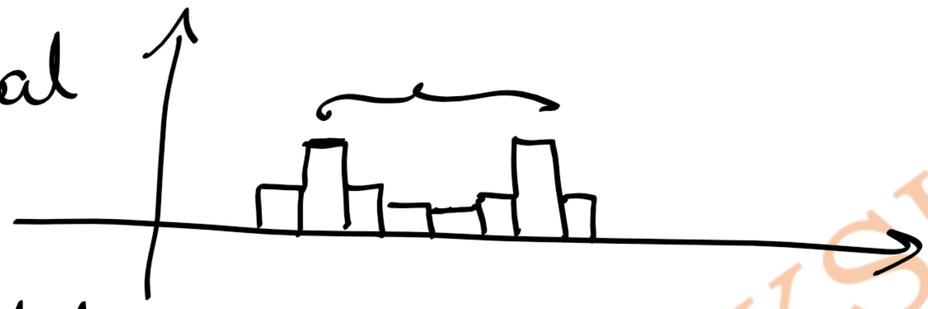
Mode

* most frequently occurring value in the data set
* ALL ✓
* Bimodal



* Multimodal

Mode of a grouped data

C.I	frequency

Modal class

$$\text{Mode} = l + \frac{h(f_1 - f_0)}{(f_1 - f_0) - (f_2 - f_1)}$$

DR. HARINAKSHI KARKERA

$$\text{Mode} = l + \frac{h(f_1 - f_0)}{(f_1 - f_0) - (f_2 - f_1)}$$

$$= l + \frac{h(f_1 - f_0)}{2f_1 - f_0 - f_2}$$

l — lower limit of mode interval

h — width of the "

f_1 — frequency of the modal class

f_0 — frequency of the class preceding the modal class

f_2 — frequency of the class succeeding the modal class

Class interval	Frequency
20-30	6
30-40	18
40-50	11
50-60	11
60-70	3
70-80	1

$$\text{Mode} = 30 + \frac{10(18-6)}{2\times 18 - 6 - 11} = \underline{36.31}$$

AKSHI KARKERA

Class interval	Frequency
20-30	6
30-40	18
40-50	11
50-60	11
60-70	3
70-80	1

frequency

6

12

13

11

3

1

Mode = 36.31

Median = ✓

unknown values

Dr. HARINAKSHI KARKERA

Quartiles

Partition values

25%	25%	25%	25%

Q_1 Q_2 Q_3

Q_2 = Median

Discrete dist^n

Q_1: $\frac{N}{4}$, c.f just greater than $\frac{N}{4}$, the corresponding value of the variable will be Q_1

Q_2: $\frac{2N}{4}$, " " $\frac{N}{2}$, " " " Q_2

Q_3: $\frac{3N}{4}$, " " " $\frac{3N}{4}$, " " " Q_3

Dr. HARINAKSHI KARKERA

Continuous distribution

C.I	frequency

Median

$$Q_2 = l + \frac{h}{f}\left(\frac{N}{2} - c\right)$$

Q_1 : $\frac{N}{4}$, find c.f just greater than $\frac{N}{4}$, the corresponding class will contain 'Q_1'

$$Q_1 = l + \frac{h}{f}\left(\frac{N}{4} - c\right)$$

Q_3 : $\frac{3N}{4}$, find c.f just greater than $\frac{3N}{4}$, the corresponding class will contain Q_3

$$Q_3 = l + \frac{h}{f}\left(\frac{3N}{4} - c\right)$$

Dr. HARINAKSHI KARKERA

Deciles

10%	10%	10%	10%	. . . . .

