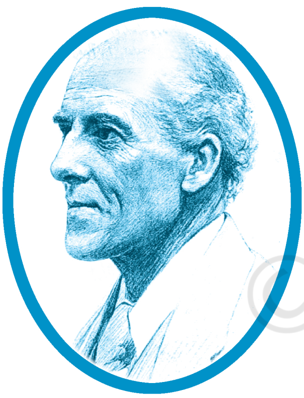
باب 15

شماریات (STATISTICS)

❖"شماریات کو صحیح طور پر اوسط اور تخمینہ کی سائنس کہا جاتا ہے۔"

❖ *A.L. BOWLEY & A. L. BODDINGTON*

15.1 تعارف (Introduction)



کارل پیرسن
(1857-1936)

ہم جانتے ہیں کہ شماریات کا استعمال کسی خاص مقصد کے لیے آنکڑوں کو اکٹھا کرنے میں کیا جاتا ہے۔ آنکڑوں کے تجزیہ اور ترجمہ کر کے ہم ان کے بارے میں فیصلے لے سکتے ہیں ۔ پچھلی جماعتوں میں آنکڑوں کو گراف کے ذریعہ ظاہر کرنا اور جدولی شکل میں لکھنے کے بارے میں پڑھ چکے ہیں۔ آنکڑوں کی یہ نمائندگی کچھ خاص خدوخال کا انکشاف کرتی ہے یا آنکڑوں کی خصوصیات کا۔ ہم پہلے ہی آنکڑوں کی نمائندگی کی قدر کو معلوم کرنے کے طریقوں کے بارے میں مطالعہ کر چکے ہیں۔ یہ قدر مرکزی ملان کو ناپنے کا طریقہ کہلاتی ہے۔ درمیانہ کو یاد کیجئے (ریاضیاتی درمیانہ)، وسطانیہ (median) اور بہتاتیہ (Mode) مرکزی ملان کو ناپنے کے تین طریقے ہیں۔ مرکزی ملان کو ناپنے سے ہمیں ناممکن صورت کا اندازہ ہو جاتا ہے جہاں آنکڑوں کے نقاط مرکز پر رہتے ہیں۔ لیکن آنکڑوں سے بہتر ترجمانی کرنے کے لیے ہمارے پاس ایک ایسا تصور ہونا چاہیے کہ کس طرح آنکڑے پھیلے ہوئے ہیں یا وہ ناپ کے مرکزی ملان کے اردگرد کس طرح گچھوں میں موجود ہیں۔

اب دو بلے بازوں کے ذریعے ان کے آخری دس میچوں میں بنائے گئے رنوں پر غور کیجئے جو اس طرح ہیں۔

بلے باز A:	30, 91, 0, 64, 42, 80, 30, 5, 117, 71
بلے باز B:	53, 46, 48, 50, 53, 53, 58, 60, 57, 52

	بلّے باز A	بلے باز B
درمیانہ	53	53
وسطانیہ	53	53

اس بات کو یاد کیجئے کہ ہم آنکڑوں کے درمیانہ (جسے $\bar{x}$ سے ظاہر کیا جاتا ہے) مشاہدوں کے مجموعہ کو مشاہدوں کے تعداد سے تقسیم کر کے حاصل کرتے ہیں یعنی؛

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

ساتھ ہی، ہم وسطانیہ کو معلوم کرنے کے لیے پہلے آنکڑوں کو بڑھتی ہوئی ترتیب یا گھٹتی ہوئی ترتیب میں رکھتے ہیں اور ذیل اصول کا استعمال کرتے ہیں۔

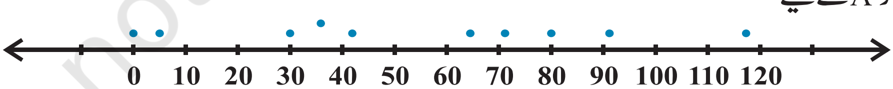
اگر شاہدوں کی تعداد طاق ہے، تب وسطانیہ $\left(\frac{n+1}{2}\right)^{th}$ مشاہدہ ہے۔

اگر مشاہدوں کی تعداد جفت ہے، تب وسطانیہ $\left(\frac{n}{2}\right)^{th}$ اور $\left(\frac{n}{2}+1\right)^{th}$ مشاہدوں کا درمیانہ ہے۔

ہم نے دیکھا کہ دونوں بلے باز A اور B کے ذریعے بنائے گئے رنوں کا درمیانہ اور وسطانیہ یکساں ہے یعنی 53۔ کیا ہم یہ کہہ سکتے ہیں کہ دونوں کھلاڑیوں کی کارکردگی یکساں ہے؟ صاف طور پر نہیں، کیونکہ بلے باز 'A' کے اسکور کا اتار چڑھاؤ 0 سے (کم سے کم) 117 (زیادہ سے زیادہ) تک ہے۔ جب کہ، بلے باز B کے بنائے گئے رنوں کے اسکور کی وسعت (range) 46 سے 60 تک ہے۔

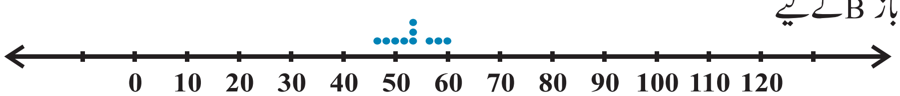
اب ہم اوپر دیئے ہوئے اسکور کو ایک عددی خط پر نقطوں کے ذریعے خاکہ تیار کرتے ہیں۔

بلے باز A کے لیے



شکل 15.1

بلے باز B کے لیے



شکل 15.2

ہم یہ دیکھ سکتے ہیں کہ بلے باز B کے مطابق نقاط ایک دوسرے کے قریب ہیں اور مرکزی ملان کی ناپ کے اردگرد جمع ہو رہے ہیں (درمیانہ اور وسطانیہ) جب کہ بلے باز A کے مطابق پھیلے ہوئے ہیں اور زیادہ پھیلے ہیں۔

اس طرح مرکز ملان کی ناپ دیئے ہوئے آنکڑوں کی مکمل جانکاری دینے کے لیے ناکافی ہے۔ تغیر (انتشار) ایک دوسرا عنصر ہے جس کے مطالعہ کی شماریات میں ضرورت ہے۔ ''مرکز ملان کو ناپنے کی طرح'' ہم ایک اکیلا عدد چاہتے ہیں جو انتشار کو بیان کرنے کے لیے ہو۔ اس اکیلے عدد کو ''انتشار کی ناپ'' کہا جاتا ہے۔ اس سبق میں ہم کچھ خاص انتشار کی ناپ کے بارے میں باجماعت اور بے جماعت آنکڑوں کے لیے پڑھیں گے۔

15.2 انتشار کی ناپ (Measures of Dispersion)

انتشار یا آنکڑوں میں پھیلا پن مشاہدوں کی بنیاد پر معلوم کیا جاتا ہے اور اس جگہ استعمال کے لیے مرکزِ ملان کی ناپ کے طریقے پر۔ یہاں مندرجہ ذیل انتشار کی ناپ دی گئی ہیں:

(i) وسعت (ii) چوتھائی انحراف (iii) درمیانہ انحراف (iv) معیاری انحراف

اس سبق میں ہم تمام انتشار ناپ کے بارے میں پڑھیں گے تین چوتھائی انحراف کو چھوڑ کر۔

15.3 وسعت (Range)

یاد کیجئے کہ اس مثال میں، جہاں دو بلے بازوں کے ذریعے بنائے گئے رن تھے، ہمیں ہر ایک سلسلے میں کم سے کم اور زیادہ سے زیادہ رنوں کی بنیاد پر کچھ تغیر کا تصور اسکورس میں ملا تھا۔ اس کے لیے ایک اکیلا عدد حاصل کرنے کے لیے، ہم زیادہ سے زیادہ اور کم سے کم قدروں کا فرق ہر ایک سلسلے کے لیے معلوم کرتے ہیں۔ یہ فرق آنکڑوں کا 'وسعت' کہلاتا ہے۔

بلے باز A کے کیس میں وسعت $= 117 - 0 = 117$ اور بلے باز B کے لیے، وسعت $= 60 - 46 = 14$ صاف طور پر B کی وسعت $<$ A کی وسعت۔ اس لیے کیس A میں اسکور پھیلے ہوئے یا انتشار میں ہیں جب کہ B میں نہ ایک دوسرے کے قریب ہیں۔

اس طرح، ایک سلسلے کی وسعت = کم سے کم قدر – زیادہ سے زیادہ قدر

آنکڑوں کی وسعت ہمیں تغیر یا پھیلاؤ کی نامکمل صورت دے دیتی ہے لیکن آنکڑوں کے انتشار کے بارے میں کچھ مرکز ملان کی رو سے نہیں بتاتی۔ اس مسئلہ کے لیے ہمیں کچھ اور متغیر کی ناپ کی ضرورت ہوگی۔ صاف طور پر، اس طرح کے پیمانے

مرکز ملان کی قدروں میں فرق یا انحراف پر منحصر ہیں۔

انتشار کا حاصل پیمانہ، جو مرکز ملان سے مشاہدوں کے انحراف پر مبنی ہے درمیانہ انحراف اور معیاری انحراف ہیں۔ ہمیں ان پر تفصیل سے بات چیت کرنی چاہیے۔

15.4 درمیانہ انحراف (Mean Deviation)

