેરફાર ન થાય. તેવી જ રીતે જો તમે ઉદ્ભવ બદલી નાખો એટલે કે બધા અવલોકનમાં કંઈક ઉમેરો અથવા બાદ કરો, તો પણ સહસંબંધ ગુણાંક ન બદલાય. તમે તમારી જાતે શોધી શકશો કે x માંથી 10 બાદ કરતા અને y માંથી 5 બાદ કરતાં અને સહસંબંધ ગુણાંક ગણો ચોથી વાત, સહસંબંધ પ્રકીર્ણન ગુણાંક ચલો વચ્ચેનો રેખીય સંબંધ દર્શાવે છે. ઉચ્ચ ક્રમના સંબંધ માટે સહસંબંધ ગુણાંક સંબંધની યોગ્ય માત્રા દર્શાવતા નથી. ઉદા. તરીકે બે ચલો x અને y માટે જો x અને y ની દરેક કિંમત માટે જો x=y_2 ત્યારે r કદાચ શૂન્ય બનશે. પરંતુ આનો અર્થ એમ નહિ કે x અને y સંકળાયેલા નથી. તેથી આપણે કહી શકીએ કે જ્યારે બે ચલો સ્વતંત્ર હોય r = 0 પરંતુ r = 0 ના જાણ પુરતું આપણે એવું ન માની શકીએ કે બે ચલો સ્વતંત્ર છે.

ઉદાહરણ લઈ r ની કિંમત શોધીએ.

ઉદાહરણ 11 :

નીચેના કોષ્ટકમાં દસ બાળકોના વજન અને ઊંચાઈની માહિતી દર્શાવી છે. ગુણન પ્રઘાત સહ સંબંધનો ગુણાંક શોધો.

ઊંચાઈ (સેમી)	વજન(કિગ્રા)	ઊંચાઈ (સેમીમાં)	વજન(કિગ્રા)
110	26	140	38
110	21	135	30
125	22	130	30
130	24	140	40
145	36	135	43

હવે ઊંચાઈને (સેમીમાં) x વડે અને વજનને (કિગ્રામાં) Y વડે દર્શાવીશું.

X_i	Y_1	$X_1(X_1-X^-)$	$Y_1(Y_i-Y^-)$	X_iY_1	X_1^2	Y_i^2
110	26	-20	-5	100	400	25
110	21	-20	-10	210	900	100
125	22	-5	-9	45	25	81
130	24	00	-7	00	00	49
145	36	15	5	75	225	25
140	38	01	07	70	100	49
135	30	05	-1	-5	25	01
130	30	00	-1	00	00	01
140	40	10	09	90	100	81
135	43	05	12	60	25	144
1300	310	--	--	--	645	1300

8.4.3 સ્પિયરમેનનો સહસંબંધ (Spearman's Rank Correlation)

અગાઉના એકમમાં, માહિતીના પ્રકાર ક્રમાંક સહસંબંધ દર્શાવતી વખતે આપણે માપદંડના ગુણોત્તર પ્રમાણ અને ક્રમિક પ્રમાણ વચ્ચે અંતર ઓળખવું રહ્યું. વ્યક્તિની ગુણાકાર કક્ષાની ઝડપ સાથે પારસ્પરિક સંબંધ લાગુ પાડી શકાય. ગુણાકાર પરિબળ સહ સંબંધ ત્યારે જ લાગુ પાડી શકાય. જ્યારે માહિતીને ગુણોત્તર પ્રમાણમાં ગણવામાં આવી હોય. પરંતુ જ્યારે અવલોકનના ખરેખર પ્રમાણને બદલે આપણી પાસે ફક્ત નંબર જ છે ત્યારે વ્યક્તિને બદલે આપણી પાસે ફક્ત નંબર જ છે ત્યારે વ્યક્તિનો ‘r’ અયોગ્ય બને છે. આવા કિસ્સામાં સ્પિયરમેનનો ક્રમ સહસંબંધ (r_s) નો ઉપયોગ થાય છે. r ગણવા માટેની પધ્ધતિ ઘણી સરળ છે. સૌ પ્રથમ ક્રમ નક્કી કરો અથવા n અવલોકનોના x અને y ની અનુક્રમ શ્રેણીમાંથી i નું અવલોકન નક્કી કરો. બીજા સ્ટેપમાં x માંનો ક્રમ અને y માનો ક્રમનો તફાવત નક્કી કરો. આ તફાવતને D_1 કહો. ત્રીજા સ્ટેપમાં નીચેનું સૂત્ર વાપરો.

$$r_s = 1 - \frac{6\sum_{i=1}^{n} D_i^2}{n(n^2-1)}$$

સ્પિયરમેનનો ક્રમાંક સહસંબંધનો વિસ્તાર ₹1 થી -1 પણ છે. તેથી ધન કિંમતો સીધી રીતે ચલો વચ્ચેનો સંબંધ દર્શાવો છે. જ્યારે ઋણ કિંમતો વિરોધી સંબંધ દર્શાવે છે. r = 0 કિંમત ચલો વચ્ચેના સંબધની ગેરહાજરી દર્શાવે છે. એક સૂચના છે કે, સ્પિયરમેનનો ક્રમાંક સહ સંબંધ ફક્ત ઉપયોગ કરી શકાય નહી કારણ કે પિયર્સનનો ગુણન પ્રધાત સહસબંધનો ગુણાંક કરતાં તેની ગણતરી સહેલી છે. આપણે ચોકક્સપણે કહી શકતો નથી કે x માં નો તફાવત એ y માં થતા તફાવત

સાથે સંકળાયેલ છે કારણ કે ક્રમમાં એકસરખો તફાવત એ લાક્ષણિકતાઓના એક સરખો તફાવત દર્શાવતા નથી.

ઉદાહરણ 12 :

નીચેની માહિતીમાં 12 સામયિકોનો ક્રમ આપ્યો છે જેમાં ક્રમાંક સહસંબંધ ગુણાંક ગણવા માટે બે જુદી જુદી રીતોનો ઉપયોગ કરેલો છે.

ક્રમ	x ના પ્રકાશનના ક્રમ	Y નાં નિર્દેશન ક્રમ	Di	Di^2
1	1	12	-11	121
2	2	09	-7	49
3	3	06	-3	9
4	4	10	-6	36
5	5	03	₹2	04
6	6	05	₹1	01
7	7	04	₹3	09
8	8	07	₹1	01
9	9	08	₹1	01
10	10	02	₹8	64
11	11	11	00	00
12	12	1	₹11	121
				$\Sigma D^2 = 416$

$r_s = 1 - \frac{6\sum D_i^2}{n(n^2-1)}$ અહીંયા ; $\sum D_i^2 = 416,\ n = 12$

તેથી $r_s = 1 - 1.454 = 0.454$

- **તમારી પ્રગતિ ચકાસો**

(1) નીચેની માહિતીનું પ્રમાણિત વિચલન શોધો

આંકડાશાસ્ત્રમાં ગુણ	વિદ્યાર્થીઓની સંખ્યા
0-20	05
20-40	10
40-60	15
60-80	08
80-100	02

(2) નીચેની કોઠામાં માહિતીમાં 10 વિદ્યાર્થીઓએ આંકડાશાસ્ત્રમાં અને ગણિતમાં મેળવેલ ગુણ છે. સહ સંબંધનો ગુણાંક શોધો

વિદ્યાર્થીનો નંબર	ગણિતના ગુણ	આંકડાશાસ્ત્રના ગુણ
1	52	48
2	60	40
3	45	38
4	38	28

5	35	42
6	62	65
7	68	53
8	28	25
9	50	28
10	62	33

3) નીચેના શબ્દો માટે સમીકરણ લખો

a) પ્રમાણિત વિચલન

b) તફાવત

c) ગુણન પ્રઘાત સહસંબંધ ગુણાંક

d) ક્રમાંક સહસંબંધ

નોંધ : (1) તમારા જવાબ નીચે આપેલી જગ્યામાં લખો

(2) એકમને અંતે આપેલા જવાબો સાથે તમારા જવાબની ચકાસણી કરો.

....

....

....

....

....

....

