Disperssion:— (15-20 marks) (Variability) => (consistency)

The values of a variable are generally ~~are~~ not all equal. In some cases the values are varry close to one another. Again in some cases, they are far apart from each other. In order to get a proper idea about the overall nature of a given set of values, it is necessary to know the ~~extend~~ extent to which the values differ among themselves.

Scatterdness:— That is we want to know how scatterd the values are about the average. This feature of a freq. distribution is called dispersion.

Measure of dispersion:—

- Absolute measure (has unit)
  - Range
  - Standard deviation (S.D) — Imp
  - Mean deviation (M.D) — Imp
  - Quartile deviation (Q.D)
- Relative measure (no unit)
  - co-efficient of M.D
  - co-efficient of Q.D
  - co-efficient of variation (C.V) — Imp

<u>Range:</u>— Let, $x_1, x_2, \cdots, x_n$ be n obs" on a variable x.

Then $\boxed{\text{range} = \max x_i - \min x_i}$

~~Where~~ when the value of range $\to 0 \Rightarrow$ dispersion is less

Q. When does the dispersion is equal to 0 ?

Ans. dispersion $= 0 \Rightarrow$ range $= 0 \Rightarrow \max x_i = \min x_i$

$\Rightarrow$ all values of obs.. are same, which is quite impossible.

ii) <u>S.D (σ)</u>

<u>definition:</u>— Let, the variable x assums values $x_1, x_2, \cdots, x_n$, then $\boxed{\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}}$

deviation/variability $= x_i - \bar{x}$

deviation $\propto \sigma$

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

<u>Note:</u>—

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

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

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 + \frac{1}{n} \sum_{i=1}^{n} \bar{x}^2 - \frac{2}{n} \sum_{i=1}^{n} x_i \cdot \bar{x} \quad [\because \Sigma K = n \cdot k]$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 + \frac{1}{\cancel{n}} \times \cancel{n} \bar{x}^2 - \frac{2}{n} \bar{x} \sum_{i=1}^{n} x_i \quad \left[\frac{1}{n} \Sigma x_i = \bar{x}\right]$$

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

$$\Rightarrow \boxed{\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2}$$

when $\sigma^2 = 0$, $\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = 0$

$\Rightarrow x_i = \bar{x}$

$\therefore x_1 = \bar{x}$

$x_2 = \bar{x}$

$x_n = \bar{x}$, which is practically impossible.

Imp. Q. Prove or disprove : $\sum_{i=1}^{50} |i-25.1| = \sum_{i=1}^{50} |i-25.2|$ (3 marks)

Ans. L.H.S $\sum_{i=1}^{50} |i-25.1| = \sum_{i=1}^{25} -(i-25.1) + \sum_{i=26}^{50} (i-25.1)$

$$= \sum_{i=1}^{25} (25.1-i) + \sum_{i=26}^{50} (i-25.1)$$

$|x-a| = x-a, \; x \geq a$

$|x-a| = -(x-a), \; x<a$

$$= \sum_{i=1}^{25} 25.1 - \sum_{i=1}^{25} i + \sum_{i=26}^{50} i - \sum_{i=26}^{50} 25.1$$

$$= 25 \times 25.1 - \sum_{i=1}^{25} i + \sum_{i=26}^{50} i - 25 \times 25.1$$

$$= -\sum_{i=1}^{25} i + \sum_{i=26}^{50} i \quad ----- (i)$$

R.H.S $\sum_{i=1}^{50} |i-25.2|$

$$= \sum_{i=1}^{25} -(i-25.2) + \sum_{i=26}^{50} (i-25.2)$$

$$= \sum_{i=1}^{25} 25.2 - \sum_{i=1}^{25} i + \sum_{i=26}^{50} i - \sum_{i=26}^{50} 25.2$$

$$= 25 \times 25.2 - \sum_{i=1}^{25} i + \sum_{i=26}^{50} i - 25 \times 25.2$$

$$= -\sum_{i=1}^{25} i + \sum_{i=26}^{50} i \quad ——— (ii)$$

~~Now, (i) = (ii)~~ $\therefore$ L.H.S = R.H.S [Proved.]