اسے دوبارہ یاد کیجئے کہ ایک اعداد وشمار x کا انحراف ایک مقرر قدر 'a' سے $x - a$ فرق ہے قدر x کا مرکز قدر 'a' سے انتشار معلوم کرنے کی ترتیب میں، ہم a پر جھکاؤ معلوم کرتے ہیں۔ انتشار کی ایک مطلق ناپ ان جھکاؤ کا درمیان ہے۔ درمیانہ معلوم کرنے کے لیے، ہمیں ان جھکاؤں کا حاصل جمع معلوم کرنا ہوگا۔ لیکن ہم یہ جانتے ہیں کہ مرکزِ ملان کا پیمانہ اعداد وشمار کے سیٹ کے لیے زیادہ سے زیادہ اور کم سے کم قدروں کے درمیان واقع ہے۔ اس لیے کچھ جھکاؤ منفی ہوں گے۔ اور کچھ مثبت۔ اس طرح جھکاؤ کا مجموعہ ختم ہوسکتا ہے۔ اور مزید یہ کہ انحراف کا مجموعہ درمیانہ $(\bar{x})$ سے صفر ہوتا ہے۔

ساتھ ہی

$$\text{درمیانہ کا جھکاؤ} = \frac{\text{انحراف کا جوڑ}}{\text{مشاہدوں کی تعداد}} = \frac{0}{n} = 0$$

اس طرح، درمیانہ سے درمیانہ جھکاؤ معلوم کرنا ہمارے لیے کسی کام کا نہیں ہے، جہاں تک انتشار کے ناپ کی بات ہے۔

یاد کیجئے کہ انتشار کا موزوں پیمانہ معلوم کرنے کے لیے، ہمیں مرکزی ملان یا ایک مقرر عدد 'a' کی ہر قدر کا فاصلہ معلوم کرنا ہے۔ دہرائیے کہ دو اعداد کی مطلق قدر کا فرق اعداد کے درمیان فاصلہ دیتا ہے جب انہیں عددی خط پر ظاہر کیا جاتا ہے۔ اس لیے ایک مقرر نقطہ 'a' سے انتشار کی ناپ معلوم کرنے کے لیے ہم مطلق قدروں کا درمیانہ مرکزی قدر کے انحراف سے لے سکتے ہیں۔ اس درمیانہ کو ''درمیانی انحراف'' کہتے ہیں۔ اس لیے ایک مرکزی قدر 'a' کے نزدیک درمیانہ انحراف مطلق قدروں کے مشاہدوں کا انحراف کا درمیانہ کہلاتا ہے۔ 'a' سے درمیانہ جھکاؤ **M. D.** سے ظاہر کیا جاتا ہے۔ (a) اس لیے

$$\text{M. D.}(a) = \frac{\text{(a) سے مطلق قدروں کے مجموعہ کا انحراف}}{\text{مشاہدوں کی تعداد}}$$

ریمارک درمیانہ انحراف کسی بھی مرکز ملان کے ناپ سے حاصل کیا جا سکتا ہے۔ حالانکہ درمیانہ انحراف، درمیانہ سے اور وسطانیہ سے عام طور پر شماریات کے مطالعہ میں استعمال کیا جاتا ہے۔

اب ہم یہ معلوم کرتے ہیں کہ کس طرح درمیانہ انحراف ، درمیانہ سے اور درمیانہ، وسطانیہ سے مختلف آنکڑوں کے لیے معلوم کیا جاتا ہے۔

15.4.1 درمیانہ انحراف غیر جماعتی آنکڑوں کے لیے (Mean Deviation for ungrouped data)

مان لیجئے n مشاہدے $x_1, x_2, x_3, ...x_n$ ہیں۔ ذیل اقدامات درمیانہ انحراف ، درمیانہ سے یا وسطانیہ سے نکالنے میں ملوث ہیں۔

قدم 1 مرکزی ملان کا پیمانہ/کی ناپ معلوم کیجئے جس سے ہمیں درمیانہ انحراف معلوم کرنا ہے۔ مان لیجئے یہ 'a' ہے۔

قدم 2 ہر ایک x_1 جھکاؤ a سے معلوم کیجئے یعنی $x_1 - a, x_2 - a, x_3 - a, ...x_n - a$

قدم 3 انحراف کی قدر مطلق معلوم کیجئے، یعنی نفی (–) کے نشان کو ہٹا دیجئے اگر موجود ہے

یعنی $|x_1 - a|, |x_2 - a|, |x_3 - a|, ..., |x_n - a|$

قدم 4 انحراف قدر مطلق کا درمیانہ معلوم کیجئے۔ اس کا مطلب ہے درمیانہ انحراف a سے یعنی،

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

اس طرح $\text{M.D.}(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n} |X_i - \bar{x}|$ جہاں $\bar{x}$ = درمیانہ ہو

اور $\text{M.D.(M)} = \frac{1}{n}\sum_{i=1}^{n} |X_i - \text{M}|$ جہاں **M** = وسطانیہ

> **نوٹ** اس باب میں ہم علامت M کا استعمال وسطانیہ کو ظاہر کرنے کے لیے کریں گے جب تک کے بتایا نہ جائے۔
> اب ہمیں ذیل مثالوں میں اوپر دیئے ہوئے طریقے کے اقدامات کا تصور کریں گے۔

مثال 1 مندرجہ ذیل آنکڑوں کے لیے درمیانہ سے درمیانہ انحراف معلوم کیجیے:

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

حل ہم ایک کے بعد ایک قدم بڑھائیں گے اور مندرجہ ذیل حاصل کریں گے۔:

قدم 1 دیئے ہوئے آنکڑوں کے درمیانہ ہے:

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

قدم 2 متعلقہ مشاہدوں کا درمیانہ $\bar{x}$ سے انحراف یعنی $x_i - \bar{x}$ ہیں۔

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

یا

$$-3, -2, 1, 3, 4, -5, -1, 3$$

قدم 3 جھکاؤ کی مطلق قدریں یعنی $|x_i - \bar{x}|$ ہیں۔

$$3, 2, 1, 3, 4, 5, 1, 3$$

قدم 4 مطلوبہ درمیانہ انحراف درمیانہ سے ہے۔

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

$$= \frac{3+2+1+3+4+5+1+3}{8} = \frac{22}{8} = 2.75$$

نوٹ ہر بار اقدامات لکھنے کی بجائے ہم حساب کریں گے، ترتیب میں بغیر اقدامات لکھے ہوئے۔

مثال 2 مندرجہ ذیل آنکڑوں کے درمیانہ سے درمیانہ انحراف معلوم کیجیے:

$$= \frac{3+2+1+3+4+5+1+3}{8} = \frac{22}{8} = 2.75$$

حل ہمیں دیئے ہوئے آنکڑوں کا درمیانہ $\bar{x}$ معلوم کرنا ہے۔

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

انحراف کے درمیانہ سے متعلقہ مطلق قدریں یعنی $|x_i - \bar{x}|$ ہیں۔

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

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

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

مثال 3 مندرجہ ذیل آنکڑوں کے لیے درمیانہ انحراف وسطانیہ سے معلوم کیجیے۔

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

حل یہاں مشاہدوں کی تعداد 11 ہے جو کہ طاق ہے۔ آنکڑوں کو بڑھتی ہوئی ترتیب میں رکھنے پر، ہمارے پاس ہے۔

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

اب $\left(\frac{11+1}{2}\right)^{th}$ = وسطانیہ یا چھٹا (6th) مشاہدہ = 9

وسطانیہ سے مطلق قدروں کی مطلق قدریں یعنی $x_i - \text{M}$ ہیں۔

6, 6, 5, 4, 2, 0, 1, 3, 9, 10, 12

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

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

15.4.2 جماعتی آنکڑوں کے لیے درمیانہ انحراف *(Mean Deviation for grouped data)*

ہم جانتے ہیں کہ جماعتی آنکڑوں کو دو طریقوں سے رکھا جاسکتا ہے:

(a) غیر مسلسل تواتر تقسیم Discrete frequency destribution

(b) مسلسل تواتر تقسیم Continuous frequency distribution

ہمیں دونوں طرح آنکڑوں کا درمیانہ انحراف (فرق) معلوم کرنے کے طریقے پر بات چیت کرنی چاہیے۔

(a) تعداد کی الگ تقسیم (Discrete frequency distribution)

مان لیجئے دیئے ہوئے آنکڑوں میں n مختلف قدریں $x_1, x_2, ..., x_n$ بالترتیب $f_1, f_2, ..., f_n$ تواتر کے ساتھ واقع ہورہی ہیں۔ان آنکڑوں کو جدولی شکل میں ظاہر کیا جاسکتا ہے جیسا کہ نیچے دیا گیا ہے،اور غیر مسلسل تواتر تقسیم کہا جاتا ہے۔

$$x : x_1 \quad x_2 \quad x_3 \quad .... \quad x_n$$

$$f : f_1 \quad f_2 \quad f_3 \quad .... \quad f_n$$

(i) درمیانہ انحراف درمیانہ سے (Mean deviation about mean)

دیئے ہوئے آنکڑوں کے لیے فارمولہ استعمال کر کے سب پہلے ہم درمیانہ $\bar{x}$ معلوم کرتے ہیں۔

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

جہاں $\sum_{i=1}^{n} x_i f_i$ ، x_i مشاہدوں کے ان کے متعلقہ تواتر کے ساتھ حاصل ضرب کے مجموعہ کو ظاہر کرتا ہے اور $N = \sum_{i=1}^{n} f_i$ تعداد کا جوڑا ہے۔

تب ہم، x_i مشاہدوں کا درمیانہ $\bar{x}$ سے فرق معلوم کرتے ہیں اور ان کی مطلق قدریں لیتے ہیں یعنی $|x_i - \bar{x}|$ تمام $i = 1, 2, ..., n$ کے لیے۔

اس کے بعد،فرق کی تمام مطلق قدروں کا درمیانہ معلوم کیجئے جو کہ درمیانہ کا مطلوبہ درمیانہ سے فرق ہے۔اس طرح

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

(ii) درمیانہ انحراف وسطانیہ سے (Mean deviation about mean)