8.5 નિયતસંબંધ વિશ્લેષણ (REGRESSION ANALYSIS)

અગાઉના એકમમાં આપણે જોયું કે સહસંબંધ સહગુણાંક એ કારણ અને અસર વચ્ચેના સંબંધોનું પ્રતિબિંબ નથી. આ એકમમાં નિયતસંબંધ વિશ્લેષણની ચર્ચા કરીશું. તે દર્શાવે છે કે એક ચલ કારણ છે અને બીજો તેની અસર છે. સામાન્ય રીતે આપણે કહી શકીએ કે ચલો બે પ્રકારના હોય છે. આધારિત ચલો અને સ્વતંત્ર ચલો. સ્વતંત્રચલ એ કારણ છે અને આધારિત ચલ એ અસર છે. નિયત સબંધ વિશ્લેષણ આંકડાશાસ્ત્રનું મહત્વનું સાધન છે કે જે ચલો અને સ્વતંત્ર ચલોની જાણીતી કિંમતોથી આધારિત ચલોની અજ્ઞાત કિંમતો વચ્ચેનો સંબંધ સમજાવામાં મદદરૂપ થાય છે.

ધારી લો કે પુસ્તકાલયમાં સરક્યુલેશન પુસ્તકોની સંખ્યા, એ વાંચકોની સંખ્યા સાથે સંકળાયેલ છે. ઉદાહરણ તરીકે, તેને વાંચકોની વધતી જતી સંખ્યા તરીકે ધારી લો અહીંયા વાચકોની સંખ્યા એ સ્વતંત્ર ચલ છે અને સરક્યુલેશનમાં પુસ્તકોની સંખ્યા એ આધારિત ચલ છે. હવે આધારિત ચલને Y તરીકે દર્શાવીએ અને સ્વતંત્ર ચલને X તરીકે દર્શાવીએ. નિયતસંબંધ વિશ્લેષણમાં આપણે સમયગાળા ઉપર માહિતી ભેગી કરીશું અથવા સમયના બિંદુએ રજુઆત કરીશુ. ધારી લઈએ કે X અને Y માં અવલોકનોની 'n' જોડીઓ ભેગી કરેલી છે. પછીના સ્ટેપમાં X અને Y વચ્ચેનો સંબંધ શોધવો પડશે.

X અને Y વચ્ચેનો સંબંધ ઘણા સ્વરૂપમાં હોય છે. સામાન્ય ટેવ એ છે કે સંબંધોને અમુક ગાણીતીય સમીકરણના સ્વરૂપમાં દર્શાવવા. આ સમીકરણોમાં સૌથી સરળ રેખીય સમીકરણ છે. આનો અર્થ એમ થયો કે X અને Y વચ્ચેનો સંબંધ સુરેખાના સ્વરૂપમાં હશે. જ્યારે સમીકરણને વક્ર સ્વરૂપે દર્શાવવામાં આવે તો નિયતસંબંધને અરેખીય અથવા વક્રરેખીય કહેવામાં આવે છે.

અહીંયા પ્રશ્ન એ થાય છે કે, આપણે કેવી રીતે સમીકરણના સ્વરૂપને ઓળખી શકીશું ? અહીંયા કોઈ ઝડપી અને કડક/સખત નિયમ નથી. સીમકરણનું સ્વરૂપ સંશોધકની વિચાર શક્તિ અને ધારણા ઉપર આધાર રાખે છે.

ગમે તેમ પણ, સંશોધક વિકીર્ણ આલેખ દોરવા માટે આલેખ પેપર પર X અને Y ચલો દર્શાવી શકે છે. વિકીર્ણ આલેખ દ્વારા આલેખ પેપર પરના બિંદુઓના સ્થાન સમીકરણ સ્વરૂપ શોધવામાં મદદ કરે છે. જો બિંદુઓની સંખ્યા સુરેખ પર વતા ઓછા પ્રમાણમાં હોય તો રેખીય સમીકરણ ધારવામાં આવે છે. જો બિદુઓ સુરેખા પર ન હોય અને વક્રના સ્વરૂપમાં હોય તો યોગ્ય અરેખીય સમીકરણ કે જે વિસ્તરણને દર્શાવતું હોય તેમ ધારવામાં આવે છે.

સંશોધકે વધુ એક ધારણા કરવી પડે છે. એટલે કે આધારિત અને સ્વતંત્ર ચલોની ઓળખ. આ બધુ સંશોધન વિશ્લેષણના તર્ક અને હેતુ પર આધાર રાખે છે કે Y એ X ઉપર અવલંબે છે અને X એ Y ઉપર અવલંબે છે. તેથી સરખી માહિતી (a) થી જ્યારે X એ Y ઉપર આધારિત ધારવામાં આવે ત્યારે ત્યાં બે નિયતસંબંધ રેખાઓ હોઈ શકે. તેને 'Y ઉપર X' રેખા દર્શાવાય અને X એ Y રેખા ઉપર એમ દર્શાવાય. રેખીય સમીકરણનું Y સાથેનું સ્વાયત ચલનું ઉદાહરણ લઈએ અને સ્વતંત્ર તરીકે X લઈએ.

$Y= 3+2X$

X ની જુદી જુદી કિંમતો મૂકતાં, આપણે Y ની કિંમત શોધી શકીએ. ઉ.દા. જ્યારે $X =1$, ત્યારે $Y=5$, જ્યારે $X = 2$, ત્યારે $Y = 7$ જો આપણે આવા બિંદુઓ (1,5)(2,7) ની જોડ આલેખ પેપર પર દર્શાવીએ તો આપણને સુરેખા મળશે.

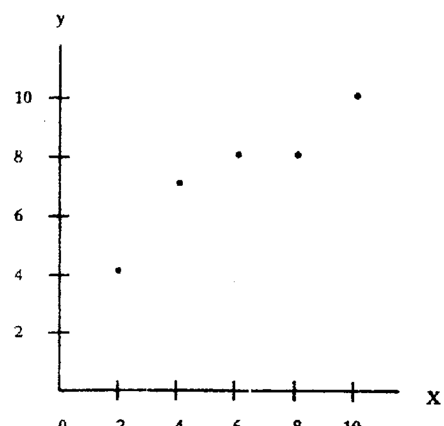
ઉપરના સંબંધને સામાન્ય સ્વરૂપે નિરૂપણ કરતા એવું કરી શકાય કે Y નું X ઉપરનું રેખીય સમીકરણ $Y=a+bX$ ના સ્વરૂપમાં મળે છે, જ્યાં a અને b અચળ છે. તેવી જ રીતે અરેખીય સમીકરણોને ઘણા સ્વરૂપમાં રજુ કરી શકાય. સામાન્ય ઉદાહરણ એ છે કે $Y=a+bX + CX2$

8.5.1 નિયત સંબંધ રેખા (Linear Regression)

નીચેની માહિતી વિચારો : અઠવાડિયાના દિવસો દરમિયાનની વાચકોની સંખ્યા અને (ઈસ્યુ કરેલી) આપેલી પુસ્તકોની સંખ્યા અંગેની માહિતી આપેલ છે.

પુસ્તકાલયમાં વાચકોની સંખ્યા (X)	6	2	10	4	8
(ઈસ્યુ કરેલા) આપેલ પુસ્તકોની સંખ્યા (Y)	8	4	10	7	8

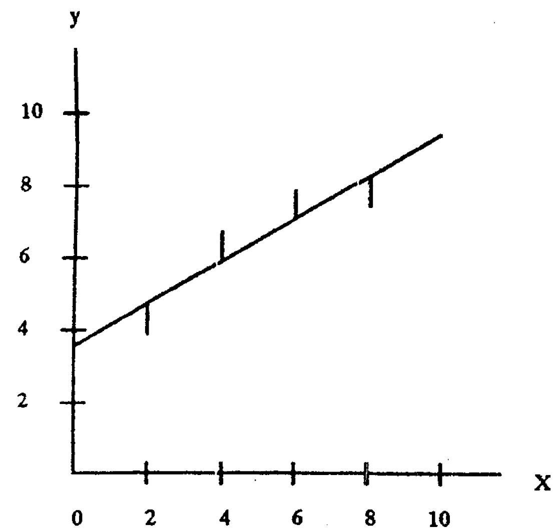
જો આપણે માહિતીને આલેખ પેપર ઉપર દર્શાવીશું તો વિસ્તરણ આકૃર્તિનીચે દર્શાવેલ આકૃતિ નીચે દર્શાવેલ આકૃતિ 8.6 જેવી મળશે.