10 parts

9 points deviding the series into 10 equal parts

D_7: 70% of the observations before it

discrete dist^n:

D_7: $\frac{7N}{10}$, C.F just greater than $\frac{7N}{10}$, the corresponding value D_7

Continuous dist^n $\boxed{\text{C.I} \mid \text{frequency}}$

D_7: $\frac{7N}{10}$, C.F just greater than $\frac{7N}{10}$, the corresponding C.I will contain D_7.

$$D_7 = l + \frac{h}{f}\left[\frac{7N}{10} - c\right]$$

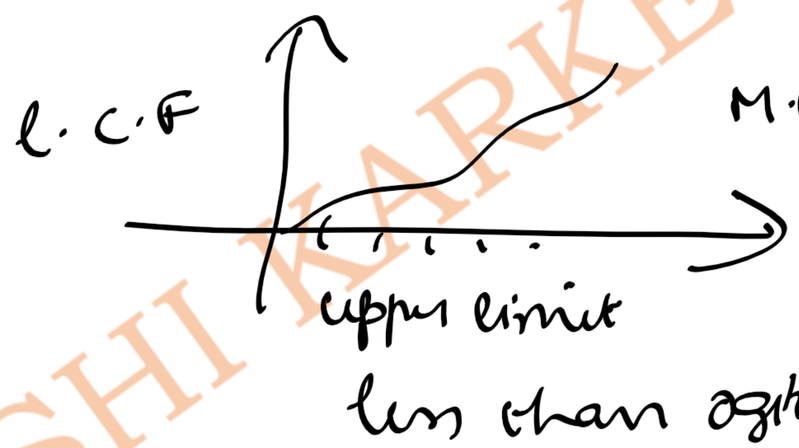
Dr. HARINAKSHI KARKERA

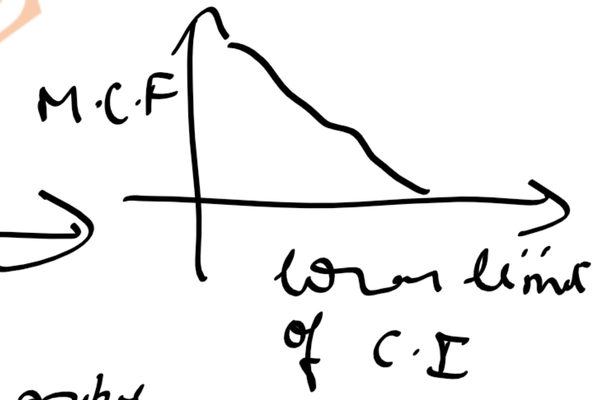
Percentiles

1%	1%	1%	. . . . .

99 points | 100 parts

P_{37} ; 37%





<u>For discrete dist^n</u>

P_{37} : $\frac{37N}{100}$, C.F just greater than $\frac{37N}{100}$, the corresponding value is P_{37}

<u>For continuous dist^n</u>

$$\boxed{P_{37} = l + \frac{h}{f}\left(\frac{37N}{100} - c\right)}$$

$\frac{37N}{100}$, C.F just greater than $\frac{37N}{100}$, corresponding C.I, contains P_{37}.

Dr. HARINAKSHI KARKERA

Measures of dispersion or Measures of variation

Mean: Students

$\underbrace{10 \text{ observations}}$

$= 55$]

$\left\{ \begin{matrix} 2 & — & 95 \\ 5 & — & 8 \end{matrix} \right] = 55$

Mode: Empirical relationship b/w Mean, Median & Mode

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

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

Dr. HARINAKSHI KARKERA

No Variability in Cash Flow
Mean
30
30
30
30
30
Variability in Cash Flow
Mean
30

Criteria

1) rigidly defined
2) calculate easily

Dr. HARINAKSHI KARKERA

Range :

$$X_{max} - X_{min}$$

Interquartile range

$$Q_3 - Q_1$$

Quartile Deviation (semi interquartile range)

$$\boxed{Q.D = \frac{(Q_3 - Q_1)}{2}}$$

Dr. HARINAKSHI KARKERA

Mean Deviation

x_i , $i = 1, 2, \ldots, n$

f_i , " "

x_i	f_i

$$\boxed{M.D = \frac{\sum_{i=1}^{n} f_i |x_i - A|}{N}}$$

A — average (Mean, Median, Mode)

$$N = \sum_{i=1}^{n} f_i$$

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

Median

Dr. HARINAKSHI KARKERA

Standard Deviation

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

$$N = \sum_{i=1}^{n} f_i$$

Variance

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

variance of x — variance of y

σ_x^2 — σ_y^2

Dr. HARINAKSHI KARKERA

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

$(a-b)^2$

Expand & simplify

$$\boxed{\sigma_x^2 = \frac{1}{N} \sum_{i=1}^{n} f_i x_i^2 - \left(\frac{1}{N} \sum_{i=1}^{n} f_i x_i\right)^2}$$