وسطانیہ سے درمیانہ فرق معلوم کرنے کے لیے ہم دیئے ہوئے تواتر تقسیم کا وسطانیہ معلوم کرتے ہیں۔ اس کے لیے مشاہدات کو بڑھتی ہوئی ترتیب میں رکھا جاتا ہے۔ اس کے بعد مجموعی تواتر (Cumutative frequencies) کو حاصل کیا جاتا ہے۔ تب ہم ان مشاہدات کی پہچان کرتے ہیں جن کا تعداد مجموعی تواتر یا تو $\frac{N}{2}$ کے برابر ہو یا بڑا ہو۔ جہاں **N** تواتر کا مجموعہ ہے۔ یہ مشاہدوں کی قدر آنکڑوں کے درمیان میں واقع ہے، اس لیے یہ مطلوبہ وسطانیہ ہے۔ وسطانیہ معلوم کرنے کے بعد، ہمیں وسطانیہ سے مطلق قدروں کے فرق کا درمیانہ حاصل ہوتا ہے۔

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

مثال 4 ذیل آنکڑوں کے لیے درمیانہ سے درمیانہ انحراف معلوم کیجئے۔

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

حل ہمیں دیئے ہوئے آنکڑوں کا جدول 15.1 بنانا چاہیے اور دوسرے کالم حساب لگانے کے بعد بنانے چاہیں۔

جدول 15.1

x_i	f_i	$f_i x_i$	$\lvert x_i - \bar{x}\rvert$	$\lvert f_i\rvert\ x_i - \bar{x}$
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

$$N = \sum_{i=1}^{6} f_i = 40, \sum_{i=1}^{6} f_i x_i = 300, \sum_{i=1}^{6} f_i |x_i - \bar{x}| = 92$$

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

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

مثال 5 مندرجہ ذیل آنکڑوں کے لیے وسطانیہ سے درمیانہ انحراف معلوم کیجیے۔

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

حل دیئے مشاہدات پہلے ہی بڑھتی ہوئی ترتیب میں ہیں۔ دیئے ہوئے آنکڑوں میں ایک قطار نظیری مجموعی تواتر میں جمع کرنے پر ہمیں حاصل ہوتا ہے۔(جدول 15.2)

جدول 15.2

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

اب $N = 30$ ، جو کہ جفت ہے۔

وسطانیہ 15 ویں اور 16 ویں مشاہدات کا درمیانہ ہے۔ یہ دونوں مشاہدات مجموعی تواتر 18 ویں میں واقع ہیں، جس کے لیے نظیری (Corresponding) مشاہدہ 13 ہے۔

اس لیے $\frac{13+13}{2} = 13$ = M وسطانیہ = 16 واں مشاہدہ + 15 واں مشاہدہ

اب وسطانیہ سے مطلق قدروں کا فرق، یعنی $|x_i - M|$ جدول 15.3 میں دکھائے گئے ہیں۔

جدول 15.3

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

$$\sum_{i=1}^{8} f_i = 30 \quad \text{and} \quad \sum_{i=1}^{8} f_i\lvert x_i - M\rvert = 149$$

اس لیے
$$\text{M.D.}(M) = \frac{1}{N}\sum_{i=1}^{8} f_i\lvert x_i - M\rvert$$
$$= \frac{1}{30} \times 149 = 4.97$$

(b) مسلسل تواتر تقسیم (Continuous frequency distribution)

تواتر کی لگاتار تقسیم ایک سلسلی ہے جس میں آنکڑوں کو مختلف کلاس، وقفہ میں بانٹا گیا ہے بغیر فاصلے کے اپنی متعلقہ تواتر کے ساتھ۔

مثال کے طور پر، 100 طلبا کے ذریعے حاصل کئے گئے نمبر ایک مسلسل تواتر تقسیم میں ظاہر کیے گئے ہیں جیسا کہ نیچے دیئے گیا ہے:

حاصل کئے گئے نمبر	0-10	10-20	20-30	30-40	40-50	50-60
طلباء کی تعداد	12	18	27	20	17	6

درمیانہ سے درمیانہ انحراف (Mean deviation from Mean)

جب درمیانہ سے ایک تعداد کی لگاتار تقسیم کا حساب لگایا جاتا ہے، ہم نے یہ مانا کہ ہر جماعت میں تعداد درمیانی نقطے پر مرکز ہے۔ یہاں بھی یہ ہر جماعت کا درمیانی نقطہ لکھتے ہیں اور آگے بڑھتے ہیں جیسا کہ تعداد کی الگ تقسیم کے لیے کیا جاتا ہے درمیانہ فرق معلوم کرنے کے لیے۔

ہم ذیل مثال لیتے ہیں۔

مثال 6 مندرجہ ذیل آنکڑوں کے لیے درمیانہ کے قریب درمیانہ انحراف معلوم کیجیے۔

حاصل شدہ نمبر	10-20	20-30	30-40	40-50	50-60	60-70	70-80
طلباء کی تعداد	2	3	8	14	8	3	2

حل دیئے ہوئے آنکڑوں سے ہم مندرجہ ذیل جدول 15.4 بناتے ہیں۔

جدول 15.4

حاصل شدہ نمبر	طلباء کی تعداد f_i	Mid-point x_i	$f_i x_i$	$\|x_i - \bar{x}\|$	$f_i \|x_i - \bar{x}\|$
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

یہاں $N = \sum_{i=1}^{7} f_i = 40, \sum_{i=1}^{7} f_i x_i = 1800, \sum_{i=1}^{7} f_i |x_i - \bar{x}| = 400$

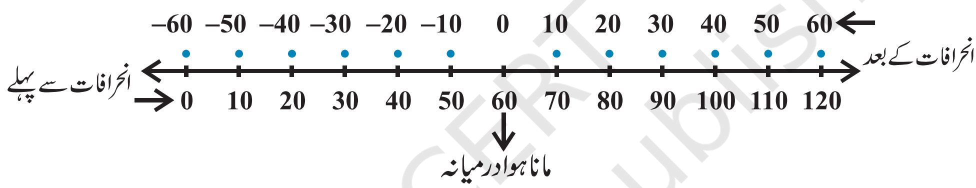
اس لیے $\bar{x} = \frac{1}{N}\sum_{i=1}^{7} f_i x_i = \frac{1800}{40} = 45$

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

درمیانہ سے مختصر طریقے سے درمیانہ انحراف کا حساب لگانا
(Shortcut method for calculating mean deviation about mean)

ہم $\overline{x}$ کا حساب لگانے کے لیے اکتا دینے والے طریقے سے بچ سکتے ہیں ذیل Step deviation کا طریقہ استعمال کر کے۔ یاد کیجیے کہ اس طریقے میں ہم ایک فرضی درمیانہ (assumed mean) لیتے ہیں جو کہ آنکڑوں میں درمیانی یا اس کے قریب ہوتا ہے۔ تب مشاہدات کا فرق (یا جماعتوں کے درمیانی نقطے) اور مانے ہوئے درمیانہ سے لیتے ہیں۔ یہ کچھ نہیں ہے بلکہ عددی خط پر صفر سے مانے ہوئے درمیان تک مبدا کو بدلنا ہے، جیسا کہ شکل 15.3 میں دیا گیا ہے۔

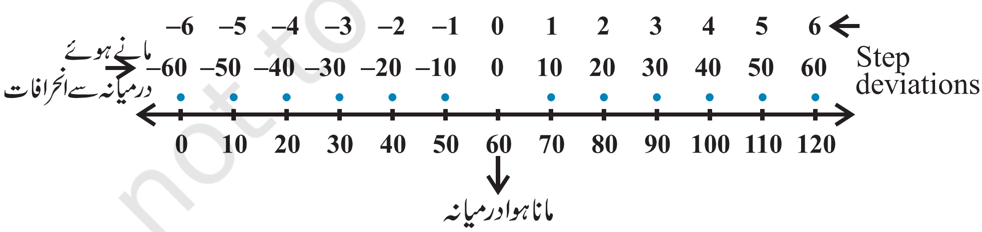
اگر تمام انحرافات کا ایک مشترک اجزائے ضربی ہے ہم انحرافات کو مختصر کرنے کے لیے اسے مشترکہ اجزائے ضربی سے



شکل 15.3

تقسیم کرتے ہیں۔ انھیں Step deviations کہا جاتا ہے۔ Step deviations لینے کا طریقہ عددی خط پر اسکیل بدلنا ہے جیسا کہ شکل 15.4 میں دکھایا گیا ہے۔

فرق اور Step deviation مشاہدات کے سائز کو چھوٹا کر دیتے ہیں، تا کہ حساب لگانا مطلب حاصل ضرب، وغیرہ



شکل 15.4

آسان ہو جاتا ہے۔ مان لیجیے، نیا متغیر $d_i = \frac{x_i - a}{h}$ سے ظاہر کیا جاتا ہے جہاں 'a' مانا ہوا درمیانہ ہے اور h مشترکہ اجزائے ضربی۔

تب درمیانہ $\overline{x}$ ، Step deviation طریقے کے ذریعے اس سے دیا گیا ہے۔

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

ہم مثال 6 کے آنکڑے لیتے ہیں اور Step deviation کے طریقے کا استعمال کر کے درمیانہ انحراف معلوم کرتے ہیں۔

مان لیا کہ درمیانہ $a = 45$ اور $h = 10$ لیجیے اور مندرجہ ذیل جدول 15.5 ہے۔

جدول 15.5

حاصل شدہ نمبر	طلباء کی تعداد f_1	Mid-Points x_i	$d_i = \frac{x_i - 45}{10}$	$f_i d_i$	$\lvert x_i - \overline{x}\rvert$	$f_i \lvert x_i - \overline{x}\rvert$
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	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

اس لیے

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

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

اور

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

> نوٹ Step deviation کا طریقہ $\bar{x}$ کا حساب لگانے کے لیے استعمال کیا جاتا ہے۔ باقی سب کارکردگی کا طریقہ وہی ہے۔