આકૃતિ 8.6 પુસ્તકાલયમાં વાચકોની સંખ્યા અને ઈસ્યું કરેલા પુસ્તકોની સંખ્યાનો આલેખ

આલેખ પરથી એ તો નિશ્ચિત છે કે બિંદુઓ સીધી રેખા ઉપર આવેલ નથી પરંતુ તે ઉપર વધતુ વલણ દર્શાવે છે કે જ્યાં સીધી રેખા બંધબેસે છે. જો આપણે વિખેરણ બિંદુ સાથે સુરેખા દોરીએ તો આલેખ આકૃતિ 8.7 માં દર્શાવ્યા મુજબ મળે.

પ્રત્યાગમન રેખા અને અવલોકન વચ્ચેનો તફાવત એટલે 'ભૂલ' ઉ.દા. તરીકે X ની કિંમત 2 હોય તો Y ની કિંમત 4 મળે આને અવલોકન કિંમત કહેવાય.



આકૃતિ 8.7

પરંતુ પ્રત્યાગમન રેખાY ની કિંમત 4.8 ની Xની કિંમત 2 વિરુધ્ધ Y ની કિંમત દર્શાવે છે. આ કિંમત કે જેને પ્રત્યાગમન રેખા પરથી ગણવામાં આવી છે તેને આશરે કિંમત કહે છે. અવલોકન કેટલી કિંમત અને ધારેલી કિંમત વચ્ચેના તફાવતને ભૂલની કિંમત તરીકે દર્શાવાય છે તેથી આપણે જોઈ શકીએ છીએ કે અવલોકેલી કિંમત એ ધારેલી કિંમત અને ભૂલ અને કિંમતનો સરવાળો છે.

પ્રત્યાગમન રેખામાં આપણું લક્ષ્ય ભૂલની કિંમતને ઓછું કરવાનું છે. આવું મોટાભાગે લઘુવર્ગની રીતે કરવામાં આવે છે. લઘુ વર્ગની રીત $\Sigma\ 8^2$ ની કિંમત ઓછી કરે છે. જ્યાં E એ અવલોકેલી કિંમત અને ધારેલી કિંમત વચ્ચેનો તફાવત છે. આપણે અહીંયા આ રીતની વિગતે ચર્ચા નહીં કરીએ. તેના બદલે લઘુ વર્ગની રીત ઉપર આધારિત બે સમીકરણો અને સામાન્ય સમીકરણો નીચે આપેલા છે.

$\Sigma y = na + b\Sigma\times$

$\Sigma \times y = a\Sigma \times b\Sigma\times 2$

અંગુઠાના નિયમ મુજબ આ સામાન્ય સમીકરણને a અને b ના સહગુણકોને રેખીય સમીકરણ સાથે ગણીને અને બધા અવલોકનોનો સરવાળો કરીને મેળવી શકાય છે.

અહીંયા રેખીય સમીકરણ $y = a + b\ x$ છે. પ્રથમ સામાન્ય સમીકરણ એ રેખીય સમીકરણ $y = a + bx$ ને સાદુરૂપ આપે છે.

$Sy = a\ x\ sb\ x$ અથવા $Sy = na \times bsx$

બીજુ સામાન્ય સમીકરણ એ રેખીય સમીકરણને X વડે ગણવાનું અને બધા અવલોકનોનો સરવાળો કરવાનો.

$S\ x\ Y = SaX + Sbx^2$ and $S\ x\ Y = aSx + bSx^2$

તે સ્પષ્ટ છે કે આ સમીકરણોમાં બધા જ પદો સંખ્યાત્મક છે કે જે a અને b સિવાય માહિતીમાંથી ગણવામાં આવ્યો છે. આધારિત ચલોની આંકેલી કિંમત મેળવવા માટે a અને b ની કિંમત ગણવી પડે તે નીચે મુજબ મેળવી શકાય.

કોષ્ટક 8.1 પરથી જોઈ શકાય છે કે x અને y ઉપરની માહિતી પ્રથમ બે સ્તંભમાં આપેલ છે. કોષ્ટકમાં આપેલ બે ક્રમિક સ્તંભો ગણતરી આપે છે કે જે ઉપર આપેલ મુજબ બે સામાન્ય સમીકરણો ઉકેલવા જરૂરી છે. y ની ધારેલી કિંમત અને ભૂલની કિંમતો અંતિમ બે સ્તંભોમાં આપેલી છે.

કોષ્ટક 8.1 (Regression) : પ્રત્યાગમન વિશ્લેષણ માટે માહિતી ગણતરી

	(1)	(2)	(3)	(4)	(5)	(6)
	X	Y	X2	XY	Yની	ભૂલ
	-	-	--	---	ધારેલી કિંમત	---
	6	8	36	48	7.4	0.6
	2	4	4	8	4.8	-0.8
	10	10	100	100	10.0	0.0
	04	07	16	28	6.1	0.9
	08	08	64	64	8.7	-0.7
કુલ	30	37	220	248	37	0.0

સામાન્ય સમીકરણ મુજબ :

Y = na + bx ....(8.8)

$\Sigma xY = a\Sigma x + b\Sigma x^2$ ...... (8.9)

આપણે કોષ્ટક 1 માંથી અનુક્રમ કિંમતો મૂકીશું.

તેથી,

37 = 5a + 30b ... (8.10)

284 = 30a + 220b ... (8.10)

જો આપણે સમીકરણ (3) ને વડે ગુણી ગુણાકારને સમીકરણ (4) માંથી બાદ કરીએ તો.

248 = 30a + 22ob

222 = 30a + 180 b

40b = 26

અથવા b = 0.65

b ની કિંમત સમીકરણ (8.10) માં મૂકતાં,

37 = 5a + 30 × 0.65

અથવા 5a = 17.5 અથવા a=3.5

તેથી પ્રત્યાગમન રેખા,

Y = 3.5 + 0.65 × ....(8.12)

જો આપણે X ની કિંમતને પ્રત્યાગમન રેખામાં મૂકીએ તો સમીકરણ (8.12) તો આપણે y ની ધારેલી કિંમત મેળવીશું. ઉદાહરણ તરીકે જ્યારે X=2

Y = 3.5 + 0.65 × 2 = 4.8

પરંતુ આપણે X = 2 ની વિરુધ્ધ Y ની ધારેલી કિંમત 4 છે. ધારેલી કિંમત્ ને અવલોકેલી કિંમત વચ્ચેના તફાવતને ભૂલ (4-4.8 = -0.8) એ 'e' છે. કહેવાય. X = 2

Y અને e ની ધારેલી કિંમત કોષ્ટક 8.1 માં સ્તંભ (5) અને (6) માં આપેલી છે.

એ નોંધ લેવી કે નમૂના માટે સરવાળાની ભૂલ શૂન્ય છે.

એટલે કે Se = 0

ગણતરીના હેતુ માટે આપણે a અને b ની કિંમત શોધવા માટે નીચેના સમીકરણનો ઉપયોગ કરીશું.

X = (X-X)

Y = (Y-Y)

XY= (X-X) (Y-Y)

જ્યાં $\overline{X}$ અને $\overline{Y}$ એ X અને Y ના એંકગણિતીય મધ્યકો છે. આ સમીકરણ આપણને નીચે મુજબ સૂત્ર આપશે.

$$b = \frac{\sum XY}{\sum X^2} \text{ ...... (8.13)}$$

$$a = \overline{Y} - b\overline{X} \text{ ....(8.14)}$$

હવે, આ સમીકરણો સામાન્ય સમીકરણોમાંથી મેળવવામાં આવે છે. આપણે આ રીતમાં 'a' અને 'b' ની કિંમત પણ સરખી જ મેળવીશું.

ગણતરીના પગલાં નીચે મુજબ છે.

1) X અને Y ની કિંમતો શોધો

2) X= (X-X) અને Y=(Y-Y) શોધો

3) XY ની કિંમત શોધો

4) X^2 ની કિંમત શોધો

5) ઉપરના સમીકરણ (8.13) અને (8.14) માં સમીકરણનો ઉપયોગ કરે 1 માં આપેલ ઉદાહરણ ઉપરનું સમીકરણ ઉપયોગમાં લેતા આપણે કોશ્ટક 2 માં આપેલ માહિતી મેળવીશું.

