

Statistics

* statistics পৰি কোনো Data কি?

⇒ Data consists of information coming from observations counts, measurements or responses.

⇒ এই বলা কল: অধিগ্রহণে আবা যাওয়া ঘোষণার ৫০% টি তিনি পথচারী।
তাইনে এটা কুল Data বা observation থেকে পাওয়া পেল।

* What is statistics?

⇒ The science of collecting, organizing, analyzing and interpreting data in order to make decisions.

[কোনো বিষয়ে যে data শুলো collect কৰা কল তা আমে আজানো ক্ষয় (organizing), এবং পর analyzing (mathematical analysis) কৰা ক্ষয় ও অন্তিমে এব বিষেচনা কৰে একটি decision (আজা ক্ষয়)]

Data Sets

i) Population : The collection of all outcomes, responses, measurements or counts that are of interest.

ii) Sample : A subset of the population.

* একটি example দিয়ে clearly সুব্ধব।

৪৭. বাংলাদেশে recently এক survey তে ২৭০৮ জন Adult কে—
জিজ্ঞাসা করা কল : “আবা কি ঘৰে কহে global warming is a problem
that requires immediate government action. তো আদেৱ ঘৰ্ণে
নং৩৯ জন কলল হ্যাঁ।”

তো এখানে population কিৰি sample কিৰি?

এখন চিন্তা কৰ, Bangladesh o adult আছে ৫০০০০ টো। এইটো লাখটো
নিম্যে পঢ়িত কল—population. আৱ survey কৰা বয়েছে ঘৰ্ণ ২৭০৮ জনৰ
ঘৰ্ণে। এই ২৭০৮ জনৰ জন কল কে ৫ লাখেৱ ঘৰ্ণে কৃত্তু এক অংশ। অতএব,
২৭০৮ জন কল sample এয় example.

Parameter: A number that describes a population
characteristics.

[এই বাংলাদেশৰ ২৭ হাজিৰি ঘৰ্ণুষ কল population এৰ অংশ। এন্ধন বলা
কল বাংলাদেশৰ ঘৰ্ণুষেৰ পত্ৰ বয়স কল ৫০ বছৰ। এইটো কল population
এৰ একটী characteristic বুবাছে। \therefore average age কল parameter]

Statistic: A number that describes a sample characteristic.

[পুলো বাংলাদেশৰ ঘৰ্ণুষ মদি একটী population এৰ অংশ। ক্ষয় তাহলে
বাংলাদেশৰ ৪ ৮ বিজাগেৰ ঘৰ্ণুষ কল sample এৰ অংশ। এখন যদি বলা ক্ষয়
বাংলাদেশৰ ৮ বিজাগেৰ ঘৰ্ণুষেৰ average age . তাহলে এটা কল sample
এৰ characteristic \therefore statistic.]

Branches of Statistics

- i) **Descriptive statistics:** Involves organizing, summarizing and displaying data e.g. table, charts
 [କୌଣସି ଡାଟା ଏଇ କୋଣସି କାହାରେ କାହାରେ କାହାରେ କାହାରେ କାହାରେ]
- ii) **Inferential statistics:** Involves using sample data to draw conclusion about a population.

[ଯେଉଁ ବାଦିଲାଗେର ଏହି କୋଣସି ଶାଖା କିମ୍ବା ଫଳ ଏକାଟି ବାଦିଲାଗେର ଏହି କିମ୍ବା ଫଳ ଏକାଟି ପରିଚୟ କରିବାରେ ଆଜାଦେଇ ଅଛି ଏହି କିମ୍ବା ଫଳ ଏକାଟି ପରିଚୟ କରିବାରେ ଆଜାଦେଇ ଅଛି]

Types of Data

Quantitative

↓
 Data can be measured
 with numbers such as
 duration or speed (ମେଧ୍ୟବାକେ)

Discrete Data

(ପ୍ରକାଶ୍ୟ, ଡିମ୍ବାଙ୍ଗରୁଣେ)

continuous data

(କ୍ଷେତ୍ରଫଳୀକାରୀଯାୟ)

Interval

Ratio

Qualitative

↓
 Non-numerical data
 that is categorical,
 such as yes/no responses
 or eye colors (ଚାନ୍ଦବାକେ)

Nominal

(କୋଣାରୀ variable
 ଏବଂ ନାମକାରନୀ ବ୍ୟବହାର କରୁଥିଲୁବାକେ)
 Ordinal
 (order of
 values
 describable
 କରୁଥିଲୁବାକେ)

Sampling Techniques

Simple Random Sample:

Every possible sample of the same size has the same chance of being selected.

[same size বা same feature এর sample মূল্যে select হওয়ার chance 'বেশি' থাকে]

Frequency Distribution

*construction of Frequency Distribution

আজয়া একটি example দিয়ে এব তু বুঝঃ

make a frequency distribution table of this example:

The following sample data set lists the number of minutes 50 Internet subscribers spent on the internet during their most recent session. construct a frequency distribution that has seven classes:

50	40	41	18	11	7	22	44	28	21	19	23	37	51	54	42	86
41	78	56	72	56	17	7	69	30	80	56	29	33	46	31	39	20
18	29	34	59	73	22	36	39	30	62	54	67	39	31	53	44	

Class	Tally	Frequency, f
7-18		6
19-30		10
31-42		13
43-54		8
55-66		5
67-78		6
79-90		2
		$\therefore \sum f = 50$

class : দুটা এখানে class অংশ্যা বল ২. কারণ ques এই বলে দিচ্ছে রাস্তিক্রমে
তবে generally class অংশ্যা 5-20 এর ছাত্রে রয়।

এখন দেখব class কিরাবে তিথি করব। এখন তা করার জন্য আগে range জানালাগতো

$$\text{range} = \text{বড়চেয়ে বড় অংশ্যা} - \text{সবচেয়ে ছোট অংশ্যা}$$

$$= 86 - ?$$

Example গুলি

number মুলি যেকোনো

ক্ষেত্র

$$= ?$$

আজ্ঞা যেটা পাও তার
পরের পূর্ণ অংশ্যা নিয়ে

আব, class width কল:

$$\frac{\text{range}}{\# \text{classes}} = \frac{?}{?} \approx 11.29 = 12$$

ques পরিদেয়া রাস্তাক্রমে

এখন দেখ, class এর ~~ফোর্ম~~ column কি complete করব কিভাবে?

class ও এর প্রতিটি গোয়েটি form কচ্ছে গুজাবে দেখ \rightarrow

lower limit — upper limit

এখন lower limit ও upper limit কিভাবে পাব তা দেখ.

lower limit

Upper limit

এখানে প্রথমে x Given ques
এর lowest number
 $x + 12 = 19$

$$19 + 12 = 31$$

$$31 + 12 = 43$$

$$43 + 12 = 55$$

$$55 + 12 = 67$$

$$67 + 12 = 79$$

এখানে upper class limit কুইজাবে নিতে পায়।

lower limit

upper limit

x

19

19

31

31

প্রাপ্ত পুরোটি

43

2nd lower limit আজাবে

55

একে বলে
continuous.

67

79

ଆଶ୍ରମୀ ଉପାୟঃ

lower limit	upper limit
୨	$\begin{array}{r} 18 \\ +12 \\ \hline 30 \end{array}$
୧୯	
୩୧	$42 \rightarrow 30+12$
୫୩	୫୪
୫୫	୬୬
୬୪	୭୪
୮୭	୯୦ ଏବଂ ବଳେ discontinuous

* ସମ୍ଭାବ୍ୟ ଯେଉଁ ଏକଟି ଫର୍ମାଟ୍ କରିବାକୁ ଆବଶ୍ୟକ କରିବାକୁ ହାତେ କରିବାକୁ ପରିବର୍ତ୍ତନ କରିବାକୁ ଆବଶ୍ୟକ କରିବାକୁ ହାତେ କରିବାକୁ

ସମ୍ଭାବ୍ୟ ଯେଉଁ ଏକଟି data ୧୮.୫ ଆଛେ ତାକେ ତୁ କିମ୍ବା discontinuous କିମ୍ବା ବାନାନୋ ବାନନ୍ତି କିମ୍ବା ଏକଟି କିମ୍ବା ବାନନ୍ତି କିମ୍ବା ବାନନ୍ତି କିମ୍ବା ବାନନ୍ତି କିମ୍ବା

କିମ୍ବା ଏକଟି data ୫୦. ବାନନ୍ତି କିମ୍ବା ବାନନ୍ତି କିମ୍ବା ବାନନ୍ତି

କିମ୍ବା ଏକଟି data ୫୩-୫୫ ଏକଟି class କିମ୍ବା କିମ୍ବା କିମ୍ବା କିମ୍ବା କିମ୍ବା

* frequency ଏବଂ ଘରେ just ଆଶ୍ରମୀ tally ଏବଂ ଘରେ କରିବାକୁ କାମ କରିବାକୁ କାମ କରିବାକୁ

Midpoint of a class :

$$\frac{\text{lower class limit} + \text{upper class limit}}{2}$$

class	Midpoint	frequency
7 - 18	$\frac{7+18}{2} = 12.5$	6
19 - 30	$\frac{19+30}{2} = 24.5$	10
31 - 42	$\frac{31+42}{2} = 36.5$	13

relative frequency :

$$\text{relative frequency} = \frac{\text{class freq.}}{\text{sample size}}$$

যানে ques. বা
কর্তজন্ম / কর্তৃ
data দেওয়া

class	frequency	relative frequency
7 - 18	6	$\frac{6}{50} = 0.12$
19 - 30	10	$\frac{10}{50} = 0.20$
31 - 42	13	$\frac{13}{50} = 0.26$

cumulative frequency of a class:

=The sum of the frequency for that class and all previous classes.