(ii) وسطانیہ سے درمیانہ انحراف (Mean deviation about medain)

وسطانیہ سے درمیانہ انحراف ایک مسلسل تواتر تقسیم سے معلوم کرنے کا طریقہ کار بالکل ایسا ہی ہے جیسا کہ ہم نے درمیانہ سے درمیانہ انحراف معلوم کرنے کے لیے کیا ہے۔ اس ردوبدل میں صرف ایک ہی فرق ہے جو کہ درمیانہ کو وسطانیہ سے بدلنے میں جب فرق لیا جاتا ہے۔

ہمیں اس طریقے کو دہرانا چاہیے جہاں مسلسل تواتر تقسیم کے لیے استعمال کیا جاتا ہے۔ پہلے آنکڑوں کو بڑھتی ہوئی ترتیب میں لکھا جاتا ہے۔ تب پہلے اس جماعت کی نشاندہی کرتے ہیں جہاں وسطانیہ واقع ہے۔ (وسطانیہ جماعت) اور پھر مسلسل تواتر تقسیم حاصل ہوتی ہے اور پھر ضابطہ (Formula) کا استعمال کرتے ہیں۔

$$\text{وسطانیہ Median } = l + \frac{\frac{N}{2} - C}{f} \times h$$

جہاں وسطانیہ جماعت، جماعت کا وقفہ جس کی مجموعی تعداد زرا $\frac{N}{2}$ سے بڑی ہے یا برابر ہے۔ N تواتر کا مجموعہ ہے، l، f، h اور C بالترتیب نچلی انتہا، تواتر، وسطانیہ جماعت کی چوڑائی (width) اور C جماعت سے بالکل پہلے کی وسطانیہ کلاس کی مجموعی تواتر ہے۔ وسطانیہ نکالنے کے بعد درمیانی نقطہ x_i کے فرق کی مطلق قیمتیں ہر جماعت کے لیے وسطانیہ سے یعنی $|x_i - M|$ سے حاصل کی جاتی ہے۔

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

طریقہ کی مندرجہ ذیل مثال میں تشریح بیان کی گئی ہے۔

مثال 7 مندرجہ ذیل آنکڑوں کے لیے وسطانیہ سے درمیانہ انحراف معلوم کیجیے۔

جماعت Class	0-10	10-20	20-30	30-40	40-50	50-60
Frequency	6	7	15	16	4	2

حل دیئے ہوئے آنکڑوں سے مندرجہ ذیل جدول 15.6 تیار کیجئے۔

جدول 15.6

Class	Frequency f_i	Cummulative frequency $(c.f.)$	Mid-Point x_i	$\|x_i - \text{Med.}\|$	$f_i\|x_i - \text{Med.}\|$
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

وہ کلاس وقفہ جس میں $\frac{N}{2}^{th}$ یا 25^{th} اندراج ہے 20-30 ہے۔اس لیے 20-30 وسطانیہ کلاس ہے۔ہم جانتے ہیں کہ

$$\text{وسطانیہ Medain} = l + \frac{\frac{N}{2} - C}{f} \times h$$

یہاں $l = 20, C = 13, f = 15, h = 10$ اور $N = 50$

اس لیے $$Median = 20 + \frac{25-13}{15} \times 10 = 20 + 8 = 28$$

اس طرح وسطانیہ،درمیانہ اس طرح دیا گیا ہے۔

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

مشق 15.1

1 تا 2 مشقں میں آنکڑوں کے لیے درمیانہ انحراف درمیانہ سے معلوم کیجیے۔

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

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

3 تا 4 مشقوں میں آنکڑوں کے لیے درمیانہ انحراف، وسطانیہ کے سے معلوم کیجیے۔

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

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

5 تا 6 مشقوں میں آنکڑوں کے لیے درمیانہ انحراف، درمیانہ سے معلوم کیجیے۔

.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

7 تا 8 مشقوں میں آنکڑوں کے لیے درمیانہ انحراف، وسطانیہ سے معلوم کیجیے۔

.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

9 تا 10 مشقوں میں آنکڑوں کے لیے درمیانہ انحراف، درمیانہ کے سے معلوم کیجیے۔

.9

روزآنہ کی آمدنی	700-800	600-700	500-600	400-500	300-400	200-300	100-200	0-100	
انسانوں کی تعداد	3	4	5	7	0	1	9	8	4

10.

اونچائی سینٹی میٹر میں	95-105	105-115	115-125	125-135	135-145	145-155
لڑکوں کی تعداد	9	13	26	30	12	10

11. ذیل آنکڑوں کے لیے درمیانہ انحراف وسطانیہ سے معلوم کیجئے۔

نمبر	0-10	10-20	20-30	30-40	40-50	50-60
لڑکیوں کی تعداد	6	8	14	16	4	2

12. نیچے دیئے گئے 100 انسانوں کی عمر کی تقسیم کے لیے درمیانہ انحراف وسطانیہ سے معلوم کیجیے۔

عمر	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55
تعداد	5	6	12	14	26	12	16	9

[**اشارہ** دیئے ہوئے آنکڑوں میں سے 0.5 کو نچلی انتہا میں سے تفریق کرنے اور 0.5 اوپری انتہا میں جمع کرنے پر مسلسل تواتر تقسیم ہر ایک کلاس وقفہ میں]

15.4.3 درمیانہ انحراف کی احاطہ بندی (پابندی) (*Limitations of mean deviation*)

کسی سلسلئ میں جہاں تغیر کا درجہ بہت بلند ہے، وسطانیہ مرکزی ملان کا نمائندہ نہیں ہے۔ اس لیے وسطانیہ سے درمیانہ انحراف کا حل اس طرح کی سلسلئ کے لیے مکمل طور پر قابلِ اعتبار نہیں ہے۔

درمیانہ سے انحراف کا جوڑ (منفی نشان کو نظر انداز کرتے ہوئے) وسطانیہ سے انحراف کے جوڑ سے زیادہ ہے۔ اس لیے درمیانہ انحراف درمیانہ سے بہت زیادہ سائن ٹیفک نہیں ہے۔ اس طرح بہت سے کیسس میں ممکن ہے درمیانہ فرق غیر مطمئن نتیجے دے سکتا ہے۔ ساتھ ہی درمیانہ انحراف فرق کی مطلق قدروں کی بنیاد پر نکالا جاتا ہے۔ اور اس لیے، الجبری عمل کے لیے کیا جاسکتا، اس کا مطلب یہ کہ ہمارے پاس کچھ اور دوسرے انتشار معلوم کرنے کے پیمانے ہونے چاہیں۔ معیاری انحراف اس طرح کا ایک انتشاری پیمانہ ہے۔

15.5 عدم مطابقت اور معیاری انحراف (Variance and Standard Deviation)

یاد کیجئے کہ جب درمیانہ انحراف درمیانہ یا وسطانیہ سے نکالا جاتا ہے، فرق کی مطلق قدریں لی گئیں تھیں۔ درمیانہ انحراف بامعنی

بنانے کے لیے مطلق قدریں لی گئیں تھیں، ورنہ انحرافات آپس میں مسنوخ ہوجائیں گے۔

اس پریشانی پر فتح حاصل کرنے کے لیے جو انحرافات کی علامت کی وجہ سے ہوتی ہے، دوسرا طریقہ یہ ہے کہ تمام انحرافات کے مربعے کیے جائیں۔ صاف طور پر ان فرق کے تمام مربع غیر منفی ہیں۔ مان لیجئے $x_1, x_2, x_3, \ldots, x_n$ مشاہدے ہیں اور $\bar{x}$ انکا درمیانہ ہے۔ تب

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

اگر یہ مجموعہ صفر ہے، تب ہر ایک $(x_i - \bar{x})$ کو صفر ہونا گا۔ اس کا مطلب ہے کہ ابھی ایسا کوئی انتشار نہیں ہے کہ تمام مشاہدات درمیانہ کے برابر ہیں۔

اگر $\sum_{i=1}^{n}(x_i-\bar{x})^2$ چھوٹا ہے، یہ ظاہر کرتا ہے کہ مشاہدے $x_1, x_2, x_3, \ldots, x_n$ درمیانہ $\bar{x}$ کے قریب ہیں اور اس لیے انتشار کا درجہ چھوٹا ہے۔ اس کے برعکس، اگر یہ مجموعہ بڑا ہے، مشاہدوں کا انتشار کا درجہ درمیانہ $\bar{x}$ سے بڑا۔ اس لیے کیا ہم کہہ سکتے ہیں کہ مجموعہ $\sum_{i=1}^{n}(x_i-\bar{x})^2$ ایک قابلِ قبول انتشار کا یا انحراف کا مقدار نما (indecator) ہے؟

ہم چھ مشاہدات 55 ,45 ,35 ,25 ,15 ,5 کے سیٹ A کو لیتے ہیں۔ ان مشاہدوں کا درمیانہ $\bar{x} = 30$ ہے۔ اس سیٹ کے 30 سے لیے گیے انحرافات کے مربعوں کا جوڑ ہے۔

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

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

اب ہم ایک دوسرا سیٹ B 31 مشاہدوں 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 لیتے ہیں۔ ان مشاہدات کا درمیانہ $\bar{y} = 30$ ہے۔ یہ نوٹ کر لیجئے کہ مشاہدات کے سیٹ A اور B کا درمیانہ 30 ہے۔

اب، سیٹ B کے لیے درمیانہ $\bar{y}$ سے مشاہدوں کے فرق کے مربعوں کا جوڑ دیا گیا ہے۔