કોષ્ટક 8.2 પ્રત્યાગમન સમીકરણની ગણતરી : ટૂંકી રીત

X	Y	X=(X-X)	Y(Y-Y)	XY	X^2
6	8	0	0.6	0	0
2	4	-4	-3.4	13.6	16
10	10	4	2.6	10.4	16
04	07	-2	-0.4	0.8	04
08	08	2	0.6	1.2	04
30	37			26.0	40

$$x^{-}\frac{1}{n}\sum \times {}^{30}\!/_{5} = 6. \quad b = \frac{\sum XY}{\sum x^2} = {}^{26}\!/_{40} = 0.65$$

$$y^{-}\frac{1}{n}\sum y \times {}^{37}\!/_{5} = 7.4. \quad y^{-}\frac{1}{n}\sum y \times {}^{37}\!/_{5} = 7.4.$$

અગાઉ દર્શાવ્યા મુજબ a અને b ની કિંમતો સરખી છે.

8.5.2 નિયત સંબંધ ગુણાંકની લાક્ષણિકતાઓ (Properties of Regression Coefficient)

ગુણાંક ‘b’ ને નિયત સંબંધ ગુણાંક કહેવાય. નોંધો કે X અને Y ની માહિતી ઉપરથી આપણે બે નિયતસંબંધ રેખા દોરી શકીએ છીએ

(a) X ઉપર Y રેખા ; Y = a + bx

(b) Y ઉપર X રેખા ; X = $\alpha + \beta$Y

બે સહઅવયવો b અને β કેટલીક રસપ્રદ લાક્ષણિકતાઓ સ્પષ્ટ કરે છે. સૌપ્રથમ બંને નિયતસંબંધ ગુણાંકનો ગુણાકાર એ r ના વર્ગ બરાબર છે. (સબસંબંધ ગુણાંક) એટલે કે βY = r^2.

તેથી એકવાર બંને નિયતસંબંધ ગુણાંક જાણ્યા પછી આપણે r^2 ની કિંમત શોધી

શકીશું. r^2 નું વર્ગમૂળ લેતા આપણે r મેળવીશું. બીજી રીતે જો નિયતસંબંધ ગુણાંક ઋણ હોય તો નિયત સંબંધ ગુણાંક પણ ઋણ થશે.

જો નિયતસંબંધ ગુણાંક ધન હોય તો સહસંબંધ પણ ધન હશે. ત્રીજુ તમે જાણો છો કે ;

$-1<r<+1$ એટલે કે, r ની કિંમત -1 અને +1 વચ્ચે હશે.

તેથી r^2 ની કિંમત શૂન્ય અને +1 વચ્ચે હશે.

નિયત સંબંધ ગુણાંક ની કિંમત સિમિત હોઈ શકે પરંતુ એક નિયત સંબંધ ગુણાંક 1 થી વધારે અને બીજો નિયત સંબંધ ગુણાંક 1 થી ઓછો હોય. બંને પ્રત્યાગમન ગુણાંક સરખા ન હોઈ શકે. તે દર્શાવે છે કે બંનેનો ગુણાકાર એટલે કે r^2 એક સમાન ન હોઈ શકે. સહસબંધના ગુણાંક વર્ગને નિર્ણાયક અને મહત્વની લાક્ષણિકતાનો ગુણાંક કહેવાય જો r^2 એ નિશ્ચાયકનો ગુણાંક એ 1 ની નજીક છે કે જેથી આપણે અનુમાન કરી શકીએ કે સ્વતંત્ર ચલ આધારિત ચલમાં ચલન સમજાવે છે. જો નિશ્ચાયકનો ગુણાંક શૂન્યની નજીક હોય તો સ્વતંત્ર ચલ આધારિત ચલમાં ફેરફાર સમજાવશે નહિં.

8.5.3 અરેખીય નિયત સંબંધ (Non-Linear Regression)

આપણે અગાઉના પેટા વિભાગમાં બે ચલોનો સમાવેશ કરનાર સામાન્ય રેખીય નિયત સંબંધની ચર્ચા કરી. એક આધારિત ચલ અને બીજો સ્વતંત્ર. નિયતસંબંધ એક આધારિત ચલને સમાવી શકે છે. આવા કિસ્સાને બહુવિધ નિયતસંબંધ કહે છે.

સમીકરણ કે જે નિયતસંબંધમાં બંધબેસે છે કે જે અરેખીય કે વક્રીય હોઈ શકે છે તે સંખ્યાત્મક સ્વરૂપમાં હોઈ શકે. સામાન્ય સ્વરૂપ કે જે બે ચલોને સમાવે છે. તે ચતુવર્ગીય હોઈ શકે. સમીકરણ નીચે મુજબ છે.

$Y = a + bx + Cx^2$

અહીં ત્રણ માપદંડો છે, કે જે a,b અને c અને સામાન્ય સમીકરણ.

$\Sigma\ Y = na + b \times + C\Sigma x^2$

$\Sigma xY = a\Sigma\ x + b\Sigma x^2 + C\Sigma x^3$

$\Sigma x^2Y = a\Sigma x^2 + b\Sigma x^3 + C\Sigma x^4$

નોંધો કે સામાન્ય સમીકરણો એ નિયતસંબંધ સમીકરણ છે કે જે a,b અને c ના ગુણાંકો વડે ગુણવામાં આવ્યા છે. અને દરેક અવલોકનો ઉપર સરવાળો કરવામાં આવ્યો છે.

ચોક્કસ અરેખીય સમીકરણોને રેખીય સમીકરણોમાં લઘુગણકની મદદથી પરિવર્તિત કરી શકાય છે. પરિવર્તિત રેખીય સમીકરણો પરથી માપદંડોની ચોક્કસ કિંમત એ અગાઉના વિભાગમાં ચર્ચેલી રીત જેટલી જ હોય છે.

અહીં નીચે વારંવાર ઉપયોગમાં લેવાતા અરેખીય સમીકરણો અને અનુક્રમે પરિવર્તિત રેખીય સમીકરણો આપેલ છે.

(1) $Yae^{b\times}$

લઘુગણક લેતા તેને નીચે મુજબ લખી શકાય ;

in $Y = \ln a + bx$

અથવા $Y' = a + \beta x^1$

કે જ્યાં $Y = \lambda na, a\text{-}\lambda na, x' = \beta = b.$

(2) $Y=ax^b$

લઘુગણક લેતા સમીકરણ નીચે મુજબ પરિવર્તિત કરી શકાય.

$\log y = \log a + b \log \times$

અથવા $Y' = a + \beta \times'$

કે જ્યાં $Y' = \log Y, \alpha = \log a, \beta = b. \quad X' = \log \times$

(3) $Y = \dfrac{1}{a + b\times}$

જો આપણે $Y^1 = {}^1\!/_Y$ લઈએ તો

$Y^1 = a + b\times$

(4) $Y = a + b\sqrt{x}$

જો આપણે X' લઈએ તો

$Y^1 = a + b\times'$

એકવાર અરેખીય સમીકરણ પરિવર્તિત થઈ જાય તો નિયતસંબંધ રેખા એ આ વિભાગમાં અગાઉ ચર્ચા કર્યા મુજબ રીત સાથે બંધબેસતો છે. આપણે સામાન્ય સમીકરણ મેવવીશું અને અવલોકેલી માહિતી પરથી ગણેલી માહિતીને તેમાં મૂકીશું. પરિવર્તિત માપદંડોમાંથી, સાચો માપદંડ પરિવર્તિત પરિવર્તન દ્વારા મેળવી શકાય છે.

8.5.4 નિયતસંબંધ નિર્દેશ (Prediction)

નિયત સંબંધ રેખાના અભ્યાસમાં ઊંડો રસ એ તેની ભવિષ્ય કરવાની ક્ષમતાને લીધે છે. શરૂઆતના ઉદાહરણોમાં આપણે ધાર્યું કે વાચકોને ઈસ્યુ કરેલા પુસ્તકો મુલાકાતીઓની સંખ્યા ઉપર આધારિત છે. આપણે અવલોકેલી માહિતીમાં રેખીય સમીકરણ બંધ બેસાડી શકીએ છીએ અને સંબંધ મેળવી શકીએ છીએ.