<u>class</u>	<u>Frequency</u>	<u>cumulative freen</u>
7-18	6	6
19-30	10	16 [10+6]
31-42	13	29 ∴ 6+10+13=29

অন্তর্বর্ণ Expanded frequency distribution: (গোল গবেষণা একা table(4))

Class	Frequency, f	Midpoint	Relative Frequency f/n	cumulative frequency
7-18	6	12.5	0.12	6
19-30	10	24.5	0.20	16
31-42	13	36.5	0.26	29
43-54	8	48.5	0.16	37
55-66	5	60.5	0.10	42
67-78	6	72.5	0.12	48
79-90	2	84.5	0.04	50

$$\sum f = 50$$

$$\sum \frac{f}{n} = 1.$$

এখনে, relative frequency টি বোবায় কেনো class এর ছাত্রে
কেনো একটি data এর পজিশন গুলো কতটুকু পথে আবও দেখব।

* relative frequency এর অবস্থানের যোগ্যতা হবে ।

Graphs of Freq. Distribution

Graph অংকণ করে আবাসের 1st কাড কল check কৰা যে,

আবাসের class মুলো continuous না কি discontinuous?

মদি continuous ক্ষয় তাঙ্কে অবস্থা নাই। কিন্তু মদি discontinuous ক্ষয় তাঙ্কে আবাসের আবেক্ষণ্ণ কাড করতে কৰে। তা কল class boundary কৰে কৰা। কিভাবে কৰে কৰব?

class Boundary কৰে কৰা:

মুক্ত প্রক্রম class : ২-১৮ এর class boundary কি কৰে?

Lower limit থেকে ০.৫ যিলোগ কৰবা

* Upper limit থেকে ০.৫ যোগ কৰবা

তাঙ্কে class : ২-১৮

$$\therefore \text{class boundary} : [2-0.5] - [18+0.5]$$

$$= 6.5 - 18.5 \text{ এটাই কল class boundary}$$

২-১৮ class এব

ତାରିଖ

class

class boundaries

୮-୧୮

୬.୫ - ୧୮.୫

୧୯-୩୦

୧୮.୫ - ୩୦.୫

୩୧-୪୨

୩୦.୫ - ୪୨.୫

ଦ୍ୱାରା class boundary

କେବଳ କିମ୍ବା class ଶୁଳ୍କ

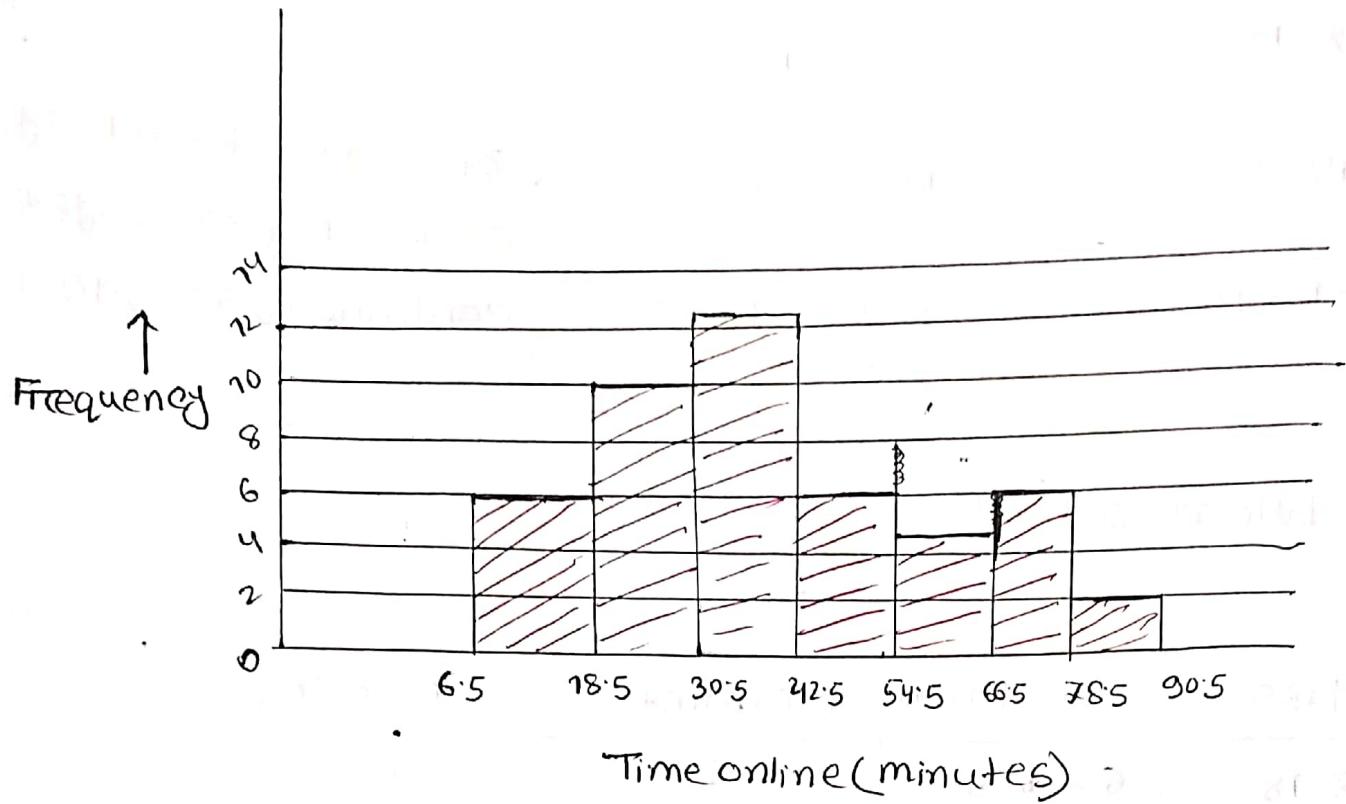
continuous କହା ଗେଛ ।

ଆଜିଲେ ପୂର୍ବୋ table ହେଉଥିବା ଏବଂ

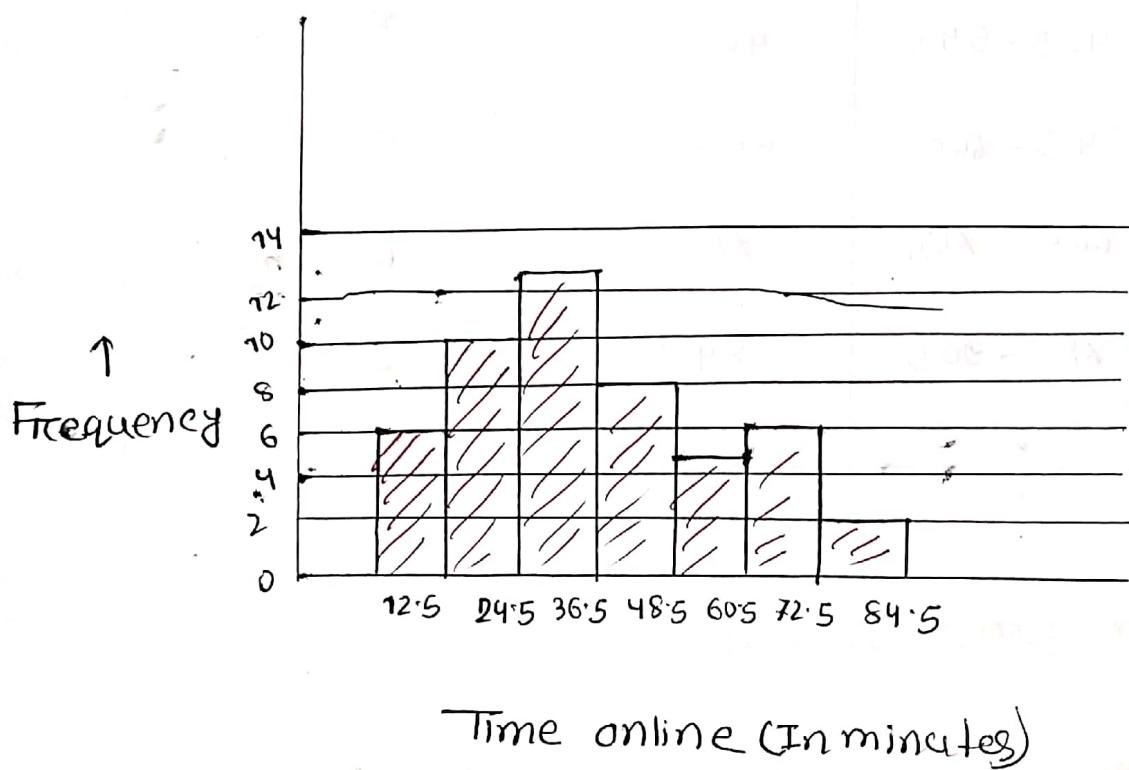
class	class Boundaries	MidPoint	frequency
୮-୧୮	୬.୫-୧୮.୫	୧୨.୫	୬
୧୯-୩୦	୧୮.୫-୩୦.୫	୨୪.୫	୧୦
୩୧-୪୨	୩୦.୫-୪୨.୫	୩୬.୫	୧୩
୪୩-୫୪	୪୨.୫-୫୪.୫	୪୮.୫	୮
୫୫-୬୬	୫୪.୫-୬୬.୫	୬୦.୫	୫
୬୭-୭୮	୬୬.୫-୭୮.୫	୭୨.୫	୬
୮୯-୯୦	୭୮.୫-୯୦.୫	୮୪.୫	୨

Frequency Histogram ହୁଅଇବେ ଆଜା ଯାଏ :

ଏହା ପଦ୍ଧତି ମଧ୍ୟରେ Page ୭:



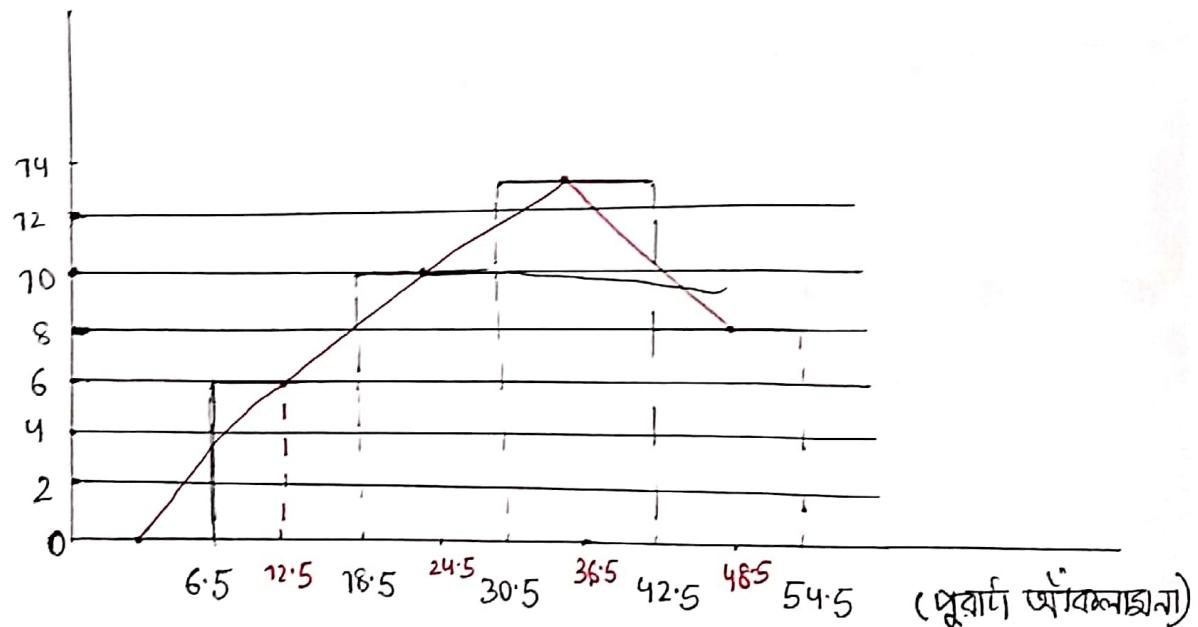
class boundary use করে— অঁকা কল এখন, আব 2nd pic'র midpoint
পুনৰ উল্লেখ করা কল graph adjust. আব কিন্তু না।



ଏହା graph ଦେଖେ କୁମା ଯାଇଛି, maximum subscribers \Rightarrow 19 ମେଟ୍ରୋସ୍ୟ
minute internet ଏ ଅଛ୍ୟ କାହିଁମେହିତେ ।

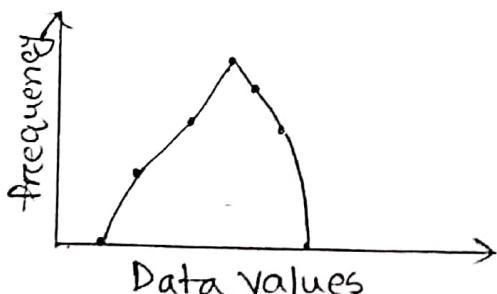
Frequency Polygon

ଆମେ ପିଂକେ ନିର୍ବିଧି, then କୁମାଟିଲି ।



ଦେଖ ଫର୍ମ୍‌ସ୍ୱର୍ଗକ class ଏବଂ midpoint ଏ ଆଶ୍ରାମକାରୀ ବର୍ତ୍ତେ dot ଦିଲାଇ ।

Then ତେମୁଣ୍ଡାକେ ହିଂସ ବର୍ତ୍ତେ ଦିଲେହି ଆଶ୍ରାମ ମେହେ ଯାଏ Frequency polygon. ତାହାରେ frequency polygon ଫାର୍ଜାଲ ଏହନୋ



Arithmetic Mean (Mean)

मत

* घोटा वा कम्प्युल्टर sample एवं मतके express करा रख \bar{x} (X bar दिय)

* यास कोना population एवं मतके express करा रख M (mu द्वारा)

Grouped data : यासा एकी example देखिलाए ५० टी अंदे देत्या आडे जोके यासा आजिये frequency division table बानियेहिलाए। एवं आजानोर भवे नास रन grouped data.

ungrouped data : table ए आजानोर आमे ५० टी Data येष्वन देत्या दिल ग्रुप्स ए. जोकी रन ungrouped data ए example]

Mean for ungrouped Data :

$$\text{Mean} = \frac{\text{sum of all values}}{\text{Number of values}}$$

sample mean : $\bar{x} = \frac{\sum x}{n}$ → यहे कर्य प्रयोगादेशेर आवृयके निये प्रकटी population आव आव उपिंगमेय आवृय कर्म कोना sample एर अंश । n कर्म sample संस्कृत

Population mean, $\mu = \frac{\sum x}{N} \rightarrow$ population

Mean for grouped data:

Sample mean; $\bar{x} = \frac{\sum fx}{n}$

Population mean; $\mu = \frac{\sum fx}{N}$

Question solving

question এয়দি আল্যা করবে population mean এর কথা না বলে

তাহলে আব্যাস sample mean র কৈবল্য করব।

example 1:

The daily rates of a sample of eight employees

at GCMS Inc. are 550, 420, 560, 500, 700, 670, 860,

480. Find the mean daily rate of employee.

Ans: কলাই আছে sample:

$$\therefore \text{Sample mean} = \frac{\sum x}{n} = \frac{550 + 420 + 560 + 500 + 700 + 670 + 860 + 480}{8}$$

$$= 592.50 \text{ (Ans)}$$

Example 2: Find the population mean of the ages of 9 middle management employees of a certain company.

The ages are: 53, 45, 59, 48, 54, 46, 51, 58 and 55

Solution:

কলারি প্যাঠে Population mean:

$$\therefore \bar{A} = \frac{\sum x}{N} = \frac{53+45+59+48+54+46+51+58+55}{9} \\ = 52.71(\text{Ans})$$

Example 3:

class limits	Frequency
18-26	3
27-35	5
36-44	9
45-53	14
54-62	11
63-71	6
72-80	2

Using the example provided in on SJS Travel Agency.

Determine the mean of the frequency distribution on the ages of 50 people taking travel tours.

Solution:

Class limits	f	x	fx
18-26	3	$\frac{18+26}{2} = 22$	66
27-35	5	31	155
36-44	9	40	360
45-53	14	49	686
54-62	11	58	638
63-71	6	67	402
72-80	2	76	152
Total	50		2459

$$\therefore \bar{x} = \frac{\sum fx}{n} = \frac{2459}{50} = 49.18 \text{ (Ans)}$$

কিছু অন্যান্য

ques প্র.তা

sample mean

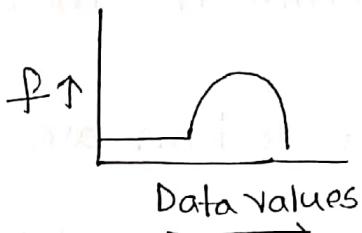
বিত্ত এর Arithmetic mean always proper measure না।

আচর্যা দেশ্বে জ্ঞান্য বেশ্যায় Arithmetic mean proper measure না।

When Arithmetic mean should not be used.

i) In Highly skewed distribution.

দেখ কিন্তু graph কতে পারে এবন



ক্ষুবুষ দিকে frequency

এবন কষা। যাবাব পৰে শিমে অনেক চেকি। তো এবন ক্ষেত্ৰে যা $\frac{f}{\text{Data}}$

এবন ক্ষেত্ৰে Arithmetic mean use কৰা যাবেনা।

ii) When the distribution is unevenly spread. concentration

being small or large at irregular points.

Example 3 এর data শুলো নিয়ে আগব্যা যদি graph আঁকি তাৰে

graph কৰে এবন f

যান evenely spreaded একটি graph

(মোখাও ঝুঁটি কৰে drop কৰিবাই) কিন্তু যদি তেন্তো graph কৰ তা

f

এবন. যান ঝুঁটি কৰে rise/fall, তথ্য বলা যাব তা

unevenly spreaded graph. এবন ক্ষেত্ৰে arithmetic mean use

কৰা যাবেনা।

iii) When an average rate of growth or change over a period of time is required.

[ঘনে কর, speed এর graph, যা time এর জাতে আবেচে change
হয়। এখন ক্ষেত্রে arithmetic mean বড় করা যাবেন।]

iv) When the observation are from geometric progression

[Geometric mean এর example দেখুব।]

v) When there are very large and very small values of observations.

[মদি data মূলো অনেক বড় বা ছোট বয়েজে আসে বলে।]

এই ক্ষেত্রে আবাও বিন্দু mean এর সাধিক্ষণ করেছে তা আবাও দেখো।

Weighted Mean

আবাদের Ruet এ CGPA এর বিজ্ঞাপ কিভাবে কর কল তো। গত কর
১টি ৩ credit এর ক্ষেত্রে আছে। তাহলে কি করা বস্তু, আবাদের
১টি subject এ প্রতি CGPA কে তাদের credit দ্বারা গুণ করে—
আগে sum করা হয় ও then total credit দিয়ে ভাগ করা
হয়। weighted mean ও same. আবাও ৩টি example দেখে
clear হবে।

Example 1: A class is graded based on weighted mean as follows:

Homework	: 20%	পঞ্চলো আঞ্চাদের example এর credit এর হাতা পঞ্চলোর নাই কল weight. তাই Homework কত mark বা এর ছুল্প (weight) কো
Quizzes	: 35%	
Tests	: 45%	

Let's say scores are 95 on homework, 82 on quizzes, and 79 on tests. What is the weighted mean?

$$\bar{x} = \frac{\sum (x \cdot w)}{\sum w}$$

দেখ Homework এ মেয়েছে 95, এটি x_1 আর Homework এর weight 20% এর w_1 . তাইলে $(x_1 \cdot w_1) = (0.2 \times 95)$ এভাবে সব পঞ্চলো আগ দিব।

$$\therefore \bar{x} = \frac{(95 \times 0.2) + (82 \times 0.35) + (79 \times 0.45)}$$

$$1 \rightarrow (20\% + 35\% + 45\%) = 100\% \therefore 100\% শান্ত এলিএট$$

$$= 83.5$$

এখন যান কব, math এ tests নাই option নাই কোনো। তখন,

$$\bar{x} = \frac{\sum (x \cdot w)}{\sum w} = \frac{95(0.2) + 82(0.35)}{0.55} = 86.7$$

Math 2: Rachel is taking a class in which her grade is determined as follows:

50% from her test mean

15% from her midterm

20% from her final

15% from her homework.