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

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

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

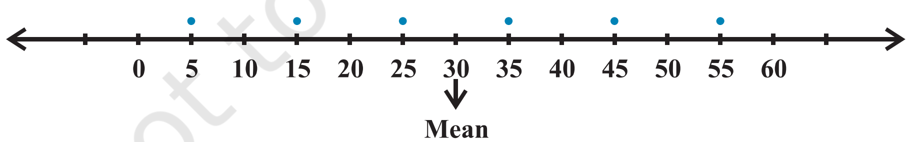
$$= 2\times\frac{15\times(15+1)\,(30+1)}{6} = 5\times16\times31 = 2480$$

(کیونکہ x طبعی اعداد کے مربعوں کا جوڑ $= \dfrac{n\,(n+1)\,(2n+1)}{6}$، یہاں $n = 15$)

اگر $\sum_{i=1}^{n}(x_i-\bar{x})^2$ آسانی سے ہمارے انتشار کا پیمانہ ہے یا درمیانہ کے قریب پھیلا ہوا ہے، ہم یہ کہنے پر مجبور ہوں گے کہ 6 مشاہدوں والا سیٹ A کم انتشار رکھتا ہے درمیانہ کے قریب 31 مشاہدوں والے سیٹ B ہے، حالانکہ سیٹ A میں مشاہدوں سے زیادہ پھیلے ہوئے ہیں۔(فرق کا دائرہ 25- سے 25 تک ہے) بجائے سیٹ B کے (جہاں فرق کا دائرہ 15– سے 15 تک ہے)

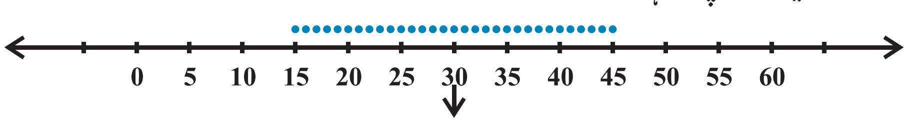
یہ مندرجہ ذیل اشکال سے بھی صاف ہے۔

سیٹ A کے لیے ہمارے پاس ہے۔



شکل 15.5

سیٹ B کے لیے ہمارے پاس ہے۔



شکل 15.6

اس طرح، ہم کہہ سکتے ہیں کہ درمیانہ سے فرق کے مربعوں کا جوڑ انتشار کا صحیح پیمانہ نہیں ہے۔اس پریشانی پر عبور حاصل کرنے کے لیے ہم فرق کے مربعوں کا درمیانہ لیتے ہیں یعنی، ہم $\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ لیتے ہیں۔سیٹ A کے کیس میں، ہمارے پاس درمیانہ $= \frac{1}{6}\times 175 = 291.67$ ہے اور سیٹ B کیس ملیں یہ $\frac{1}{31}\times 2480 = 80$ ہے یہ ظاہر کرنا ہے کہ بکھرا ہوا یا انتشار سیٹ A میں سیٹ B سے زیادہ ہے۔جس کی تفریق دونوں سیٹ کے جیومیٹریائی اظہار سے ہوتی ہے۔

اس طرح، ہم اسے $\frac{1}{n}\sum(x_i-\bar{x})^2$ مقدار کے طور پر لے سکتے ہیں جو انتشار کی صحیح ناپ کے طرف لے جاتا ہے۔ یہ عدد یعنی درمیانہ سے لیے گیے انحراف کے مربعوں کا درمیانہ عدم مطابقت **(Variance)** کہلاتا ہے۔اس کو σ^2 سے ظاہر کرتے ہیں جسے (سگما مربع) (Sigma Square) پڑھتے ہیں۔اس لیے، n مشاہدوں $x_1, x_2, x_3, \ldots, x_n$ کی عدم مطابقت کو $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ سے ظاہر کیا جاتا ہے۔

15.5.1 معیاری انحراف (*Standard Deviation*)

عدم مطابقت کو حل کرنے میں، ہم نے یہ حل نکالا ہے کہ انفرادی مشاہدات x_i کی اکائیاں اور ان کی اکائی کا درمیانہ $\bar{x}$ انحراف سے مختلف ہے، کیونکہ انحراف میں $(x_i-\bar{x})$ کے مربعوں کا مجموعہ ملوث ہے۔اس وجہ کے لیے، انتشار کا صحیح پیمانہ ایک مشاہدات کے سیٹ کے درمیانہ سے عدم مطابقت مثبت جزر المربع ہے اور اسے معیاری انحراف کہا جاتا ہے۔اس لیے معیاری انتشار کو 'σ' سے ظاہر کیا جاتا ہے اور اسے اس طرح دیا گیا ہے۔

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

ہم عدم مطابقت کے حل کی تشریح کرنے کے لیے ذیل مثال لیتے ہیں اور اس طرح، غیر جماتی آنکڑوں کا معیاری انحراف۔

مثال 8 ذیل آنکڑوں کی عدم مطابقت (Variance) معلوم کیجیے۔

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

حل دیئے ہوئے آنکڑوں سے ہم ذیل جدول 15.7 بنا سکتے ہیں Step deviation طریقے سے 14 کو درمیانہ مان کر درمیانہ حل کیا گیا ہے۔مشاہدوں کی تعداد $n = 10$ ہے۔

جدول 15.7

x_i	$d_i = \frac{x_i - 14}{2}$	Deviation from mean $(x_i - \bar{x})$	$(x_i \bar{x})$
6	– 4	– 9	81
8	– 3	– 7	49
12	– 2	– 5	25
12	– 1	– 3	9
14	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

اس لیے $\text{Mean } \bar{x} = \text{assumed mean} + \frac{\sum_{i=1}^{n} d_i}{n} \times h = 14 + \frac{5}{10} \times 2 = 15$

اور $var\,iance\,(\sigma^2) = \frac{1}{n}\sum_{i=1}^{10}(x_i - \bar{x})^2 = \frac{1}{10} \times 330 = 33$

اس لیے معیاری انحراف $= (\sigma) = \sqrt{33} = 5.74$

15.5.2 غیر مسلسل تواتر تقسیم کا معیاری انحراف (Standard Deviation of discrete frequency distribution)

مان لیجیے، دی ہوئی غیر مسلسل تواتر تقسیم ہے۔

$$
\begin{array}{llllll}
x: & x_i, & x_2, & x_3 & ,\dots, & x_n \\
y: & f_i, & f_2, & f_3 & ,\dots, & f_n
\end{array}
$$

اس کیس میں معیاری انحراف

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

جہاں $N = \sum_{i=1}^{n} f_i$

ہمیں مندرجہ ذیل مثال لینی چاہیے۔

مثال 9 ذیل آنکڑوں کی عدم مطابقت (Variance) اور معیاری انحراف (Standard deviation) معلوم کیجیے۔

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

حل آنکڑوں کو جدول شکل 15.8 میں ظاہر کرنے پر، ہمیں حاصل ہوتا ہے۔

جدول 15.8

x_i	f_i	$f_i x_i$	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	$f_i(x_i - \overline{x})^2$
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	8	324	324
	30	420			1374

$$N = 30, \sum_{i=1}^{7} f_i x_i = 420, \ \sum_{i=1}^{7} f_i\left(x_i - \overline{x}\right)^2 = 1374$$

اس لیے $\overline{x} = \dfrac{\sum_{i=1}^{7} f_i x_i}{N} = \dfrac{1}{30} \times 420 = 14$

اس طرح $\text{variance}\,(\sigma^2) = \frac{1}{N} \sum_{i=1}^{7} f_i (x_i - \bar{x})^2$

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

اور $\text{Stanard deviation } (\sigma) = \sqrt{45.8} = 6.77$

15.5.3 مسلسل تواتر تقسیم کا معیاری انحراف (Standard Deviation of a Continuous frequency distribution)

دی ہوئی مسلسل تواتر تقسیم غیر مسلسل تواتر تقسیم میں ظاہر کیا جا سکتا ہے۔ ہر کلاس کو اس کے درمیانی نقطے سے تبدیل کر کے۔ تب معیاری انحراف کو اس عمل سے حل کیا جا سکتا ہے جو کہ غیر مسلسل تواتر تقسیم میں استعمال کیا گیا ہے۔

اگر ایک تواتر تقسیم n کلاسوں کا ہے جس کی ہر کلاس اس کے درمیانی نقطے x_i سے define کی گئی ہے تواتر f_i کے ساتھ، معیاری انحراف ذیل فارمولے کی مدد سے حاصل کیا جائے گا۔

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

جہاں $\bar{x}$ درمیانہ ہے اور $N = \sum_{i=1}^{n} f_i$

معیاری انحراف کے لیے ایک دوسرا فارمولہ (Another formula for standard deviation)

ہم جانتے ہیں کہ:

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

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

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

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

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

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

اس طرح $$\text{(standard deviation)}(\sigma) = \frac{1}{N}\sqrt{N\sum_{i=1}^{n} f_i x_i^2 - \left(\sum_{i=1}^{n} f_i x_i\right)^2}$$

مثال 10 مندرجہ ذیل تقسیم کے لیے درمیانہ، عدم مطابقت اور معیاری انحراف معلوم کیجیے۔

کلاس	30-40	40-50	50-60	60-70	70-80	80-90	90-100
تعداد	3	7	12	15	8	3	2

حل دیئے ہوئے آنکڑوں سے ہم نے مندرجہ ذیل جدول 15.9 تیار کیا ہے۔

Class	Frequency (f_i)	Mid-point (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	528	1587
90-100	2	95	190	1089	2178
	50		3100		10050

اس طرح $$\text{Mean } \bar{x} = \frac{1}{N}\sum_{i=1}^{7} f_i x_i = \frac{3100}{50} = 62$$

$$\text{variance}\left(\sigma^2\right)=\frac{1}{\text{N}}\sum_{i=1}^{7} f_i(x_i-\overline{x})^2$$