$Y = 3.5 + 0.65x$

આ સમીકરણ દ્વારા આપણે મુલાતાકીઓની સંખ્યા પરથી વાચકોને ઈસ્યુ થતા પુસ્તકોનું ભવિષ્ય કથન કરી શકીએ.

$Y = 3.5 + 0.65 \times 30 = 2$

આ રીતમાં નિયતસંબંધ સમીકરણમાં X ની કિંમત મૂકવાની અને અપેક્ષિત Y નું મૂલ્ય મેળવવાનું છે.

અહી સવાલ ઊભો થાય છે કે, આપણે ધારેલી કિંમત સાચી મળશે ? તેનો આધાર નિશ્ચાયકના ગુણાંક પર રહેલો છે. જો નિશ્ચાયકનો ગુણાંક એકની નજીક હોય, તો ત્યાં મોટા પ્રમાણમાં કહી શકાય કે ધારણા સાચી ઠરશે. ગમે તેમ, ધારેલી કિંમતને ઉદ્દેશ કરેલા નંબરોથી મેળવી શકાય છે કે જે માણસની વર્તણૂંક સાથે અને અજાણ્યાં તત્ત્વો સાથે સંકળાયેલા છે.

◆ **તમારી પ્રગતિ ચકાસો**

4) નીચે X અને Y ની માહિતી આપેલી છે.

X : 15	17	20	22	25	33
Y : 25	22	30	31	37	35

i) X ની ઉપર Y ની નિયતસંબંધ રેખા શોધો

ii) Y ની ઉપર X ની નિયતસંબંધ રેખા શોધો

iii) નિશ્ચાયકનો ગુણાંક શોધો

iv) પિયર્સનનો પરિબળ ગુણાકાર સહસંબંધનો ગુણાંક મેળવો.

નોંધ : (1) નીચેની જગ્યામં જવાબ લખો

(2) તમારા જવાબોને એકમના અંતે અપાયેલા જવાબ સાથે સરખાવો

....

....

....

....

....

....

8.6 સમય શ્રેણીનું વિશ્લેષણ (TIME SERIES ANALYSIS) :

નિયત સંબંધ વિશ્લેષણમાં આપણે આધારિત અને સ્વતંત્ર ચલો વચ્ચેના સંબંધમાં કારણો અને અસરોનો અભ્યાસ કર્યો. સ્વતંત્ર ચલ એ કોઈપણ લાક્ષણિકતા હોઈ શકે. જ્યારે સ્વતંત્ર ચલ 'સમય' છે, ત્યારે આપણે તેને 'સમય શ્રેણી' કહીશું. આપણે ધારીશું કે આધારિત ચલ એ સમય અથવા સમયના જુદા જુદા ફેરફાર પર આધારિત છે.

આપણી રોજિંદુ જિંદગીમાં આપણને કેટલીક ઘટનાનો ભેટો થઈ જાય છે. ગ્રંથાલયમાં જથ્થો (Stock), વપરાશકર્તા પાછળ સુવિધા માટે કરાતો ખર્ચ વગેરે ઉદાહરણો છે.

ગ્રંથપાલ અને માહિતી અધિકારી ભવિષ્ય માટે ખ્યાલ બોધે તે પહેલા એક મહત્વનું કાર્ય ઉદાહરણ તરીકે, એક પ્રકાશક તેના પ્રકાશનનું આવતા વર્ષમાં વેચાણની સંભાવના જાણવા માંગે છે., તેથી તે તેના દ્વારા પ્રકાશિત કરાયેલા પુસ્તકો વેચાણમાં વિસ્તરણની ઉણપ કે વધેલા માલની સંભાવના રાખવા માટે યોગ્ય પગલા ભરી શકે, ગ્રંથપાલ પુસ્તકના વલણનો અભ્યાસ કરવા માંગે છે કે જેથી ભવિષ્યમાં પ્લાન બનાવવા માંગે તો યોગ્ય માપદંડ લઈ શકે. આ બધા હેતુ માટે, દરેક જણ માહિતી મેળવી શકે કે જેને સમયના યોગ્યગાળા સાથે ભેગી કરવામાં અને રેકોર્ડ કરવામાં આવી હોય. આવી આંકડાશાસ્ત્રીય (મહિતી)ને 'સમય શ્રેણી' ની માહિતી કહી શકાય.

સમય શ્રેણીની માહિતીનું ઉદાહરણ ગ્રંથાલયમાં ગ્રંથલાય અને માહિતી વિજ્ઞાનમાં 1995 થી 2004 સુધીમાં ઈસ્યુ થયેલા પુસ્તકોની સંખ્યા. આ માહિતીને નીચે મુજબ સંગ્રહિત કરી શકાય.

કોષ્ટક 8.3 સમય સાથે ઈસ્યુ થયેલ પુસ્તકોની અનુમાનિત માહિતી સંગ્રહ (નોંધ)

Hypothesis Data Record

વર્ષ	ઈસ્યુ થયેલા પુસ્તકોની સંખ્યા
1995	356
1996	350
1997	391
1998	289
1999	408
2000	412
2001	405
2002	482
2003	497
2004	469

બારીકાઈથી નિરીક્ષણ કરશો તો જોવા મળશે કે પુસ્તકની માંગ કેટલીક ચડઉતર સાથે વધતી રહી છે. ત્યાં કેટલાંક કારણો છે કે જેથી ચોક્કસ સમયગાળા દરમિયાન વધારો કે ઘટાડો થાય છે.

8.6.1 સમય શ્રેણીના ઘટકો (Components of Time Series)

સમયશ્રેણીમાં ચડઉતરને કેટલાક પ્રકારના ફેરફારમાં વર્ગીકૃત કરી શકાય છે

(a) પ્રચલિત વલણ (Secular trend)

(b) મોસમી ફેરફાર (Seasonal Variation)

(c) ચક્રીય ફેરફાર (Cyclical Variation)

(d) અનિયમિત ફેરફાર (Irregular Variation)

સમયશ્રેણીને કદાચ ઉપરના ઘટકોમાં સમાવવામાં આવે છે.

(a) પ્રચલિત વલણ (Secular trend)

ફેરફાર કે જે માહિતીના સામાન્ય વલણના પરિણામ તરીકે જગ્યા લે છે. જેથી વધારો કે ઘટાડો પ્રચલિત ફેરફાર તરીકે ઓળખાય છે. લાંબા સમયગાળા ઉપરની સામાન્ય ફેરફાર ચાલુ રહે તેને પ્રચલિત વલણ કહેવાય. ઈસ્યુ થયેલ પુસ્તકોની સંખ્યા માટેનું ઉપરનું સમય શ્રેણીનું ઉદાહરણ એ પ્રચલિત વલણનું ઉદાહરણ છે.

(b) મોસમી ફેરફાર (Seasonal Variation)

બદલાવ કે ફેરફાર કે જે આબોહવા, હવામાન, મહત્વની ઘટના વગેરેના એક વર્ષના પરિણામ તરીકે વગેરેના એક વર્ષના પરિણામ તરીકે મોસમ પ્રમાણે ઉદ્ભવે તે કદાચ સંભવ છે કે ઉતાર ના ઉદાહરણમાં ઈસ્યુ થયેલા પુસ્તકોની માહિતી (જો શક્ય હોય તો જાતિના પ્રમાણે જોવું) વર્ષની શરૂઆતમાં ઓછી સંખ્યા સાથે શરૂ થાય છે, તે મધ્યમાં ટોચ પર પહોંચે વર્ષને અંતે નીચે ઉતરે છે. આમ, ચઢાવ-ઊતાર જોવા મળે છે. આવા પ્રકારના એક વર્ષના ફેરફારને મોસમી ફેરફાર કહેવાય છે.

(c) ચક્રીય ફેરફાર (Cyclical Variation)

બદલાવ કે જે ચક્રીય ફેરફાર જેવા કે સમૃધ્ધી અને તણાવને ચક્રીય ફેરફાર કહી શકાય. દરેક ધંધાકીય ચક્રમાં ચાર સમયગાળા હોય છે.

(i) સમૃદ્ધિ (ii) ઘટાડો (iii) તણાવ (iv) સુધારો. ચક્રીય ફેરફાર એક વર્ષના સમયગાળા કરતા વધુ હોય છે.

(d) અનિયમિત ફેરફાર

બદલાવ કે જે પરિબળને કારણે ઉદ્ભવે કે જે હિંસક ધમાલ, કુદરતી આફતો વગેરેની ધારણા કરી શકાતી નથી તેઓને અનિયમિત ફેરફાર ગણાય.