Her scores are 86 (tests), 96 (midterm) and 100 (homework). What does Rachel need to get on her final exams to receive a 90% in class?

$$\text{目标 } \bar{x} = 90.$$

$$\therefore \bar{x} = \frac{\sum(x \cdot w)}{\sum w} = \frac{86(.5) + 96(.15) + 100(.15) + n(.2)}{100\% = 1} = 90$$

↖ & final \bar{x}
↓
60%

$$70.9 + .2n = 90$$

$$\therefore .2n = 19.1$$

$$\therefore n = 95.5 \text{ (Ans)}$$

\therefore Rachel needs to get a 95.5% on the final to earn a 90% in the class.

Example 3: For the month of April, a checking account has a balance of \$523 for 24 days, \$2415 for 2 days and \$250 for 4 days. What is the account's mean daily balance for April?

$$= 523 \times \left(\frac{24}{30}\right) + 2415 \times \left(\frac{2}{30}\right) + 250 \times \left(\frac{4}{30}\right)$$

\nearrow
April মাস 30 দিনে

$$= 612.73 \text{ (Ans)}$$

Geometric Mean

মদি data মুলো শক্তিবা কাবে প্রকাশ করা যাবে তোহানে geometric mean সূত্রাব ক্ষেত্রে পারব।

$$G.M. = \sqrt[n]{(x_1)(x_2)(x_3)\dots(x_n)}$$

Geometric mean ক্ষুর্দ্ধাণ্য ungrouped data এর জন্য formula.

এব. তোহার একটি company আছে। যার নাম 'মাহিন বই শব'। তো তুঁরি জনতে চাও তোহার company'র একটি নির্দিষ্ট period এ কি পরিমাণ sale হচ্ছে বা production কোভে তা বেব করার formula কল;

$$G.M. = \sqrt[n-1]{\frac{\text{value at the end of the period}}{\text{value at the start of the period}}} \rightarrow$$

Example 1: Suppose the profits earned by the MSS construction company on five projects were 5, 6, 4, 8, and 10 Percent, respectively. What is the geometric mean profit?

Solution:

$$G.M. = \sqrt[5]{(x_1)(x_2)(x_3)(x_4)(x_5)}$$

$$= \sqrt[5]{(5)(6)(4)(8)(10)}$$

$$= \sqrt[5]{9600}$$

$$= 6.26 \text{ (Ans)}$$

Example 2: Badminton as a sport grew rapidly in 2008. From January to December 2008 the number of Badminton clubs in Metro Manila increased from 20 to 155. Compute the mean monthly percent increase in the number of badminton clubs.

Answer: এখানে একটি নিম্ন ক্ষয়ের দুটা বলা আছে। তার পুনরাবৃত্তির মূল মূল্য 20 B Period এর জেনে 155.

$$\therefore GCM = \sqrt[n-1]{\frac{\text{value at the end of period}}{\text{value at the start of period}}} - 1$$

$$= \sqrt[11]{\frac{155}{20}} - 1$$

12 মাসের interval

সুব, $n=12$

$$= \sqrt[11]{\frac{155}{20}} - 1$$

$$= \sqrt[11]{7.75} - 1 = 0.2046$$

\therefore Badminton clubs are increasing at a rate of almost 0.2046 or 20.46% per month.

Combined Mean

মনে কর ৩টি শ্রেণী আছে। Group 1, Group 2, Group 3. Group 1 এ Data আছে 12টি ও তাদের Sample mean 10, Group 2 এ Data 10টি ও Sample mean 40. and Group 3 এ Data 15টি ও তাদের Sample mean 50.

एकल घासव combined mean है—

$$\bar{X}_{CM} = \frac{N_1 \bar{X}_1 + N_2 \bar{X}_2 + N_3 \bar{X}_3}{N_1 + N_2 + N_3}$$

प्रमाण।

Group 1 is sample
एक समूह
एक डेटा
संग्रह
mean

#Problem 1: A study comparing the typical household incomes for 3 districts in the city of Manila was initiated to see where differences in household incomes lie across districts. The mean household incomes for a sample of 45 different families in three districts of Manila are shown below. Calculate a combined mean to obtain the average household income for all 45 families in the Manila Sample.

District 1

$$\bar{X}_1 = 30,400$$

$$N_1 = 12$$

District 2

$$\bar{X}_2 = 27,300$$

$$N_2 = 18$$

District 3

$$\bar{X}_3 = 42,500$$

$$N_3 = 15$$

Solution:

$$\text{if } \bar{X}_{CM} = \frac{N_1 \bar{X}_1 + N_2 \bar{X}_2 + N_3 \bar{X}_3}{N_1 + N_2 + N_3}$$

$$= \frac{(30,400 \times 12) + (27,300 \times 18) + (42,500 \times 15)}{12 + 18 + 15}$$

$$= 33,793.33.$$

Thus the combined mean in three district of Manila is 33,793.33.TK.

Median

(মুর্দ্ধক)

Median কল Midpoint of the data array.

Data array কি কিনিঃ? Data array কল ascending / Descending order এ আজাগো অনেকগুলো Data.

Median এবং ফিল্ড বিশিষ্টি:

i) এই অবশ্যই unique হবে। ঘানে কেনো array থেকে শুরুজ্ঞাত ১টি Median হওয়া পাওয়া গিয়ে।

ii) Data কত বড় বা ছোট কল তার জায়ে Median এর গুরুত্ব নেই।

iii) Median বের করার জন্য অবশ্যই Data গুলোকে ascending / Descending order এ আজাগো হবে।

Median for ungrouped Data

* যদি বিজোড় অংশুক Data থাকে

অর্থাৎ if the n is odd, the median is the middle ranked.

and if n is even, the median is the average of the two middle ranked values. (example এ দেখো)

যদি n odd হয় তাহলে Median হবে $\frac{n+1}{2}$ th Data টি।

#Example 1: Find the median of the ages of 9 middle management employees of a certain company. The ages are 53, 45, 59, 48, 54, 46, 51, 58 and 55.

Solve:

Step 1: Data পুরোকে ascending order এ আজাব / Ascending

45, 46, 48, 51, 53, 54, 55, 58, 59

Data পুরোটি $\therefore n=9$ (odd)

\therefore Median হল $\frac{n+1}{2}$ th Data

$$= \frac{10}{2} \text{ th Data}$$

$$= 5 \text{ th Data}$$

5th Data কল 53 \therefore Median 53.

Hence the median age is 53 years.

Example 2:

The Daily rates of a sample of eight employees at GTMS Inc. are 550 taka, 420 Tk, 560 Tk, 500 Tk, 700 Tk, 670 Tk, 860 Tk, 480 Tk. Find the median Daily rate of employee.

Solution:

Descending order / Ascending order এ আজাব:

420, 480, 500, 550, 560, 670, 700, 860
↑ ↑
4th 5th

এখানে Data অংশটি 8 $\therefore n=8$

$$\text{অর্থনৈতিক} \quad \text{Median} = \frac{n+1}{2} = \frac{9}{2} = 4.5$$

4.5 mean যাকে 4th & 5th Data এর average. যদি 7.5 হাতে পাইল

7th & 8th Data এর average.

$$\therefore \text{Median} = \frac{550+560}{2} = \frac{1110}{2} = 555$$

\therefore The median Daily rate is 555 taka.

Median for grouped Data:

একটি example দিয়ে আবৃত্ত পুরোটি বুকেব:

Using the example provided in on SJS Travel Agency.

Determine the Median of the frequency distribution.

on the ages of 50 people taking travel tours.

class-limits	Frequency
18-26	3
27-35	5
36-44	9
45-53	14
54-62	11
63-71	6
72-80	2

Solution:

class-limits	freqn (f)	cumulative freqn (cf)
18-26	3	3
27-35	5	$3+5=8$
36-44	9	$8+9=17$
45-53	14	$17+14=31$
54-62	11	$31+11=42$
63-71	6	$42+6=48$
72-80	2	$48+2=50$
	Total = 50	

এখানে Data আছে 50টি. কারণ $\frac{\text{total}}{2} = \frac{50}{2} = 25$ $\therefore N=50$

$$\frac{N}{2} = \frac{50}{2} = 25.$$

আবাসের মাঝে বর্তে 25 no Data কোন class এ পড়তে।

মধ্য, 18-26 class এ Data 3টি ও 27-35 এ Data 5টি, শান্ত আবাস
 $3+5=8$ টি Data চলে গেল। অবশ্য 36-44 class এ Data 9টি, শান্ত
আবাস $8+9=17$ টি Data চলে গেল। 45-53 এ Data আছে 14টি.

শান্ত 17টি Data এ পরবর্তে 14টি Data আছে 45-53 class এ শান্ত
বলা যায় 25th Data এ আছে 45-53 class এ। তাই এর নাম Median
class.