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

مثال 11 مندرجہ ذیل آنکڑوں کے لیے معیاری انحراف معلوم کیجیے:

x_i	3	8	13	18	32
f_i	7	10	15	10	6

حل ہمیں مندرجہ ذیل جدول 15.10 بنانا چاہیے۔

جدول 15.10

x_i	f_i	f_ix_i	x_i^2	$f_ix_i^2$
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

اب فارمولہ (3) سے، ہمارے پاس ہے۔

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

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

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

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

اس لیے معیاری فرق $(\sigma) = 6.12$

15.5.4 عدم مطابقت اور معیاری انحراف معلوم کرنے کا مختصر طریقہ بعض اوقات مسلسل تقسیم میں x_i کی قدریں یا مختلف کلاسوں کا درمیانہ نقطہ x_i، مسلسل تقسیم میں کافی بڑے ہوتے ہیں اور اس لیے درمیانہ اور عدم مطابقت کا حساب لگانا اکتا دینے والا اور زیادہ وقت لینے والا ہو جاتا ہے۔ Step deviations کے طریقے کا استعمال کرکے، عمل کو آسان کیا جانا ممکن ہے۔

مان لیجئے مانا ہوا درمیانہ 'A' ہے اور اسکیل $\frac{1}{h}$ گنا چھوٹا کرلیا جائے۔ (h کلاس وقفہ کی چوڑائی ہے)

مان لیجئے Step deviation یا نئی قدر y_i ہے۔

یعنی $x_i = A+hy_i$ $\quad y_i = \frac{x_i - A}{h}$ ....(1)

ہم جانتے ہیں کہ $\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{N}$ ....(2)

x_i کو (1) سے (2) میں رکھنے پر، ہمیں حاصل ہوتا ہے۔

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

$$= \frac{1}{N}\left(\sum_{i=1}^{n} f_1 A + \sum_{i=1}^{n} hf_i y_i\right) = \left(\frac{1}{N} A\sum_{i=1}^{n} f_i + h\sum_{i=1}^{n} f_i y_i\right)$$

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

اسطرح $\bar{x} = A + h\bar{y}$ ....(3)

اب متغیر x کی عدم مطابقت (Variance) ہے $\sigma_x^2 = \frac{1}{N}\sum_{i=1}^{n} f_i\left(x_i - \bar{x}\right)_2$

(1) اور (3) کا استعمال کرنے پر $= \frac{1}{N}\sum_{i=1}^{n} f_i\left(A + hy_i - A - h\bar{y}\right)^2$

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

متغیرہ y_i کی عدم مطابقت $= h^2 \times \frac{h^2}{N}\sum_{i=1}^{n} f_i\left(y_i - \bar{y}\right)^2$ =

یعنی $\sigma_x^2 = h^2\sigma_y^2$

یا $\sigma_x = h\sigma_y$ ....(4)

(3)اور(4)سے، ہمارے پاس ہے۔

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

فارمولہ (5) کو استعمال کر کے ہم مثال (11) کو مختصر طریقے سے حل کر سکتے ہیں۔

مثال 12 مندرجہ ذیل تقسیم کے لئے درمیانہ، عدم مطابقت اور معیاری انحراف معلوم کیجئے۔

Classes	30–40	40–50	50–60	60–70	70–80	80–90	90–100
Frequency	3	7	12	15	8	3	2

حل مان لیجئے مانا ہوا درمیانہ A=65. یہاں h=10

دئے ہوئے آنکڑوں سے ہم نے مندرجہ ذیل جدول 15.11 حاصل کیا ہے۔

جدول 15.11

Class	Frequency f_i	Mid-point 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	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

اس طرح $\bar{x} = A + \frac{\sum fiyi}{50} \times h = 65 - \frac{15}{50} = 62$

Variance (عدم مطابقت) $\sigma^2 = \frac{h^2}{N^2} N\sum f_i y_i^2 - \left(\sum f_i y_i\right)^2$

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

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

معیاری انحراف

مشق 15.2

مشق 1 تا 5 میں ہر ایک آنکڑے کے لئے درمیانہ اور عدم مطابقت معلوم کیجئے

1. 6, 7, 10, 12, 13, 4, 8, 12
2. پہلے n طبعی اعداد کیلئے
3. پہلے 10، 3 کے اضعاف کے لئے
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. مختصر طریقہ کا استعمال کرکے درمیانہ اور معیاری انحراف معلوم کیجئے۔

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

7. مشق 7 اور 8 میں ذیل تواتر تقسیم کے لئے درمیانہ اور عدم مطابقت معلوم کیجئے۔

Classes	0–30	30–60	60–90	90–120	120–150	150–180	180–210
Frequencis	2	3	5	10	3	5	2

8.

Classes	0–10	10–20	20–30	30–40	40–50
Frequencies	5	8	15	16	6

9. مختصر طریقہ کا استعمال کرکے درمیانہ، عدم مطاقت اور معیاری انحراف معلوم کیجئے۔

H in	70–75	75–80	80–85	85–90	90–95	95–100	100–105	105–110	110–115
No.ch	3	4	7	7	15	9	6	6	3

10. نیچے ایک ڈیزائن میں کھینچے گئے دائروں کے قطر (mm میں) دئے گئے ہیں۔

Diameters	33–36	37–40	41–44	45–48	49–52
No of circl.	15	17	21	22	25

دائروں کا درمیانہ قطر اور معیاری انحراف معلوم کیجئے۔

[اشارہ پہلے آنکڑوں کو لگاتار بنائیے اسطرح کلاس بنا کر 36.5–40.5, 40.5–44.5,44.5–48.5,48.5–52.5
32.5–36.5, اور پھر آگے بڑھیں]

15.6 تواتر تقسیم کا تجزیہ (Analysis and Frequency Distributions)

پچھلے سیکشن میں ہم انتشار کو ناپنے کے کچھ طریقوں کے بارے میں پڑھ چکے ہیں۔ درمیانہ انحراف اور معیاری انحراف کی یکساں اکائیاں ہیں جن میں دئے ہوئے آنکڑے ہیں۔ ہم جب بھی دو سلسلئی کی یکساں درمیانہ کے ساتھ متغیر کا موازنہ کرنا چاہتے ہیں، جو کہ مختلف اکائیوں میں سے ناپا جاتا ہے، ہم صرف انتشار کی پیائش کو ہی نہیں نکالتے بلکہ ہمیں وہ پیانے درکار ہیں جن میں اکائیاں نہیں ہوتیں۔ متغیر کی وہ پیائش جو اکائیوں سے مبرّا ہے عدم مطابقت کا ضریب کہلاتا ہے (جسے cv سے

ظاہر کیا جاتا ہے)

عدم مطابقت کے ضریب کو اس طرح بیان کرتے ہیں۔

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

جہاں σ اور $\bar{x}$ آنکڑوں کے معیاری انحراف اور درمیانہ ہیں۔

دو سلسلئی کے متغیر یا انتشار کا موازنہ کرنے کے لئے، ہم ہر سلسلئی کے لئے عدم مطابقت کا ضریب (Coefficient of Variance) حل کرتے ہی۔ جس سلسلئی کا **C.V.** بڑا ہوتا ہے، کہا جاتا ہے کہ وہ دوسروں سے زیادہ تغیر پذیر ہوتی ہے۔ جس سلسلئی کا **C.V.** چھوٹا ہوتا ہے کہا جاتا ہے کہ وہ دوسروں سے زیادہ استقلال پذیر (Consistant) ہوتی ہے۔

15.6.1 مساوی درمیانہ والے دو تواتر تقسیموں کا موازنہ *(Comparison of two frequency distributions with same mean)*

مان لیجئے $\bar{x}_1$ اور σ_1 پہلی تقسیم کے درمیانہ اور معیاری انحراف ہیں، اور $\bar{x}_2$ اور σ_2 دوسری تقسیم کے درمیانہ اور معیاری انحراف ہیں۔

تب $\quad C.V.(\text{پہلی تقسیم}) = \frac{\sigma_1}{\bar{x}_1} \times 100$

اور $\quad C.V.(\text{دوسری تقسیم}) = \frac{\sigma_2}{\bar{x}_2} \times 100$

مان لیجئے $\bar{x}_1 = \bar{x}_2 = \bar{x}$ دیا ہوا ہے

اسلئے $C.V.(\text{1st distribution}) = \frac{\sigma_1}{\bar{x}} \times 100 \quad \text{(1)....}$

اور $C.V.(\text{2nd distribution}) = \frac{\sigma_2}{\bar{x}} \times 100 \quad \text{(2)....}$

(1) اور (2) سے یہ صاف ہے کہ صرف σ_1 اور σ_2 کی قدروں کی بنیاد پر دو c.v's کا موازنہ کیا جا سکتا ہے اسطرح ہم کہتے ہیں کہ دو برابر درمیانہ والی سلسلئی، جس سلسلئی کا معیاری انحراف (یا عدم مطابقت) بڑا ہے زیادہ تغیر پذیر ہے یا انتشار حاصل کئے ہوئے ہے اور ساتھ ہی جس سلسلئی کے معیاری انحراف (یا عدم مطابقت) کی قدر کم ہے دوسری سلسلئی سے زیادہ استقلال پذیر ہے۔

ہمیں مندرجہ ذیل مثالیں لینی چاہئیں:

مثال 13 ایک فیکٹری کے دو پلانٹ **A** اور **B** اپنے مزدوروں کی تعداد اور انہیں دی گئی تنخواہوں کے ذیل نتیجے دکھاتے ہیں

	A	B
مزدوروں کی تعداد	5000	6000
اوسط مہینے وار تنخواہیں	Rs 2500	Rs 2500
تنخواہوں کی تقسیم کی عدم مطابقت	81	100

کس پلانٹ میں **A** یا **B** میں انفرادی تنخواہوں میں زیادہ تغیر ہے؟

حل پلانٹ **A** میں تنخواہوں کی تقسیم میں عدم مطابقت $= \left(\sigma_1^{\ 2}\right) = 81$