સમય શ્રેણીના ઘટકો કે જે મોસમી, વલણ, ચક્રીય અને અનિમયમિત વગેરેને અલગ કરી શકાય છે. પરંપરાગત સમય શ્રેણી વિશ્લેષણમાં એવું ધારવામાં આવે છે કે ત્યાં આ ચાર વિભાગો વચ્ચે બહુવિધ સંબંધ છે. તેને સાંકેતિક રીતે નીચે મુજબ દર્શાવી શકાય.

$(Y = T \times S \times C \times I)$

કે જ્યાં T = પ્રચલિત ફેરફાર

S = મોસમી ફેરફાર

C = ચક્રીય ફેરફાર

I = અનિયમિત ફેરફાર

આને બહુવિધ મોડેલ કહેવાય. બીજા અભિપ્રાયથી એવું ધારવામાં આવે છે કે

$Y = T + S + C + I$

આ પ્રકારની કક્ષાના સૂત્રીકરણને 'ઉમેરેલો નમૂનો' કહે છે.

8.6.2 પ્રચલિત (સેક્યુલર) વલણના માપદંડો (Measurement of Secular Trends)

પ્રચલિત વલણ નક્કી કરવા માટે ઘણી રીતો છે વારંવાર ઉપયોગમાં લેવાતી રીતો

(1) ચલિત સરેરાશ પધ્ધતિ

(2) લઘુવર્ગની રીત

(a) ચલિત સરેરાશ પધ્ધતિ (Method of Moving Average)

સમય શ્રેણીના એકબીજા પર સંપાત (Coverlapping) થતા સમયગાળા વડે સરેરાશ શ્રેણી ગણતા ચઢાવ-ઊતારને સરખો કરવાની રીત છે. ચલિત સરેરાશ ગણતાં પહેલા એક વાત જરૂરી છે કે ચલિત સરેરાશનો યોગ્ય સમયગાળો નક્કી કરવો જોઈએ જેમ કે ત્રણ વર્ષીય, પંચવર્ષીય વગેરે

જો સમયગાળો m વર્ષ હોય તેમ નક્કી કરીએ તો ચલિત સરેરાશને શ્રેણીના સતત m કિંમતો, આપાત સમયગાળાને આવરી લેતાની મધ્ય કિંમતથી ગણવામાં આવે છે પ્રથમ m કિંમતોનો મધ્યક $\frac{1}{m}(Y_1 + Y_2 + ... + Y_m)$ વડે ગણવામાં આવે છે.

તે પ્રથમ વર્ષની પ્રથમ M વર્ષની સરેરાશ કિંમત બનશે. બીજી વર્ષની ચલિત સરેરાશ કિંમત મેળવવા માટે બીજા વર્ષથી (M+1) વર્ષનો સરેરાશ સમય ગણવામાં આવે છે. ઉદાહરણ તરીકે $\frac{1}{m}(Y_1 + Y_2 + ... + Y_m + 1)$ એ હવે પછીની ચલિત સરેરાશ હશે. અંતિમ અવલોકન આવરી ન લેવાય ત્યાં સુધી આ રીત ચાલુ રાખવાની.

ત્રણ વર્ષની ચલિત સરેરાશ નીચે મુજબ મળે ;

$$\frac{1}{3}(Y_1 + Y_2 + Y_3), \frac{1}{3}(Y_2 + Y_3 + Y_4), \frac{1}{3}(Y_3 + Ya + Y_s)$$

અને પાંચ વર્ષની ચલિત સરેરાશ નીચે મુજબ છે.

$$\frac{1}{5}(Y_1 + Y_2 + Y_3 + Y_s), \frac{1}{5}(Y_2 + Y_3 + Y_4 Y_s + Y_6)$$

નીચેના ઉદાહરણ દ્વારા ઉપરની રીત સમજવી સહેલી છે.

ઉદાહરણ :

પ્રકાશકના વેચાણ આંકડાની માહિતી ઉપરથી 3 વર્ષની ચલિત સરેરાશ શોધો.

કોષ્ઠક 8.4 ચલિત સરેરાશની ગણતરી કરવા માટે અનુમાનિત માહિતી

વર્ષ	વેચાણ (સો એકમ)
1990	05
1991	07
1992	09
1993	12
1994	11
1995	10
1996	08
1997	12
1998	13
1999	17
2000	19
2001	14
2002	13
2003	12
2004	15

પંદર વર્ષ ના પુસ્તકોના વેચાણની માહિતી આપેલી છે. ત્રણ વર્ષની ચલિત સરેરાશ ગણવા માટે આપણે પ્રથમ ત્રણ વર્ષનું કુલ ટોટલ લઈશું. તે 5₹7₹9x=21 અને આ કિંમતને મધ્ય અવલોકનની સામે મૂકો. એટલે કે 1991 ની સામે કોષ્ટક 8.5 ના સ્તંભ 4 માં ચલિત સરેરાશ એટલે કે 3 વર્ષના કુલ ટોટલને સમયગાળા વડે ભાગીને કોષ્ટકમાં આપેલ છે. તેથી $21/3 = 7$ માટે પ્રથમ પ્રવેશ કરશે. નીચેના કોષ્ટકમાં એજ વસ્તુને ગણતરી કરીને પરિણામ દર્શાવ્યું.

કોષ્ટક 8.5 : 3 વર્ષની ચલિત સરેરાશની ગણતરી

વર્ષ (1)	વેચાણ (2)	3 વર્ષીય કુલ (3)	3 વર્ષની ચલિત સરેરાશ (4)
1990	5	-	-
1991	07	21	7.0
1992	09	28	9.33
1993	12	32	10.67
1994	11	33	11.00
1995	10	29	09.67
1996	08	30	10.00
1997	12	33	11.00
1998	13	42	14.00
1999	17	49	16.33
2000	19	50	16.67
2001	14	46	15.33
2002	13	39	13.00
2003	12	40	13.33
2004	15	--	---

ઉપરના ઉદાહરણથી નોંધો કે આપણે શરૂઆતના વર્ષની અને અંતિમ વર્ષની ચલિત સરેરાશ શોધી નથી. પાંચ વર્ષની ચલિત સરેરાશના કિસ્સામાં આપણે શરૂઆતના બે વર્ષ અને અંતિમ બે વર્ષની ચલિત સરેરાશ ગુમાવીશું. આ ગુમાવેલી માહિતી ચલિત સરેરાશના વધવાના સમયગાળા સાથે વધતી જશે. ચલિત સરેરાશથી ભવિષ્ય માટે ચોક્કસ આંકડા ધારી શકાતા નથી.. તે ફક્ત ભૂતકાળનું વિશ્લેષણ છે.

(બી) લઘુવર્ગની રીત :

આપણે અગાઉ આ રીત નિયતસંબંધ રેખાઓ મેળવવા વાપરી હતી. આ રીત નિયતસંબંધ રેખાની અનુકુળતાને અનુરૂપ છે. અહીંયા સ્વતંત્ર ચલ 'સમય' t છે. પ્રથમ પગલાંમાં 'પ્રચલિત વલણ'ના સ્વરૂપને મેળવાનું છે. તમે જાણો છો કે, બંને સીધી રેખાઓ લઘુ વર્ગની રીતથી બંધ બેસે છે. જો Yએ આધારિત ચલ હોય તો સીધી રેખા Y = a+b બંધ બેસે છે. ફરીથી આપણે અવલોકીત અને ધારેલી કિંમત વચ્ચેની ભૂલ ઓછી કરવાની છે. લઘુ વર્ગની રીત આપણને ભૂલનાં પદોના વર્ગનો સરવાળો ઓછો હોવો હોઈએ. આ પધ્ધતિથી, આપણે આધારિત અને સુરેખ ચલ વચ્ચેના સંબંધને સામાન્ય સમીકરણથી અનુમાનીત કરી શકીએ છીએ. રેખીય સમીકરણ Y=a+bt માટે સામાન્ય સમીકરણો નીચે મુજબ થશે.

$$\sum Y = na + b\sum t$$

$$\sum +Y = a\sum t + b\sum t^2$$

અચળાંક 'a' અને 'b' આ બે સમીકરણોથી મેળવી શકાય છે અને 'n' એ ઉદાહરણમાં રહેલા અવલોકનોની સંખ્યા દર્શાવે છે.