Now, formula

$$\text{Median} = LB + \left(\frac{\frac{N}{2} - Cf}{f} \right) Ci$$

আবাস যাপনে Median class এর class boundary বেঁকবিব। কারণ
table এ দেখা গেল continuous type class বা।

Median class বল: 45-53

\therefore class Boundary: 44.5-53.5

LB? : LB হল Median class Boundary এর lower limit.

$$\therefore LB = 44.5$$

cf? Median class এর আগের ক্লাসের cf = 17

f? Median class এর frequency = 14

(i)? Median class Boundary এর interval = $53.5 - 44.5$
 $= 9$

Median = $LB + \left(\frac{\frac{N}{2} - cf}{f} \right) (i)$

$$= 44.5 + \left(\frac{25 - 17}{14} \right) (9)$$

$$= 49.64$$

Quartiles, Deciles & Percentiles

আঞ্চলিক এবং গৃহ শাস্ত্র ব্যাখ্যা ও math বর্ণনা

Quartiles এর formula

ungrouped
$$Q_k = \frac{k(N+1)}{4}$$

grouped
$$Q_k = LB + \left(\frac{\frac{kN}{4} - cf}{f} \right) (i)$$

Deciles এর formula

$$D_k = \frac{k(N+1)}{10}$$

$$D_k = LB + \left(\frac{\frac{kN}{10} - cf}{f} \right) (i)$$

Percentiles এর formula

$$P_k = \frac{k(N+1)}{100}$$

$$P_k = LB + \left(\frac{\frac{kN}{100} - cf}{f} \right) (i)$$

#Math 78: Find the 1st, 2nd and 3rd quartiles (গুরুত্ব k=1, 2, 3) of the ages of 9 middle-management employees of a certain company. The ages are 53, 45, 59, 48, 54, 46, 51, 58, 59

solution:

আবাদের আগে কাজ হল Data খুলোকে Ascending order করানো:

45, 46, 48, 51, 53, 54, 55, 58, 59

$$\text{Now, } Q_1 = \frac{\frac{k}{4}(N+1)}{4} = \frac{\frac{1}{4}(9+1)}{4} = 2.5$$

∴ আবাদের উপরের জাজনো array তে 2.5th data কোনটি?

2.5th Data কাজ 23 3 no. data এর average.

$$\therefore Q_1 = \frac{46+48}{2} = 47$$

$$\therefore Q_2 = \frac{2(N+1)}{4} = \frac{2 \times (9+1)}{4} = \frac{20}{4} = 5$$

∴ 5th data. ∴ Q₂ = 53.

$$Q_3 = \frac{3(N+1)}{4} = \frac{3 \times 10}{4} = \frac{30}{4} = 7.5$$

কাজ 7 no 13 8 no data এর average.

$$\therefore Q_3 = \frac{55+58}{2} = \frac{113}{2} = 56.5 \text{ (Ans)}$$

Grouped Data এর solve একদম Demian বের বচ্চায় ঘটা।

#Math2:

Using the example provided in on SJS Travel Agency.

Determine the $Q_1, Q_2, Q_3, D_2, P_{22}$ of the frequency distribution on the ages of 50 people taking travel tours.

Class Limits	Frequency
18-26	3
27-35	5
36-44	9
45-53	14
54-62	11
63-71	6
72-80	2

solution:

Class Limits	Frequency	cF
18-26	3	3
27-35	5	8
36-44	9	17
45-53	14	31
54-62	11	42
63-71	6	48
72-80	2	50
N = 50		

for Q1 :

Q1 ঘাণ, $k=1$. আবাদের জ্যাব আগে যেব বজ্যা লাগবে - Quartiles এর

$$\text{কষ্টে } \frac{kN}{4} = ?$$

এখানে, $N=50$ (ques এর বলা)

$$k=1$$

$$\therefore \frac{kN}{4} = \frac{50}{4} = 12.5.$$

কষ্টে, 12.5 , Data টি কত class ও পড়ছে? 18-26 class এ Data

আছে $3f$ (f এর value), 27-35 এ আছে $5f$. তাহলে 18-35

এর ঘাণে $8f$ আছে। আবার 36-44 এর ঘাণে Data আছে $9f$ ।

তাহলে 8 মেরে $8+9=17$ এর Data সুলো আছে 36-44 class এ.

$\therefore 12.5$ এর Data আছে 36-44 class এ.

এখন, তাহলে Desired class = 36-44

\therefore class Boundary : $36.5 - 44.5$

$\therefore LB = 35.5 ; cf = 8$ [36-44 class এর আগের class এর cf টি]

and $f = 9$ [36-44 class এর f] and interval = 9

$$\therefore Q_1 = LB + \left(\frac{\frac{kN}{4} - cf}{f} \right) (i) = 35.5 + \left(\frac{\frac{1 \times 50}{4} - 8}{9} \right) (9) = 40$$

For Q₂:

$$\frac{KN}{4} = \frac{2 \times 50}{4} = 25$$

Given 25 Data with 45-53 class

∴ Class Boundary : 44.5 - 53.5

∴ LB = 44.5 ; cf = 17 ; f = 14 ; i = 9

$$\therefore Q_2 = LB + \left(\frac{\frac{KN}{4} - cf}{f} \right) (i)$$

$$= 44.5 + \left(\frac{\frac{2 \times 50}{4} - 17}{14} \right) (9)$$

$$= 49.64$$

For Q₃:

$$\frac{KN}{4} = \frac{3 \times 50}{4} = 37.5$$

37.5 Data with 54-62 class

∴ Class Boundary : 53.5 - 62.5

LB = 53.5, cf = 31, f = 11, i = 9

$$\therefore Q_3 = LB + \left(\frac{\frac{KN}{4} - cf}{f} \right) (i) = 53.5 + \left(\frac{\frac{3 \times 50}{4} - 31}{11} \right) (9) = 58.82$$

for D₂:

Deciles এর জন্য আগে check করতে হবে $\frac{KN}{10}$.

$$\therefore \frac{KN}{10} = \frac{22 \times 50}{100} = 11.$$

জুন 35th Data কম 54-62 class এ.

class Boundary : 53.5 - 62.5

$$\therefore LB = 53.5, cf = 31, f = 11 \quad (i) = 9$$

$$\therefore D_2 = LB + \left(\frac{\frac{KN}{10} - cf}{f} \right) (i)$$

$$= 53.5 + \left(\frac{\frac{22 \times 50}{100} - 31}{11} \right) (9) = 56.22 \text{ (Ans)}$$

for P₂₂:

Percentiles এর জন্য আগে check করতে হবে $\frac{KN}{100}$.

$$\therefore \frac{KN}{100} = \frac{22 \times 50}{100} = 11$$

11th Data আছে 36-44 class এ.

∴ class Boundary : 35.5 - 44.5

$$\therefore LB = 35.5, cf = 8, f = 9; (i) = 9$$

$$\begin{aligned} \therefore P_{22} &= LB + \left(\frac{\frac{22N}{100} - cf}{f} \right) (i) = 35.5 + \left(\frac{\frac{22 \times 50}{100} - 8}{9} \right) \times 9 \\ &= 38.5 \text{ (Ans)} \end{aligned}$$

Harmonic Mean

Harmonic mean শুরু হওয়া concept.

For ungrouped Data:

ungrouped data এর formula হল:

$$\frac{n}{\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} + \dots + \frac{1}{n_n}}$$

example কিরণ বুলি:

#Example 1:

Find Harmonic mean of the numbers 3, 6, 24, 48.

সারল number/Data আছে এখন 4টি

$$\therefore n = 4$$

$$\text{H.M} = \frac{n}{\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} + \frac{1}{n_4}}$$

$$= \frac{4}{\frac{1}{3} + \frac{1}{6} + \frac{1}{24} + \frac{1}{48}}$$

$$= 2.11 \text{ (Ans)}$$

For grouped Data:

grouped Data ഏ ഫോറ്റേ formula ക്ക് : H.M. = $\frac{\sum f}{\sum (\frac{f}{m})}$

example എൻ ക്ലേർ ബാബു !

Example 2 : Find Harmonic mean

Marks	f	m(midpoint)	$\frac{f}{m}$
10-20	6	15	0.4
20-30	14	25	0.56
30-40	22	35	0.629
40-50	07	45	0.16
50-60	1	55	0.02

$$\sum f = 50$$

$$\therefore \sum \left(\frac{f}{m} \right) = 1.769$$

$$\therefore H.M. = \frac{\sum f}{\sum \left(\frac{f}{m} \right)} = \frac{50}{1.769} = 28.26$$

Dispersion

Dispersion করা actual value & average value এর difference.

Dispersion এর মান প্রতিটি মাত্রার মধ্যে আছে। মাত্রা এবং তা একটি কারণ কিম্বা:

Range:

The difference of the highest value and the lowest value in the data set.

মানুষের Data set:

Example: The daily rates of a sample of eight employees at GCSM Inc. are 550, 420, 560, 500, 700, 670, 860, 480. Find the range.

Answer:

আগে Ascending order এ আজায়: (না আজালে শু হয়)

420, 480, 500, 550, 560, 670, 700, 860

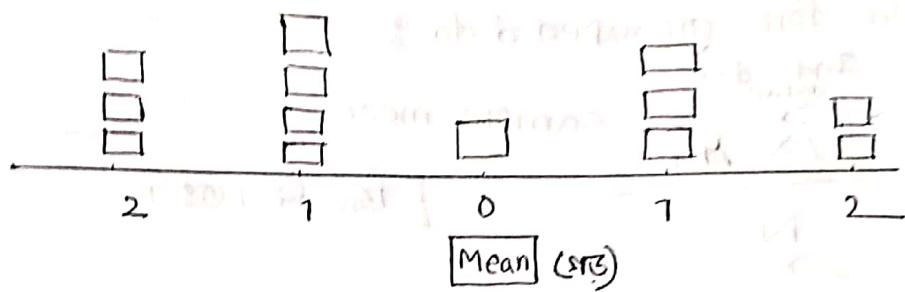
Highest value = 860

Lowest value = 420

$$\therefore \text{range} = 860 - 420$$

$$= 440$$

Average Deviation: ~~বিভিন্ন কাছের অন্তরের গড়~~
It is the absolute difference between the element and a given point.



দুই mean কেন একপাশ দুই side এর দ্বারা ছাপ্পাইয়া দেওয়া হল । এখন কোথা কল 0 এর বাজে
পাশে -1, -2 কেন লিখলাইনা। বগুন কল 6 জাতের বলা আছে absolute
value. এখন -1 এ Data আছে 4ff (4ff box) তাকেন কল তো -1 point
যাকা Data শুলোর Deviation mean point থেকে কল 2। আবার
2 point-এ যাকা Data শুলোর mean point থেকে Deviation 2.
Mean point-এর উপরিকের জন্যও same deviation.