اس لئے پلانٹ **A** میں تنخواہوں کی تقسیم میں معیاری انحراف $\left(\sigma_1\right) = 9$

ساتھ ہی پلانٹ **B** میں تنخواہوں کی تقسیم میں انحراف $= \left(\sigma_2\right)^2 = 100$

اسلئے پلانٹ **A** میں تنخواہوں کی تقسیم میں معیاری انحراف $\left(\sigma_2\right) = 10$

کیونکہ دونوں پلانٹ میں اوسطاً مہینہ وار تنخواہیں برابر ہیں یعنی 2500 روپیے، اس لئے، وہ پلانٹ جس میں معیاری فرق زیادہ ہے متغیر (Varibility) بھی زیادہ ہوگی۔

اس طرح، پلانٹ B انفرادی تنخواہوں کے نظریئے سے زیادہ تغیر پذیر ہے۔

مثال 14 دو تقسیم کے عدم مطابقت کے ضریب 60 اور 70 ہے اور ان کے معیاری انحراف بالترتیب 21 اور 16 ہیں۔ ان کے حسابی درمیانہ کیا ہیں؟

حل دیا ہوا ہے۔

$$\text{C.V.(پہلی تقسیم)} = 60, \sigma_1 = 21$$

$$\text{C.V.(دوسری تقسیم)} = 70, \sigma_2 = 16$$

مان لیجئے پہلی اور دوسری تقسیم کے درمیانہ بالترتیب x_1 اور x_2 ہیں تب

C.V. (پہلی تقسیم) $= \dfrac{\sigma_1}{\bar{x}_1} \times 100$

اس لیے $60 = \dfrac{21}{\bar{x}_1} \times 100$ or $\bar{x}_1 = \dfrac{21}{60} \times 100 = 35$

اور C.V. (دوسری تقسیم) $= \dfrac{\sigma_2}{\bar{x}_2} \times 100$

یعنی $70 = \dfrac{16}{\bar{x}_2} \times 100$ or $\bar{x}_2 = \dfrac{16}{70} \times 100 = 22.85$

مثال 15 گیارویں کلاس کے ایک سیکشن کے متعلقہ طلباء کی اونچائی اوروزن کے حساب کی قدریں مندرجہ ذیل ہیں۔

	اونچائی	وزن
درمیانہ (Mean)	$162.6\,cm$	52.36kg
تبدیلی (Variance)	$127.69\ cm^2$	$23.136\ kg^2$

کیا ہم کہہ سکتے ہیں کہ وزن اونچائی سے زیادہ تغیر پذیر ہیں؟

حل تغیر کا موازنہ کرنے کے لیے ہمیں ان کے عدم مطابقت کے ضریب کا حساب لگانا ہے۔

دیا ہوا ہے اونچائی کی عدم مطابقت $= 127.69 cm^2$

اس لیے اونچائی کا معیاری انحراف $= \sqrt{127.69}\text{cm} = 11.3\text{cm}$

ساتھ ہی وزن کی عدم مطابقت $= 23.1361 kg$

اس لیے وزن معیاری انحراف $= \sqrt{23.1361}\text{ kg} = 4.81\text{kg}$

اب، فرق کے ضریب (C. V.) دیئے گئے ہیں۔

(C.V.) اونچائی میں $= \dfrac{\text{معیاری انحراف} \times 100}{\text{درمیانہ}}$

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

اور (C.V.) وزن میں $= \dfrac{4.81}{52.36} \times 100 = 9.18$

صاف طور پر وزن میں C.V. بڑا ہے اونچائی میں C.V. سے۔

اس لیے، ہم کہہ سکتے ہیں کہ وزن زیادہ تغیر پذیر ہے اونچائی سے۔

مشق 15.3

1. نیچے دیئے ہوئے آنکڑوں سے دکھائیے کہ کون سا گروپ A یا B زیادہ تغیر پذیر ہے؟

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

2. نیچے دیئے گئے حصے (Shares) x اور y کی قیمتوں سے معلوم کیجیے کہ قدریں کون زیادہ بھروسے مند ہے۔

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

3. دو فرم A اور B کے مزدوروں کی ماہانہ تنخواہ کی تحلیل (توڑا گیا) کی گئی جو ایک ہی صنعت کے ہیں۔ مندرجہ ذیل نتائج دیتے ہیں۔

	فرم A	فرم B
تنخواہ پانے والوں کی تعداد	586	648
ماہانہ تنخواہوں کا درمیانہ	Rs 5253	Rs 5253
تنخواہوں کی عدم مطابقت	100	121

(i) کون سی فرم A یا B ماہانہ تنخواہوں کے لیے زیادہ پیسہ دیتی ہے؟

(ii) کون سی فرم A یا B انفرادی تنخواہوں میں زیادہ تغیر (Variability) دکھاتی ہے؟

4. ٹیم A کے ذریعے ایک فٹ بال اجلاس (Session) میں اسکور کیے گئے لوگوں کے ریکارڈ ذیل ہیں۔

اسکور کیے گئے لوگوں کی تعداد	0	1	2	3	4
میچوں کی تعداد	1	9	7	5	3

ٹیم B کے لیے، ایک میچ میں اسکور کیئے گئے گولوں کا درمیانہ 2 ،اور معیاری انحراف 1.2 گول تھا۔ معلوم کیجیے کہ کون سی ٹیم کو زیادہ قابلِ اعتبار سمجھا جانا چاہیے؟

.5 پودوں کی پیداوار میں 50 پودوں لمبائی x (سینٹی میٹر میں) اور وزن y (گرام میں) کے مجموعے اور مربعوں کے مجموعے کے بالترتیب نیچے دیئے گئے ہیں:

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

کون زیادہ تغیر پذیر ہے، لمبائی یا وزن میں؟

متفرق مثالیں

مثال 16 20 مشاہدات کی عدم مطابقت 5 ہے۔ اگر ہر مشاہدے کو 2 سے ضرب کیا گیا ہو ، نتیجتاً مشاہدات کی نئی عدم مطابقت (Variance) معلوم کیجیے۔

حل مان لیجیے $x_1, x_2, \ldots, x_{20}$ مشاہدے ہیں اور $\bar{x}$ انکا درمیانہ ہے۔ دیا ہوا ہے عدم مطابقت=5 اور n = 20۔ ہم جانتے ہیں کہ

$$var\,iance\left(\sigma^2\right) = \frac{1}{n}\sum_{i=1}^{20}(x_i - \bar{x})^2 \quad e.i \quad 5 = \frac{1}{20}\sum_{i=1}^{20}(x_i - \bar{x})^2$$

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

اگر ہر مشاہدے کو 2 سے ضرب کیا گیا ہے اور نئے نتیجتاً مشاہدے y_i ہیں، تب

$$y_i = 2x_i \quad \text{یعنی} \quad x_i = \frac{1}{2}\ y_i$$

اس لیے $$\bar{y} = \frac{1}{n}\sum_{i=1}^{20} y_i = \frac{1}{20}\sum_{i=1}^{20} 2x_i = 2 \cdot \frac{1}{20}\sum_{i=1}^{20} x_i$$

یعنی $$\bar{x} = \frac{1}{2}\bar{y} \quad \text{یا} \quad \bar{y}2\bar{x}$$

x_i اور $\bar{x}$ کی قدریں (1) میں رکھنے پر ہمیں حاصل ہوتا ہے۔

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

اس لیے نئے مشاہدے کی تبدیلی $=\frac{1}{20}\times 400=20=2^2\times 5$

نوٹ پڑھنے والے کو یہ نوٹ کر لینا چاہیے کہ اگر ہر ایک مشاہدے کو ایک مستقل 'k' سے ضرب کیا جائے نتیجتاً مشاہدوں کی عدم مطابقت (Variance) اصل تبدیلی کا k^2 مرتبہ ہو جائے گا۔

مثال 17 5 مشاہدوں کا درمیانہ 4.4 ہے اور ان کی عدم مطابقت 8.24 ہے۔ اگر تین مشاہدے 1، 2 اور 6 ہیں، دوسرے دو مشاہدے معلوم کیجیے۔

حل مان لیجیے دو مشاہدے x اور y ہیں۔

اس لیے سلسلہ 1, 2, 6, x, y ہے۔

اب درمیانہ $\overline{x}=4.4=\frac{1+2+6+x+y}{5}$

یا $22=9+x+y$

اس لیے $x+y=13$ ......(1)

ساتھ ہی تبدیلی $=8.24=\frac{1}{n}\sum_{i=1}^{5}(x_i-\overline{x})^2$

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

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

اس لیے $x^2+y^2=97$ ......(2)

لیکن (1) سے، ہمارے پاس ہے۔

$x^2+y^2+2xy=169$ ......(3)

(2)اور(3)سے، ہمارے پاس ہے۔

$$2xy = 72 \qquad (4).....$$

(4)کو(2)سے تفریق کرنے پر، ہمیں حاصل ہوتا ہے۔

$$x^2 + y^2 - 2xy = 97 - 72 \quad \text{یعنی} \quad (x-y)^2 = 25 \qquad (5).....$$

یا $x - y = \pm 5$

اس لیے، (1)اور(5)سے ہمیں حاصل ہوتا ہے۔

جبکہ $x - y = 5$ $\quad x = 9, \quad y = 4$

جبکہ $x - y = -5$ $\quad x = 4, \quad y = 9$ یا

اس لیے بقیہ مشاہدے 4اور 9 ہیں۔

مثال 18 اگر ہر ایک مشاہدے $x_1, x_2, ..., x_n$ کو 'a' سے بڑھایا جائے جہاں 'a' منفی یا مثبت عدد ہے، دکھائیے کہ عدم مطابقت(Variance) بغیر تبدیلی ہوئے رہ جاتی ہے۔