આપણે સમયના ઉદ્ભવ તરીકે મધ્યબિંદુ લઈને ચલ 't' શોધીશું ધારો કે 'N=5' વર્ષ પછી ઉદ્ભવ તરીકે સમયનું ત્રીજુ વર્ષ લઈએ તો આપણે 't= -2,-1,0,1,2 મેળવીશું. અહીં જોઈ શકાય છે કે ત્રીજા વર્ષ માટે 't=0'

આવરી લેવાયેલા વર્ષોની સંખ્યા બેકી એટલે કે 6 હોય તો, ઉદ્ભવબિંદુ વચ્ચેના બે વર્ષનું મધ્યબિંદુ હશે. એટલે કે ત્રણ વર્ષ પછીના 6 મહિના ગણવામાં આવશે, આવા કિસ્સામાં 'T' ની કિંમતો -5,-3,-1,₹1,₹3,₹5.

ઉપરના 8.5 ના કોષ્ઠકના ડેટા લઈ સુરેખ પ્રચલિત રેખામાં દર્શાવો.

કોષ્ટક 8.6 માં ગણેલી જરૂરી માહિતી આપેલી છે.

કોષ્ટક 8.6 સમયના વલણની માહિતી ગણતરી

વર્ષ	't'	વેચાણ	tY	t2
1990	-7	05	-35	49
1991	-6	07	-42	36
1992	-5	9	-45	25
1993	-4	12	-48	16
1994	-3	11	-33	09
1995	-2	10	-20	04
1996	-1	08	-8	01
1997	0	12	0	0
1998	01	13	13	1
1999	02	17	34	04
2000	03	19	57	09
2001	04	14	56	16
2002	05	13	65	25
2003	06	12	72	36
2004	07	15	105	49
કુલ	--	**00**	**177**	**280**

સામાન્ય સમીકરણો ફરીથી યાદ કરીએ જે નીચે મુજબ છે.

$$\sum Y = na + b\sum t$$

$$\sum +Y = a\sum t + b\sum t^2$$

કોષ્ઠકોમાંથી અનુક્રમે કિંમતો મૂકતાં

177 = 15 × a + b × 0

177 = a × 0 + b × 280

અથવા

15 a = 177 or a = 11.8

280 b= 171 or b = 0.61

તેથી પ્રચલિત રેખા Y=11.8 + 0.61t

યાદ રાખો કે 't' એ 197 સાથે ઉદ્ભવ તરીકે પધ્ધતિસર ઘડેલું છે.

લઘુવર્ગની રીત આપણને 'y' માટેની કિંમતો ધારવા માટે સક્ષમ બનાવે છે આવું 't' ની કિંમત સમીકરણમાં મૂકતા શક્ય બને છે.

ઉપરના ઉદાહરણોમાં, 2006માં વેચાયેલા પુસ્તકો 17.29 હશે.

કોષ્ઠકોમાંથી અનુક્રમે કિંમતો મૂકતાં

$Y= 11.8 + 0.61\times9=17.29$

જ્યારથી 1997 થી શરૂઆત કરી છે t ની જગ્યાએ 9 મૂક્યો છે. એટલે કે ઉદ્ભવનું વર્ષ 2006 એ 9 વર્ષ થશે. 2006 માં પુસ્તકોના વેચાણની ધારણા (સોના એકમમાંથી) 17.29 થશે.

તમે જાણો છો તમે, અરેખીય વલણ પણ અવલોકેલી માહિતી સાથે બંધ બેસે છે. તેથી તેમના પર પણ ધારણા કરવી જોઈએ. વલણ રેખાના વિશ્લેષણમાં, આપણે ગૂઢ ધારણા કરી હતી કે પૂર્વ વર્તણૂક ભવિષ્યના સમયમાં પણ ચાલુ રહે છે તેથી પૂર્વ વર્તણૂકમાં બદલાવ કદાચ ધારણાને ભરોસાપાત્ર રહેતી નથી.

8.6.3 અન્ય વિભાગોનાં માપદંડ (Measurement of Other Components)

મોસમી ફેરફાર અને ચક્રીય ચઢાવઉતાર એ માહિતીમાં સમયગાળા અને પુનવર્તિત ચળવળ છે. ઉપર દર્શાવવામાં આપ્યું છે કે મોસમી ફેરફાર એ પ્રકૃતિમાં ટૂંકા ગાળા અને સામાન્ય રીતે સમયગાળો એક વર્ષ કરતા ઓછો હોય છે. આ તફાવતમાં ચક્રીય ફેરફાર વર્ષ કરતાં વધુ લાંબો હોય છે.

અહીં મોસમી ફેરફારોને માપવાની ઘણી બધી રીતો છે, જેવી કે તે વલણ પધ્ધતિ માટે ગુણોત્તર, ચલિત સરેરાશની રીતનો ગુણોતર, જોડાણ સંબંધી ગુણોત્તર ચક્રીય ચઢાવઉતાર ને સંવાદિત વિશ્લેષણની માપી શકાય છે. વર્ણપટ વિશ્લેષણ વગેરે. આવી પધ્ધતિઓમાં કંટાળાજનક પધ્ધતિઓ રહેલી છે. તેથી અહીં ચર્ચા કરી નથી. જો તમે આવી પધ્ધતિમાં રસ ધરાવતા હોય તો તમે આ એકમને અંતે આપેલ પુસ્તકોના નામમાંથી યોગ્ય પુસ્તક વાંચી શકો છો. ગમે તેમ સમય શ્રેણીને તેના ચાર વિભાગો મોસમી, વલણ, ચક્રીય અને વ્યવહારમાં વિભાજીત કરવા કેટલીક પધ્ધતિઓ અનુસરવી પડે છે.

નીચે આ પધ્ધતિઓ ટૂંકમાં સમાવેલ છે.

(એ) ઉપર દર્શાવ્યા મુજબ મોસમી વિભાગોને ચલિત સરેરાશની મદદથી અંદાજિત કરી શખાય છે. વિભાગોને બાદબાકી દ્વારા મુખ્ય અવલોકનોમાંથી રદ પણ કરી શકાય છે. જો તમે ઉમેરાતો વધતો નમૂનો ધારેલ હોય તો.

(બી) મોસમી ગોઠવેલી માહિતનું વલણ લઘુવર્ગની સીધી રેખા અથવા બીજા વિધેયોના સંબંધ દ્વારા લેટેસ્ચ સ્વેર વડે ગણી શકાય ઉપર દર્શાવ્યા મુજબ લઘુવર્ગની રીત માહિતી (સીઝનલ) ગોઠવી શકાય છે.

(સી) બાકીનું કે જે મુખ્ય સમય શ્રેણીમાંથી મોસમી અને વલણ વિભાગોને રદ કરતા વધે છે તેને સંગ્રહિત કરી શકાય છે અને આલેખ સ્વરૂપે દર્શાવી શકાય છે. માહિતી શ્રેણીનો વધેલો ફેરફાર એ ચક્રીય અને વ્યવહારની લાક્ષણિકતા છે.

8.7 સારાંશ (SUMMARY)

આ એકમમાં આપણે માહિતીની વિશ્લેષણ માટે જુદી જુદી આંકડાશાસ્ત્રીયે પધ્ધતિઓની ચર્ચા કરો. તેને વારંવાર માહિતીના ગણ અથવા શ્રેણી માટે સારાંશના આંકડા પૂરા પાડવા પડે છે. આવા આંકડાઓ કેન્દ્રિય વલણના માપ છે જેવા કે તફાવત, પ્રમાણભૂત વિચલન અને તફાવતની સહગુણાંક.

અહીંયા કિસ્સાઓ કે જ્યાં ઉદાહરણીય એકમની એક કરતાં વધુ લાક્ષણિકતા છે. આવા પ્રકારની માહિતીમાં આપણે સહસંબંધ ગુણાંકને શોધી શકાય અથવા આપણે નિયતસંબંધ સમીકરણને બંધ

બેસાડી શકાય છે. યાદ રાખો કે સહસંબંધ એ કારણો અને અસરો વચ્ચેનો સંબંધ દર્શાવતો નથી તે ફક્ત સંબંધોની મજબૂતાઈ દર્શાવે છે. નિયતસંબંધ વિશ્લેષણમાં ચલને બે ભાગમાં વહેંચવામાં આવે છે. સ્વતંત્ર અને આધારિત નિયંતસંબંધ સમીકરણ સીધી રેખા અથવા વક્ર હોઈ શકે કે જે બંધબેસતા સમીકરણના પ્રકાર ઉપર આધાર રાખે છે.