এখন এই deviation শুলোর একটি average করা কল 2 যোগ্য পাও
average deviation.

Average Deviation for ungrouped Data :

$$AD = \frac{\sum |x - \bar{A}|}{N}$$

Individual
Data → sample
Mean
Population

Sample $AD = \frac{\sum |x - \bar{x}|}{n}$

Average Deviation for grouped data :

$$AD = \frac{\sum f |x - \bar{x}|}{N}$$

Frequency → Individual
value → sample mean
Population

[শিখি লিখি অঙ্ক]

Sample $AD = \frac{\sum f |x - \bar{x}|}{n}$

Population

At Math এখনে আবার clear হবে

Example 1: The daily rates of a sample of eight employees at GEMS Inc. are 550, 420, 560, 500, 700, 670, 860, 480. Find the average deviation.

Solution: Average Deviation एवं formula: $AD = \frac{\sum |x - \bar{x}|}{n}$

एकान्त sample of eight employees $\bar{x} \text{ तथा } \bar{x} = \frac{\sum x}{n}$

$$\therefore n=8 \quad \bar{x} = \frac{550+420+560+500+700+670+860+480}{8} = 592.5$$

x	$x - \bar{x}$	$ x - \bar{x} $
550	-42.5	42.5
420	-172.5	172.5
560	-32.5	32.5
500	-92.5	92.5
700	107.5	107.5
670	27.5	27.5
860	267.5	267.5
480	-112.5	112.5

$$\sum (x - \bar{x}) = 0$$

$$\sum |x - \bar{x}| = 905$$

$$\therefore AD = \frac{\sum |x - \bar{x}|}{n} = \frac{905}{8} = 113.125$$

#Example 2:

The Data below shows the frequency distribution of the amounts of electric consumption of a typical household in Batangas city for the month of January, 2009. Find the average deviation.

Amount of Electric Bill.	Number of Families
700 - 849	2
850 - 999	9
1000 - 1149	15
1150 - 1299	9
1300 - 1499	5

Solutions

Midpoint

Class limits	f	x	fx	$ x - \bar{x} $	$f x - \bar{x} $
200-849	2	774.5	1549	322.5	645.0
850-999	9	924.5	8320.50	172.5	1552.5
1000-1149	15	1024.5	16,117.50	22.5	337.5
1150-1299	9	1124.5	11020.50	127.5	1147.5
1300-1499	5	1324.5	6872.50	277.5	1387.5
	$\sum f = 40$		$\sum fx = 43,880$		

$$\sum f|x - \bar{x}| = 5070$$

দেখ, আক্ষর grouped Data এর জন্য \bar{x} মেতাবু কিভাবে?

$$\bar{x} = \frac{\sum fx}{n} = \frac{43,880}{40} = 1097$$

অবশ্য $|x - \bar{x}|$ হিসাব করা।

$$\therefore AD = \frac{\sum f|x - \bar{x}|}{n} = \frac{5070}{40} = 126.75$$

Variance

#

T formula:

$$\text{For ungrouped Data, } s^2 = \frac{\sum (x - \bar{x})^2}{n-1}$$

$$= \frac{\sum x^2 - \sum \bar{x}^2}{n-1}$$

mean

এখন হেয়াল কর, $\bar{x} = \frac{\sum x}{n}$

$\sum \bar{x}^2$ কে লিখা যায় $\frac{(\sum x)^2}{n}$ \rightarrow এটি population. তাই square করে

$$\therefore s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$$

এই formula কে আব্দি variance এর math করব

s^2 বেঁধে করে root করে finally বেঁধে করব S (sample standard Deviation)

for grouped Data :

$$s^2 = \frac{\sum f(x - \bar{x})^2}{n-1}$$

$$= \frac{\sum f x^2 - \sum f \bar{x}^2}{n-1}$$

$$= \frac{\sum f x^2 - \frac{(\sum f x)^2}{n}}{n-1}$$

[grouped data ওর জন্য
 \bar{x} হিসেবে $\frac{\sum f x}{n}$]

$$\therefore s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1} \rightarrow [(\bar{x})^2 \text{ নিয়ে একটা}] \rightarrow \text{এই formula use করব always.}$$

root করে last ০.৫ দেয় করা।



এই টি কল Sample standard Deviation

Math 1: The daily rates of a sample of eight employees at GMS Inc. are 550, 420, 560, 500, 700, 670, 860, 480. Find the average sample standard Deviation / S.D.

solution:

$$\text{আমে } \bar{x} \text{ দেয় করে নিষ্ঠ; } \bar{x} = \frac{\sum x}{n} = \frac{550+420+560+500+700+670+860+480}{8} \\ = \frac{4740}{8} \\ = 592.50$$

$$\text{আবার ফর্মুলা: } s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$$

তাই x^2 ও বের করতে করে আবার আমে।

X	X^2
550	302,500
420	176,400
560	313,600
500	250,000
700	490,000
670	448,900
860	739,600
480	230,400

$$\therefore \sum x = 4740$$

$$\sum x^2 = 2951,400$$

$$\therefore s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$$

$$= \frac{2951400 - \frac{4740^2}{8}}{8-1}$$

$$= 20,421.43$$

$$\therefore s = \sqrt{20,421.43}$$

$$= 142.90$$

(Ans)

Math 2 : Using the example provided in chapter 2 on SJS Travel Agency . Determine the variance and standard deviation of the frequency distribution on the ages of 50 people taking travel tours.

Class Limits	Frequency
18 - 26	3
27 - 35	5
36 - 44	9
45 - 53	14
54 - 62	11
63 - 71	6
72 - 80	2

Solution:

Class limits	f	x(midPoint)	f_x^1	f_x^2
18 - 26	3	22	66	1452
27 - 35	5	31	155	4805
36 - 44	9	40	360	14400
45 - 53	14	49	686	33674
54 - 62	11	58	638	37004
63 - 71	6	67	402	26934
72 - 80	2	76	152	11552

$$\sum f_n = 2459 \text{ and } \sum f_n^2 = 129,761$$

$$\therefore s^2 = \frac{\sum f_n^2 - \frac{(\sum f_n)^2}{n}}{n-1}$$

$$= \frac{129,761 - \frac{(2459)^2}{850}}{50-1}$$

$$= 780.75$$

$$\therefore s = 28.22 \text{ (Ans)}$$

Population Variance

প্রকৃতি আবাসের মধ্যে variance টি হিল sample variance. কাবন তেওঁদের আবাস use করেছিলাম n . তা একইভাবে population variance বের করা যায়। শুরু formula তে population mean (M) রচে sample mean (\bar{x}) এর জায়মায়।

$$\begin{array}{c} \text{Individual} \\ \text{Data: } \sum (x_i - M)^2 \\ \therefore \sigma^2 = \frac{\sum (x_i - M)^2}{N} \end{array}$$

Population
mean

Population
variance

$$\therefore \sigma = \sqrt{\sigma^2}$$

\rightarrow

Population Standard Deviation

#Example The monthly income of the five research directors of recoletos schools are: 55,000, 59,000, 62,500, 57,000 and 61,000. Find the ^{population} variance and standard deviation.

Ans: এই দেখ, কোথাও sample/ population এর Data যাবা নাই।
মধ্যন কাশে এ sample বা population একই জিমিস। $n=N$.

$$\text{Now, } \bar{X} = \frac{\sum X}{N} = \frac{55,000 + 59,000 + 62,500 + 57,000 + 61,000}{5} \\ = 59,000.$$

X	$X - \bar{X}$	$(X - \bar{X})^2$
55,000	-4000	16,000,000
59,500	500	250,000
62,500	3500	12,250000
57,000	-2000	4000000
61,000	2000	4000000
$\sum X = 295,000$		$\sum (X - \bar{X})^2$ $= 36,500,000$

$$\begin{aligned}\therefore \sigma^2 &= \frac{\sum (n-\bar{x})^2}{N} \\ &= \frac{36,500,000}{5} \\ &= 7,300,000\end{aligned}$$

$$\therefore \sigma = 2701.85 \text{ (Ans)}$$

Interquartile range

ଆগ୍ରହୀ already Quartiles part ପାଇଁ Q_1, Q_2, Q_3 ଏକିମୋ ଦେଇ କାହା

କିନ୍ତୁ ଏଥିର କିମ୍ବା ଏଥିର Interquartile range ଏବଂ formula କିମ୍ବା

$$\text{Interquartile Range (IQR)} = Q_3 - Q_1$$

Math : Find the interquartile range of the ages of 9 middle management employees of a certain company.

The ages are 53, 45, 59, 48, 54, 46, 51, 58, and 55.

Solve:

ଆଗେ Q_1 ଓ Q_3 ଦେଇ କାହାରେ।

Ascending order ମାଜାଲେ,

45 46 48 51 53 54 55 58 59

ungrouped Data ഏ ജ്ഞ formula:

for Q₁: $\frac{K(N+1)}{4} = \frac{1 \times (9+1)}{4} = 2.5$ താഴെ Data

ഇന്ന് 2 3 3 ന് Data ഏ average.

$$\therefore Q_1 = \frac{46+48}{2} = 47$$

and for Q₃: $\frac{K(N+1)}{4} = \frac{3 \times (9+1)}{4} = 7.5$

ഇന്ന് 7 8 8 ന് Data ഏ average.

$$\therefore Q_3 = \frac{55+58}{2} = \frac{113}{2} = 56.5$$

∴ Interquartile range is = Q₃ - Q₁ = 56.5 - 47 = 9.5.