Dr. HARINAKSHI KARKERA

variance and consequently standard deviation is independent of change of origin.

$$\sigma_x^2 = \frac{1}{N}\sum_i f_i (x_i - \bar{x})^2 = \frac{1}{N}\sum_i f_i (x_i - A + A - \bar{x})^2$$

$$= \frac{1}{N}\sum_i f_i (d_i + A - \bar{x})^2, \text{ where } d_i = x_i - A.$$

$$\sigma_x^2 = \frac{1}{N}\sum_i f_i [\, d_i^2 + (A - \bar{x})^2 + 2(A - \bar{x})\, d_i \,]$$

$$= \frac{1}{N}\sum_i f_i d_i^2 + (A - \bar{x})^2 + 2(A - \bar{x}) \cdot \frac{1}{N}\sum_i f_i d_i$$

We know that if $d_i = x_i - A$ then $\bar{x} = A + \frac{1}{N}\sum_i f_i d_i$

$$\therefore \qquad A - \bar{x} = -\frac{1}{N}\sum_i f_i d_i$$

Hence

$$\sigma_x^2 = \frac{1}{N}\sum f_i d_i^2 + \left(-\frac{1}{N}\sum_i f_i d_i\right)^2 + 2\left(-\frac{1}{N}\sum_i f_i d_i\right)\left(\frac{1}{N}\sum_i f_i d_i\right)$$

$$= \frac{1}{N}\sum_i f_i d_i^2 - \left(\frac{1}{N}\sum_i f_i d_i\right)^2$$

$$\Rightarrow \qquad \sigma_x^2 = \sigma_d^2$$

variance is not independent of change of scale.

If $d_i = \frac{x_i - A}{h}$, then

$$x_i = A + hd_i \quad \text{and} \quad \bar{x} = A + h \cdot \frac{1}{N} \sum_i f_i d_i = A + h\bar{d}$$

Obviously $$x_i - \bar{x} = h(d_i - \bar{d})$$

$$\therefore \quad \sigma_x^2 = \frac{1}{N} \sum_i f_i (x_i - \bar{x})^2 = h^2 \cdot \frac{1}{N} \sum_i f_i (d_i - \bar{d})^2 = h^2 \sigma_d^2$$

$$\sigma_x^2 = h^2 \sigma_d^2$$

Dr. HARINAKSHI

The variance σ^2 of the combined series

If n_1, n_2 are the sizes,

$\bar{x}_1, \bar{x}_2$ means

σ_1, σ_2 S.D of

The combined variance

$$\sigma^2 = \frac{1}{n_1+n_2}\left[n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)\right]$$

$d_1 = \bar{x}_1 - \bar{x}$

$d_2 = \bar{x}_2 - \bar{x}$

$\bar{x} = \dfrac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}$, combined mean

Dr. HARINAKSHI KARKERA

$$\boxed{\sigma^2 = \frac{1}{n_1+n_2}\left[n_1(\sigma_1^2+d_1^2)+n_2(\sigma_2^2+d_2^2)\right]}$$

$$\boxed{\sigma^2 = \frac{1}{n_1+n_2}\left[n_1\sigma_1^2+n_2\sigma_2^2+\frac{n_1n_2}{n_1+n_2}(\bar{x}_1-\bar{x}_2)^2\right]}$$

Generalized $\quad n_1, n_2, \ldots, n_k$

$\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$

$$\bar{x} = \frac{n_1\bar{x}_1+n_2\bar{x}_2+\cdots+n_k\bar{x}_k}{n_1+n_2+\cdots+n_k}$$

$$\sigma^2 = \frac{1}{n_1+n_2+\cdots+n_k}\left[n_1(\sigma_1^2+d_1^2)+n_2(\sigma_2^2+d_2^2)+\cdots+n_k(\sigma_k^2+d_k^2)\right]$$

$d_i = \bar{x}_i - \bar{x}$

Dr. HARINAKSHI KARKERA