વારંવાર આપણી પાસે સમયના ચોક્કસ ગાળા માટે ચોક્કસ ગાળાની માહિતી હોય છે. આવી માહિતીને સમયશ્રેણી અને ચોક્કસ વિભાગો સમાયેલા છે. જેવા કે પ્રચલિત વલણ, ચક્રીય ફેરફાર, મોસમી ફેરફાર અને અનિયમિત તફાવત.

8.8 તમારી પ્રગતિ ચકાસોના ઉત્તરો (ANSWERS TO SELF CHECK EXERCISES)

1) પ્રમાણભૂત વિચલન 21.07 છે. પુસ્તકમાં આપેલ સૂત્રનો ઉપયોગ કરો અને જવાબ ચકાસો

2) સહસંબંધનો ગુણાંક r = +0.61.

3) (a) તેને ફેરફારના ધન વર્ગમૂળ તરીકે દર્શાવાય છે. અને σ વડે દર્શાવાય છે. $\sigma = \sqrt{+\sigma^2}$

(b) ફેરફાર એ વિસ્તરણના માપ માટે બહોળા પ્રમાણમાં ઉપયોગ થાય છે તેને σ^2 વડે દર્શાવાય છે અને નીચે મુજબ વ્યાખ્યાયિત કરવામાં આવે છે.

$$\text{ફેરફાર} = \sigma^2 = \frac{1}{N}\Sigma fi(X_i - X^-)^2 \frac{1}{N}\Sigma(X_i - X^-)^2$$

આવૃતિ વિસ્તરણના કિસ્સામાં ફેરફાર નીચે મુજબ દર્શાવાય છે.

$$\sigma^2 = \frac{1}{N}\Sigma fi(X_i - X^-)^2$$

જ્યાં $N = \sum nfi$, અવલોકનોની કુલ સંખ્યા

ગણતરીનું સાદુરૂપ આપવા માટે નીચેના સૂત્રનો ઉપયોગ કરીશું.

$$\sigma^2 = \frac{1}{N}\Sigma fixi^2 - X^{-2}$$

(c) $$r = \frac{\frac{\sum n}{i=i}(X_1 - X^-)(Y_i - Y^-)}{\frac{1}{N}\sigma \times \sigma Y}$$

(d) $$r_s = 1 - \frac{\frac{6}{1}\frac{\sum n}{=1}D_1^2}{n(n^2-1)}$$

(4) (1) તમારે સામાન્ય સમીકરણ શોધવું પડે અને સમીકરણમાં કિંમતો મૂકવી પડે.

અંદાજિત નિયતસંબધ $Y = 13.94 + 0.73\times$ રેખા હશે.

(2) X માટે Y ઉપર નિયતસંબંધ રેખા X = a + By હશે.

તેવી જ રીતે સામાન્ય સમીકરણો

$\Sigma X = n\alpha + \Sigma\beta Y$

$\Sigma XY = \alpha\Sigma Y + \beta\Sigma Y^2$

અંદાજિત નિયતસંબંધ રેખા X=-5.6 +92Y હશે.

(3) નિશ્ચાયકનો ગુણાંક એ નિયતસંબંધ $0.73 \times 0.92 = 0.067$ અને બંને નિયતસંબંધ રેખાનો ગુણાકાર છે.

જેથી $0.73 \times 0.92 = 0.067$

(4) સહસંબંધ ગુણાંક એ નિશ્ચાવાચકના ગુણાંકનું વર્ગમૂળ છે એટલે કે (r = 0.067 = 0.82) નિયત સંબંધગુણાંક ધન હોવાથી સબસંબંધ ગુણાંક પણ ધન હશે.

8.9 ચાવીરૂપ શબ્દો (KEY WORDS)

અંકગણિતીય મધ્યક (Arthmetic Mean) : અવલોકનોની કિંમતનો કુલ સરવાળાને અવલોકનોની કુલ સંખ્યા વડે ભાગતાં અંકગણિતીય મધ્યક મળે છે.

મધ્યસ્થ (Median) : અવલોકનોના ગણમાં એકદમ વચ્ચે રહેલી કિંમત કે તેમને જ્યારે વિસ્તારના પ્રમાણમાં ગોઠવવામાં આવ્યા હોય.

બહુલક (Mode) : અવલોકનોના ગણમાં જ સંખ્યા સૌથી વધુ વખત આવતી હોય તેને બહુલક કહેવાય.

ફેરફારનો સહગુણક (Co-efficient of vgriation) : આ વિસ્તરણ સંબંધી માપ છે કે જે માપદંડના એકમ તરીકે સ્વતંત્ર છે. વિરોધ તરીકે જોઈએ તો પ્રમાણભૂત વિચલન એ વાસ્તવિક વ્યાખ્યા છે.

વિસ્તાર (Range) : આપેલ અવલોકનોના ગણમાં મોટામાં મોટા અને નાનામાં નાના અવલોકન વચ્ચેના તફાવતને વિસ્તાર કહેવાય.

પ્રમાણભૂત વિચલન (Standard Deviation) : પ્રમાણિતભૂત વિચલન એ ફેરફારનું ઘન વર્ગમૂળ છે.

તફાવત/ફેરફાર (Variance) : વિચલિત અવલોકનોનો વર્ગનો અંકગણીતીય મધ્યક છે.

સામાન્ય સમીકરણો (Normal Equations) : સામાન્ય સમીકરણો ક્રમિક સમીકરણોનો ગણ કે જે લઘુવર્ગની રીતનો સારાંશ છે. ઉદાહરણ તરીકે પ્રત્યાગમ વિશ્લેષણ તેનો ઉપયોગ નમૂનાનાં માપદંડનો અંદાજ કાઢવા માટે થાય છે.

નિયતસંબંધ (Regression) : તે બે કે બે થી વધુ ચલો વચ્ચેનું આંકડાશાસ્ત્રીય માપદંડનો માહિતીના ઉદ્ભવના પદમાં સરેરાશ સંબંધ છે.

ચક્રીય ફેરફાર (Cyclical Yariations) : સમય શ્રેણીની આંદોલિત ફેરફારો કે જ્યાં આંદોલનનો સમયગાળો એક કરતાં વધુ હોય તેને ચક્ર કહેવાય.

અનિયમિત ફેરફાર **(Irregular Movement)** : સમય શ્રેણીની અનિયમિત ગતિ કે જે બીજા વિભાગો દ્વારા સમજાવી શકાતી નથી. આ અર્થમાં એ બીજા વિભાગોનો બાકી ભાગ છે.

લઘુવર્ગની રીત (Method of Least Squares) : જ્યારે બહુપદી વિધેય સમય શ્રેણીમાં બંધબેસતી હોય ત્યારે લઘુવર્ગની રીતને વિધેયના માપદંડોની એવી રીતે પસંદ કરવા જોઈએ કે જેથી ખરેખર અવલોકનો અને ધારેલી કિંમતના વિચલનના વર્ગનો સરવાળો પસંદ કરવો જોઈએ.

મોસમી ફેરફાર (Seasonal Variation) : આવર્તીય ફેરફાર કે જ્યાં આવર્તકાળ એક વર્ષ કરતાં વધુ ન હોય.

પ્રચલિત વલણ (Secular Trend) : સમય શ્રેણીનો સુરેખ નિયમિત અને લાંબી મુદતનો ફેરફાર. સમય શ્રેણીનો વલણનો ઝોક ઉપરની તરફ હોઈ શકે. નીચે તરફ અથવા ઢળવું અથવા વધુ રહેવું અથવા સમય સાથે સતત નીચે જવું.

8.10 સંદર્ભ અને વિશેષ વાંચન (REFERENCE AND FURTHER READING)

સેન્ડર્સ ડી. એચ. (1980) : આંકડાશાસ્ત્ર : નવું વલણ. નવી દિલ્હી, મેકગ્રો હિલ.

રાવ, I.K.R Quantitiative Methods for library and information science, New Delhi : wiley eastern