Quartile Deviation: കുറച്ച നാം ജുസ്റ്റ് മുമ്പ് യാഥെവാ $QD = \frac{Q_3 - Q_1}{2}$

ഡാറ്റയെ പരിഗ്രഹിച്ചുള്ളൂടെ Quartile Deviation കണ്ടു.

$$\therefore QD = \frac{Q_3 - Q_1}{2}$$

$$= \frac{9.5}{2} = 4.75$$

coefficient of variation

(CV)

sample mean এর আমেন্টে, $CV = \frac{s}{\bar{x}} \times 100\%$

sample
mean

sample standard deviation

population mean এর আমেন্টে, $CV = \frac{\sigma}{\mu} \times 100\%$

population mean

population standard deviation

* এছাঁ দেখে যদি বুকা না যায় কেননি? Population আর কোনী sample
এভাবে case এ যাই বকলা চিহ্ন রয়ে। তবে আরু sample টিকে নিয়ে
যোগানোই চেষ্টা prefer করি।

Example: The average age of the engineers at
VSAS pipeline corporation is 33 years, with a standard
deviation of 3. The average monthly salary of the
engineers is 45,000 with standard deviation of 3150.
Determine the coefficient of variations of age and
salary.

solution :

দেশ বুকাতে পাবছিনা population দেওয়া নাকি sample

এখি sample ঘৰেই আগাৰ যাই

∴ for age:

average age দেওয়া 33 years : - $\bar{x} = 33$

আৰ standard deviation 3 ∴ $s = 3$

$$\therefore CV = \frac{s}{\bar{x}} \times 100\% = \frac{3}{33} \times 100\% = 10\%$$

for salary;

$s = 3150$ and $\bar{x} = 45,000$

$$\therefore CV = \frac{s}{\bar{x}} \times 100\%$$

$$= \frac{3150}{45000} \times 100\% = 7\%$$

∴ coefficient of variation is larger for age, the ages are more variable than salary.

Example 2:

The mean of commissions of Educational Insurance over 1 year period is 12,500 and the standard deviation is 7350. The mean of the number of sales is 60 and the variance is 56.25. Compare the variations of the two.

Solution:

For commission,

$$CV = \frac{s}{\bar{x}} \times 100\% = \frac{7350}{12,500} \times 100\% = 10.8\% \text{ commission}$$

For sales,

$$CV = \frac{s}{\bar{x}} \times 100\% = \frac{\cancel{60} 56.25}{60} \times 100\% = 72.5\%$$

∴ the coefficient of variation is larger for sales, the sales are more variable than the commissions

Moments

Moments এর type কোনটি:

i) Moments about an Arbitrary point

ii) Moments about mean.

Moments about an Arbitrary point:

কিন্তু Ungrouped M_r' দ্বারা express কৰা কৰ্য। যদি Arbitrary point A রয়ে

$$M_r' = \frac{\sum (x-A)^r}{n} \quad \text{for } r^{\text{th}} \text{ moment.}$$

রয়ে থাকায় r^{th} moment / r^{th} moment.

for grouped Data,

$$M_r' = \frac{\sum f(x-A)^r}{\sum f} \quad \text{for } r^{\text{th}} \text{ moments.}$$

[$\sum f$ মে কখন n একই কৰ্যা কৰানো শুধুমাত্র change আব্যাস। এখানে just
f পুনৰ কৰ্য]

Moments about mean: M_r দ্বারা express কৰা কৰ্য। Mean এর sign

আছে জানি \bar{x}

$$\text{For ungrouped data: } M_{rc} = \frac{\sum (x - \bar{x})^r}{n}$$

$$\text{For grouped data: } M_{rc} = \frac{\sum f(x - \bar{x})^r}{\sum f}$$

এখন দেখ, $\sum f$ ক্যে কথা N অথবা n একই কথা। কারণ $\sum f = N$ হী আজ।
আব্বা— already দেখলাব।

$$M_{rc} = \frac{\sum f(x - \bar{x})^r}{\sum f}$$

$$\text{একে নেওয়া শায় } M_{rc} = \frac{1}{N} \sum_{i=1}^n f_i (x_i - \bar{x})^r; r=0, 1, 2, 3, \dots$$

মনি, $r=0$ ক্যে তাঁল

$$M_0 = \frac{1}{N} \sum_{i=1}^n f_i (x_i - \bar{x})^0 = 1$$

এখন, $r=1$ ক্লে,

$$M_1 = \frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^1 = 0 \quad \leftarrow \text{এইটি zero আজো always আস্থায় রাখবে, } r=1 \text{ এলে } M_1=0$$

$r=2$ ক্লে,

$$M_2 = \frac{1}{N} \sum_{i=1}^n (x_i - \bar{x})^2 = \sigma^2 \quad [\text{গাওয়া রাখব}]$$

Population standard deviation

for $n=3$

$$M_3 = \frac{1}{N} \sum_{i=1}^n f_i (x_i - \bar{x})^3$$

for $n=4$

$$M_4 = \frac{1}{N} \sum_{i=1}^n f_i (x_i - \bar{x})^4$$

এব যেমি আৰ লাখবেনা।

ওৱাৰ, moments about an arbitrary point A

আছিবা দেখোচি প্ৰয়োজ্য আমেৰি

$$M'_n = \frac{1}{N} \sum_{i=1}^n f_i (x_i - A)^n, n=0, 1, 2, 3$$

মদি, $n=0$ ক্ষম,

$$M'_0 = \frac{1}{N} \sum_{i=1}^n f_i (x_i - A)^0 = 0$$

$n=1$ কল, $M'_1 = \frac{1}{N} \sum_{i=1}^n f_i (x_i - A)^1 = 0$ [আমেৰ ছত্ৰ $n=1$ এৰ জন্য
 M'_2 এৰ ছত্ৰ M'_1 ও ০]

$n=2$ কল, $M'_2 = \frac{1}{N} \sum_{i=1}^n f_i (x_i - A)^2$.

$n=3$ কল, $M'_3 = \frac{1}{N} \sum_{i=1}^n f_i (x_i - A)^3$

$n=4$ কল, $M'_4 = \frac{1}{N} \sum_{i=1}^n f_i (x_i - A)^4$

Relation between moments about
mean and moments about an
arbitrary point A

Very important proof/Derivation.

$$\begin{aligned} \text{We know, } M_n &= \frac{1}{N} \sum_{i=1}^n f_i (x_i - \bar{x})^n \\ &= \frac{1}{N} \sum_{i=1}^n f_i (x_i - A + A - \bar{x})^n \dots \dots \dots (i) \end{aligned}$$

$$\text{Again, } M'_1 = \bar{x} - A \Rightarrow A - \bar{x} = -M'_1$$

\therefore from (i),

$$M_n = \frac{1}{N} \sum_{i=1}^n f_i [(x_i - A) - M'_1]^n$$

Now by Binomial theorem,

$$\begin{aligned} M_n &= \frac{1}{N} \sum_{i=1}^n f_i [(x_i - A)^n - n c_1 (x_i - A)^{n-1} M'_1 + n c_2 (x_i - A)^{n-2} M'^2_1 \\ &\quad - n c_3 (x_i - A)^{n-3} M'^3_1 + \dots \dots \dots + (-1)^n M'^n_1] \end{aligned}$$

$$\begin{aligned} \therefore M_n &= \frac{1}{N} \sum_{i=1}^n f_i (x_i - A)^n - n c_1 M'_1 \frac{1}{N} \sum_{i=1}^n f_i (x_i - A)^{n-1} \\ &\quad + n c_2 M'^2_1 \frac{1}{N} \sum_{i=1}^n f_i (x_i - A)^{n-2} \end{aligned}$$

$$-r c_{c_3} M_1' \frac{1}{N} \sum_{i=1}^n f_i (x_i - A)^{r-3} + \dots + (-)^r M_1'^r$$

$$M_R = M'_R - r c_{c_1} M_1' -$$

$$M_R = M'_R - r c_{c_1} M'_{R-1} M_1' + r c_{c_2} M'_{R-2} M_1'^2 - r c_{c_3} M'_{R-3} M_1'^3 + \dots + (-)^R M_1'^R$$

where for $R = 2, 3, 4$ respectively.

(Done)

For $R = 2$

$$\begin{aligned} M_2 &= M'_2 - 2 c_1 M_1' M_1' + 2 c_2 M_1'^2 \cdot M_0' \\ &= M'_2 - 2 M_1'^2 + M_1'^2 M_0' \\ &= M'_2 - 2 M_1'^2 + M_1'^2 \quad [\because M_0' = 1] \\ &= M'_2 - M_1'^2 \end{aligned}$$

$R = 3$

$$\begin{aligned} M_3 &= M'_3 - 3 c_1 M'_2 M_1' + 3 c_2 M_1'^2 M_1' - 3 c_3 M_1'^3 M_0' \\ &= M'_3 - 3 M'_2 M_1' + 3 (M_1')^2 \cdot M_1' - (M_1')^3 \quad [\because M_0' = 1] \\ &= M'_3 - 3 M_1' M_2' + 2 (M_1')^3 \end{aligned}$$

For $n=4$ we get

$$\begin{aligned}
 M_4 &= M'_4 - 4c_1 M'_3 M'_1 + 4c_2 M'_2 (M'_1)^2 - 4c_3 (M'_1)^3 M'_1 + 4c_4 M'_1^4 M'_0 \\
 &= M'_4 - 4M'_3 M'_1 + 6M'_2 (M'_1)^2 - 4(M'_1)^4 + (M'_1)^4 \quad [\because M'_0 = 1] \\
 &= M'_4 - 4M'_3 M'_1 + 6M'_2 (M'_1)^2 - 3(M'_1)^4
 \end{aligned}$$