حل مان لیجئے $x_1, x_2, ..., x_n$ کا درمیانہ $\bar{x}$ ہے۔ تب عدم مطابقت(Variance) اسی طرح دی گئی ہے۔

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

اگر ہر ایک مشاہدے میں 'a' جمع کیا جائے، نئے مشاہدے ہوں گے۔

$$y_i = x_i + a$$

مان لیجئے نئے مشاہدوں کا درمیانہ $\bar{y}$ ہے۔ تب

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n}(x_i + a)$$

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

یعنی $\bar{y} = \bar{x} + a$

اس لیے نئے مشاہدوں کی عدم مطابقت (Variance)

$$\sigma_2^{\,2} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i + a - \overline{x} - a)^2$$ [(1) اور (2) کا استعمال کرنے پر]

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

اس طرح، نئے مشاہدوں کی تبدیلی اصل مشاہدوں کی عدم مطابقت جیسی ہے۔

> **نوٹ** ہم یہ نوٹ کر سکتے ہیں جمع کرنا (یا گھٹانا) ایک مثبت عدد کا (یا سے) ایک گروپ کے ہر مشاہدے میں تبدیلی بے اثر نہیں ڈالتا۔

مثال 19 100 مشاہدات کا درمیانہ اور معیاری انحراف حساب لگانے پر بالترتیب 40 اور 5.1 ہے ایک طالب علم کے ذریعے جس نے غلطی سے ایک مشاہدہ 40 کی بجائے 50 لے لیا، بتایئے صحیح درمیانہ اور معیاری انحراف کیا ہیں؟

حل دیا ہوا ہے کہ مشاہدوں کی تعداد (n) = 100

غلط درمیانہ $\overline{x} = 40$

غلط معیاری انحراف $(\sigma) = 5.1$

ہم جانتے ہیں کہ $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$40 = \frac{1}{100}\sum_{i=1}^{100} x_i$ یا $\sum_{i=1}^{100} x_i = 4000$

یعنی مشاہدوں کا غلط جوڑ $= 4000$

اس طرح مشاہدوں کا صحیح مجموعہ = غلط مجموعہ $- 50 + 40$

$4000 - 50 + 40 = 3990$

اس لیے، صحیح درمیانہ = صحیح جوڑ = $\frac{3990}{100} = 39.9$

ساتھ ہی $\sigma=\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^{\,2}-\frac{1}{n^2}\left(\sum_{i=1}^{n}x_i\right)^2}$ انحراف(S.D)

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

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

یا $26.01=\frac{1}{100}\times \text{Incorrect}\sum_{i=1}^{n}x_i^{\,2}-1600$

اس لیے $incorrect\sum_{i=1}^{n}x_i^{\,2}=100(25.01)=162601$

اب $correct\sum_{i=1}^{n}x_i^{\,2}=incorrect\sum_{i=1}^{n}x_i^{\,2}-(50)^2+(40)^2$

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

اس لیے صحیح معیاری انحراف

$$=\sqrt{\frac{\text{Correct}\ \ \sum x_i^{\,2}}{n}-(\text{Correct mean})^2}$$

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

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

باب 15 پر متفرق مشق

.1 آٹھ مشاہدات کا درمیانہ اور عدم مطابقت بالترتیب 9اور 9.25 ہیں۔ اگر چھ مشاہدے 6، 7، 8، 10، 12، 12اور 13 ہیں، باقی دو مشاہدے معلوم کیجیے۔

.2 7مشاہدات کا درمیانہ اور عدم مطابقت بالترتیب 8اور 16ہیں۔ اگر پانچ مشاہدے 2، 4، 10، 12اور 14ہیں۔ باقی دو مشاہدات معلوم کیجیے۔

.3 چھ مشاہدات کا درمیانہ اور معیاری انحراف بالترتیب 18 اور 4 ہیں اگر ہر مشاہدے کو 3 سے ضرب کیا جائے، نتیجتاً مشاہدات کا نیا درمیانہ اور معیاری انحراف معلوم کیجئے۔

.4 n مشاہدات کا درمیانہ $\bar{x}$ اور عدم مطابقت (σ^2) دی ہوئی ہے۔ ثابت کیجئے کہ مشاہدات $ax_1, ax_2, ax_3, ..., ax_n$ کا درمیانہ اور عدم مطابقت بالترتیب $a\bar{x}$ اور $a^2\sigma^2$ ہیں $(a \neq 0)$

.5 20 مشاہدوں کا درمیانہ اور معیاری انحراف بالترتیب 10 اور 2 حاصل ہوا ہے۔ دوبارہ جانچ کرنے پر، یہ حاصل ہوا کہ ایک مشاہدہ 8 غلط تھا۔ ذیل ہر ایک کیس میں صحیح درمیانہ اور معیاری انحراف کا حساب لگایئے۔

(i) اگر غلط چیز نکال دی جائے (ii) اگر اسے 12 سے بدل دیا جائے۔

.6 تین مضمون ریاضی، فزکس اور کیمسٹری میں ایک کلاس میں 50 طلباء کے ذریعے حاصل شدہ نمبروں کے درمیانہ اور معیاری فرق نتیجے دیئے گئے ہیں۔

مضمون	ریاضی	فزکس	کیمسٹری
درمیانہ	42	32	40.9
معیاری فرق	20	15	20

تینوں مضمونوں میں کون سا مضمون نمبروں میں سب سے زیادہ تغیر پذیر ہے اور کون سب سے کم تغیر پذیر ہے؟

.7 100 مشاہدات کا درمیانہ اور معیاری انحراف بالترتیب 20 اور 3 ہے۔ بعد میں یہ پایا گیا کہ تین مشاہدے غلط ہیں، جو کہ 21، 21 اور 18 پائے گئے۔ اگر غلط مشاہدات کو نکال دیا جائے تو درمیانہ اور معیاری فرق معلوم کیجئے۔

خلاصہ (Summary)

- انتشار کے پیمانے، وسعت، چوتھائی انحراف، درمیانہ انحراف، عدم مطابقت، معیاری انحراف انتشار کے پیمانہ ہیں۔

وسعت = کم سے کم قدر – زیادہ سے زیادہ قدرت

- غیر جماعتی انکڑوں کے لیے درمیانہ فرق

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

- جماعتی آنکڑوں کے لیے درمیانہ فرق

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

- غیر جماعتی آنکڑوں کے لیے عدم مطابقت اور معیاری انحراف

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

- غیر مسلسل تواتر تقسیم کے لیے عدم مطابقت اور معیاری انحراف

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

- مسلسل تواتر تقسیم کے لیے عدم مطابقت اور معیاری انحراف

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

- عدم مطابقت اور معیاری انحراف معلوم کرنے کے لیے مختصر طریقہ

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

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

عدم مطابقت کا ضریب $(\textbf{C.V.}) = \dfrac{\sigma}{\bar{x}} \times 100, \ \bar{x} \neq 0.$

- برابر درمیانے والی سلسلئے کے لیے، وہ سلسلئے جس کا معیاری انحراف کم ہے زیادہ دیرپا ہے یا کم پھیلی ہوئی ہے۔

تاریخ کے اوراق سے

شماریات لیتن (Latin) لفظ درجہ (Status) سے نکالا گیا ہے جس کا مطلب سیاسی صوبہ ہے۔ یہ اشارہ دیتا ہے کہ شماریات اتنا ہی پرانا ہے جتنی انسانی تہذیب۔ سال 3050B.C میں شاید پہلی مردم شماری مصر میں ہوئی تھی۔ ساتھ ہی ہندوستان میں تقریب 2000 سال پہلے ہمارے پاس ایک انتظامیہ شماریات جمع کرنے کا اثر انداز طریقہ تھا، خاص طور پر، چندر گپت موریا (324-300 B.C) کے دورِ حکومت میں۔ کوتلیا ارتھ شاسترا (Kautilya's Arthshastra) کے زمانے

میں پیدائش اور موت کے آنکڑوں کو جمع کرنے کے طریقہ کا ذکر ملتا ہے۔ اکبر کے دورِ حکومت میں انتظامیہ سروے کا مفصل حساب و کتاب آئینہ اکبری میں دیا گیا ہے جو ابوالفضل نے لکھی ہے۔

برطانیہ کے کیپٹن جان گراؤنٹ (Captain John Graunt) کو شماریات حیات کا بانی کہا جاتا ہے۔ اس کی پیدائش اور موت شماریات کی جانکاری کے لیے۔ جے کب برنولی (Jacob Bernoulli) (1654-1705) نے اپنی کتاب ''اورس کنجیک ٹانڈی'' میں۔ زیادہ اعداد کے قانون کو بیان کیا ہے جو کہ 1713 میں شائع ہوئی تھی۔

شماریات کا نظری بڑھاؤ سترہویں صدی نصف میں ہوا، اور اس کے بعد کھیلوں کے نظری تعارف۔ اور واقع (یعنی اتفاق) کے۔ فرانسیسی گیلٹن (1822-1921) ایک انگریز آدمی نے شماریاتی طریقوں کو بائیو میٹری (Biometry) کے میدان میں استعمال کیا۔ کرل پیرس (1857-1936) نے شماریاتی مطالعہ کو آگے بڑھانے میں بہت زیادہ مدد کی اس نے چی مربع ٹیسٹ (Chi square test) کی ایجاد کی اور انگلینڈ میں (1911) میں شماریاتی تجربہ گاہ کی سنگِ بنیاد رکھی۔

سر رونالڈ لسے (1890-1962) جسے ہم جدید شماریات کا بانی کہتے ہیں نے اسے بہت سے مختلف دوسرے میدانوں میں استعمال کیا جیسا کہ تخلیقات (Genetics)، بائیومیٹری (Biometry)، تعلیم (Education)، کاشتکاری (Agriculture) وغیرہ وغیرہ ہیں۔