Example 18

Group	x	f	f_n	$f(x-\bar{x})^2$	$f(x-\bar{x})^3$	$f(x-\bar{x})^4$
1.5 - 2.0	1.75	2	3.5	5.445	-8.98425	14.824013
2.0 - 2.5	2.25	2	4.5	2.645	-3.04175	3.4980125
2.5 - 3.0	2.75	5	13.75	2.1125	-1.373125	0.8925313
3.0 - 3.5	3.25	15	48.75	0.3375	-0.050625	0.0075937
3.5 - 4.0	3.75	8	30	0.98	0.343	0.72005
4.0 - 4.5	4.25	6	25.5	1.335	3.68475	3.1320375
4.5 - 5.0	4.75	2	9.5	3.645	4.92075	6.6430125
Total	1.75 - 4.75	50	100	100	100	100

$$\text{দেরি, } \bar{x} = \frac{\sum f x}{\sum f} = \frac{135 \cdot 5}{40} = 3.4 ; \quad \sum f = 40$$

$$\sum f(x-\bar{x})^2 = 19.5 \quad \sum f(x-\bar{x})^3 = -4.50125 \quad \sum f(x-\bar{x})^4 = 29.11725$$

$$\therefore \mu_1 = \frac{\sum f(x-\bar{x})^2}{\sum f}$$

$$\text{for } \mu_2 = \frac{\sum f(x-\bar{x})^3}{\sum f}$$

$$= \frac{(19.5)}{40} = 0.5$$

$$\text{for } \mu_3 = \frac{\sum f(x-\bar{x})^3}{\sum f} = \frac{-4.50125}{40} = -0.1125$$

$$\text{for } \mu_4 = \frac{\sum f(x-\bar{x})^4}{\sum f} = \frac{29.11725}{40} = 0.7279$$

আবার কিন্তু example আছে গ্র্যাবের দেওয়া। আরু অশুল্লা slide আগামে add
করবে দিন্তি। same process সব।

slide from Nazmus Sharif Sir

Example of Moments (Continued)

Example

Class	f	X	fx	$X - \bar{X}$	$f(X - \bar{X})$	$f(X - \bar{X})^2$	$f(X - \bar{X})^3$	$f(X - \bar{X})^4$
0.0-0.8	5	0.4	2	-1.97	-9.84	19.37	-38.11	75.00
0.8-1.6	9	1.2	10.8	-1.17	-10.51	12.28	-14.34	16.75
1.6-2.4	15	2	30	-0.37	-5.52	2.03	-0.75	0.28
2.4-3.2	10	2.8	28	0.43	4.32	1.87	0.81	0.35
3.2-4.0	6	3.6	21.6	1.23	7.39	9.11	11.22	13.82
4.0-4.8	2	4.4	8.8	2.03	4.06	8.26	16.78	34.10
4.8-5.6	1	5.2	5.2	2.83	2.83	8.02	22.71	64.32
5.6-6.4	2	6	12	3.63	7.26	26.38	95.82	348.03
Total	50		118.4		0	87.31	94.14	552.65

$$\bar{X} = \frac{\sum fX}{\sum f} = \frac{118.4}{50} = 2.37$$

$$m_2 = \frac{\sum f(X - \bar{X})^2}{\sum f} = \frac{87.31}{50} = 1.75$$

$$m_3 = \frac{\sum f(X - \bar{X})^3}{\sum f} = \frac{94.14}{50} = 1.88$$

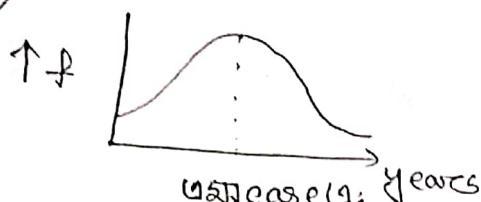
$$m_4 = \frac{\sum f(X - \bar{X})^4}{\sum f} = \frac{552.65}{50} = 11.05$$

skewness

* Any departure from symmetry is called skewness

অস্তু কুরু ৩ বিনেৰ কৱতো পাবো:

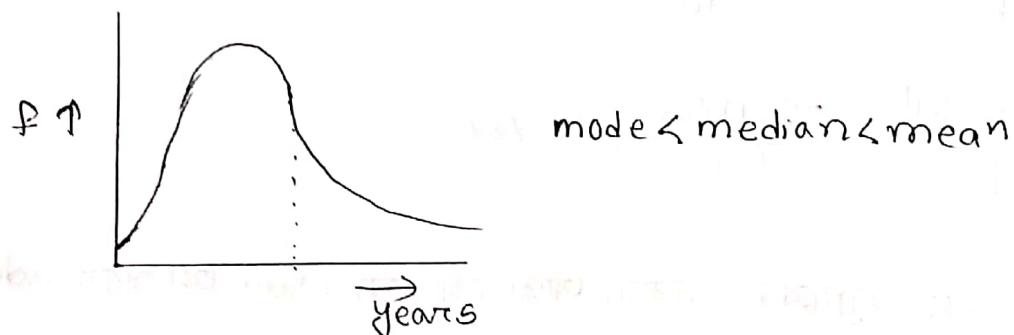
i)



$$\text{mean} = \text{mode} \\ = \text{median}$$

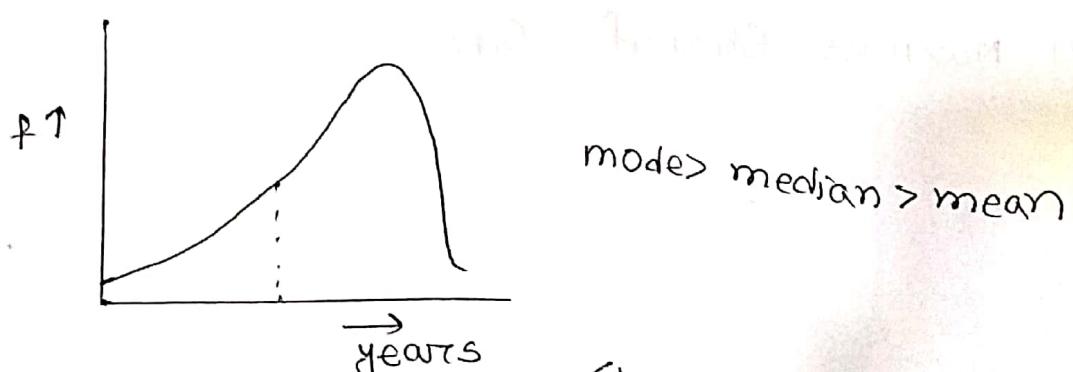
* middle point এৰ অল্যা মাছিৰ অব্যান । একে case ১ শকেন্সৰল
zero.

ii)



এছেয়ে middle point এৰ বাছামাছিৰ graph-ৰ বেশি। এভন ক্ষেত্ৰে—
skewness positive.

iii)



Skewness negative.

formulas

জাবার already mean, mode, standard deviation বেব কো পিছেই

$$\therefore Sk = \frac{\text{Mean} - \text{Mode}}{SD}$$

in some cases,

$$\text{Mode} = 3 \text{ Median} - 2 \text{ Mean}$$

$$\therefore Sk = \frac{\text{Mean} - 3 \text{ Median} + 2 \text{ Mean}}{SD}$$

$$= \frac{3 \text{ Mean} - 3 \text{ Median}}{SD}$$

$$= \frac{3(\text{Mean} - \text{Median})}{SD}$$

Sk ranges between -3 to +3

জাবার Bowley নাম্বৰ একজন scientist skewness তৰ কিন্তু অন্ধিষ্ঠিত:

$$Sk = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{Q_1 - 2Q_2 + Q_3}{Q_3 - Q_1} = \frac{Q_1 - 2 \text{ Median} + Q_3}{Q_3 - Q_1}$$

Q_1 mean কৰে ৫০% পাৰি Q_2 = Median কৰে যাব।

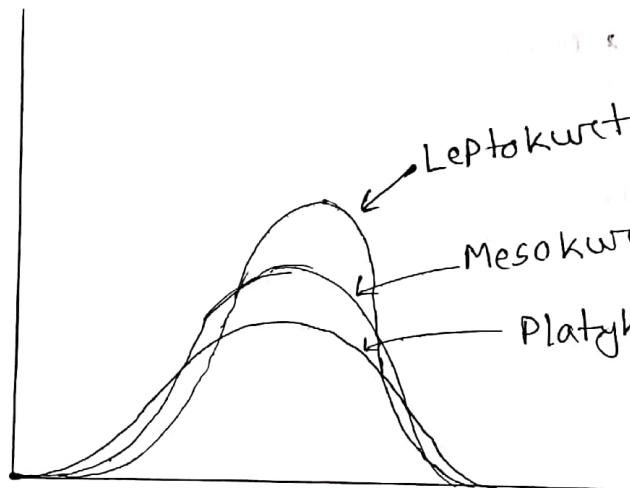
Skewness ଏବଂ ଆରେକଣ୍ଡ ଶ୍ରୀଅଛେ

moment about mean.

$$Sk = \frac{M_3}{\sigma^3} \quad \text{for population data}$$

$$Sk = \frac{m_3}{\sigma s^3} \quad \text{for sample Data.}$$

Kurtosis



Kurtosis mainly indicate କାଣେ ଯେ, curve ଏବଂ peakedness

* Normal curve ପୁଲୋ ଶ୍ରୀଏ ଲେଖିଗୁଣ୍ଠିତିରେ high ଓ ନା ଆଦାୟ flat ଓନା।

Mesokurtic କିମ୍ବା normal curve.

* ଏହା ମଧ୍ୟ କୋଣେ curve normal curve ଏବଂ ଛୁଲନାୟ flat କିମ୍ବା ତାଙ୍କିଳେ ଏବଂ Platykurtic

* ଆଦାୟ ମଧ୍ୟ କୋଣେ curve normal curve ଏବଂ ଛୁଲନାୟ high କିମ୍ବା ତାଙ୍କିଳେ Leptokurtic.

kurtosis is usually measured by the moment ratio (B_2)

$$\text{kurt} = B_2 = \frac{m_4}{m_2^2} \quad \text{for population Data}$$

$$\text{kurt} = b_2 = \frac{m_4}{m_2^2} \quad \text{for sample Data.}$$

* math ଟ ଆବଶ୍ୟକ clear ହେ ।

* kurt ଏବଂ value 3 କଣ୍ଠରେ normal curve denote ହେବେ ।

* kurt ଏବଂ value 3 ଏବଂ ବେଳେ କଣ୍ଠରେ leptokurtic curve denote ହେବେ

* kurt ଏବଂ value 3 ଏବଂ କଣ୍ଠରେ platykurtic curve denote ହେବେ ।