Property of varience:

i) $var(X) = 0$ $\quad [var(X) = \sigma^2]$

$\Rightarrow$ all obsns are equal

Imp. ii) $y = a + bx \Rightarrow var(Y) = b^2 var(X)$

(a: origin, b: scale)

$y_i = a + bx_i, \; i = 1, 2, \ldots, n$

Pf:- Let, $y_i = a + bx_i$

$\Rightarrow \sum_{i=1}^{n} y_i = \cancel{a + b\sum_{i=1} x_i}\ \sum_{i=1}^{n} (a + bx_i)$

$$\Rightarrow \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} a + \sum_{i=1}^{n} b x_i$$

$$\Rightarrow \sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} y_i = a + \frac{b}{n}\sum_{i=1}^{n} x_i$$

$$\Rightarrow \bar{y} = a + b\bar{x}$$

$$\therefore y_i - \bar{y} = a + bx_i - (a + b\bar{x})$$

$$= \cancel{a} + bx_i - \cancel{a} - b\bar{x}$$

$$= b(x_i - \bar{x})$$

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

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

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

$$= b^2\, var(x)$$

$\therefore var(y) = b^2\, var(x)$ [Proved]

Q. Find the S.D of 1, 3, 5, 7, ..., 25. (C.U) (3 marks)

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

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

$n = 13$

$$\sum_{i=1}^{n} x_i = 1 + 3 + 5 + \cdots + 25$$

$$= \frac{13}{2}(1 + 25) = 169$$

Sum of 1st odd N $= \frac{n}{2}(a + l)$

$$\therefore \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{169}{13} = 13$$

$\Sigma x_i^2 = 1^2+3^2+5^2+\cdots+25^2$

$= \left(1^2+2^2+3^2+4^2+\cdots+25^2+26^2\right) - \left(2^2+4^2+\cdots+26^2\right)$

$= \dfrac{\overset{13}{\cancel{26}}\times\overset{9}{\cancel{27}}\times 53}{\cancel{6}_2} - 2^2\left(1+2^2+3^2+\cdots+13^2\right)$

$= 6201 - 2^2\times\dfrac{13\times\overset{7}{\cancel{14}}\times\overset{9}{\cancel{27}}}{\cancel{6}_2}$

$\boxed{\text{Sum of sq. of 1st } N = \dfrac{n(n+1)(2n+1)}{6}}$

$= 2925$

$\therefore\ \sigma^2 = \frac{1}{13}\times 2925 - 13^2$

$= 225-169 = 56$

$\frac{25\times51\times49}{3}$

$\therefore\ S.D = \sqrt{56}$ Ans.

H.W For 10 values of x, $\Sigma u = 4$, $\Sigma u^2 = 144$, $u = \dfrac{x-10}{5}$. Find Σx^2.

Ans. $u = \dfrac{x-10}{5}$

$\Rightarrow u^2 = \frac{1}{25}\left(x^2+100-20x\right)$

$\Rightarrow \Sigma u_i^2 = \frac{1}{25}\left(\Sigma x_i^2 + 100\times 10 - 20\Sigma x_i\right)$

Now, $\Sigma u_i = 4 \Rightarrow \Sigma\left(\dfrac{x_i-10}{5}\right) = 4$

$\Rightarrow \Sigma x_i - 10\times 10 = 20$

$\Rightarrow \Sigma x_i = 120$

$\therefore\ \Sigma u_i^2 = \frac{1}{25}\left(\Sigma x_i^2 + 100\times 10 - 20\times 120\right)$

$\Rightarrow 144\times 25 = \Sigma x_i^2 + 1000 - 2400$

$\Rightarrow \Sigma x_i^2 = 5000$

Mean deviation:-

Let, a variable x assumes the values $x_1, x_2, \cdots, x_n$ & c be a origin (central value).

Then $M.D_c = \frac{1}{n}\sum_{i=1}^{n}|x_i - c|$

Note: (i) $c = \bar{x}$

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

ii) $c = \hat{x}$

$$M.D_{\hat{x}} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{x}|$$

[$\hat{x}$ = median] $\downarrow$ x curl

Result

$$S.D \geq M.D_{\bar{x}}$$

Pf:- Let, $y_1, y_2, \cdots, y_n$ be n obs^{n}s of a variable y.

Then $\sum_{i=1}^{n}(y_i - \bar{y})^2 \geq 0$

$$\Rightarrow \sum_{i=1}^{n}(y_i^2 - 2y_i\bar{y} + \bar{y}^2) \geq 0$$

$$\Rightarrow \sum_{i=1}^{n} y_i^2 - 2\bar{y}\sum_{i=1}^{n} y_i + \sum_{i=1}^{n}\bar{y}^2 \geq 0$$

$$\Rightarrow \sum_{i=1}^{n} y_i^2 - 2\bar{y}\cdot n\bar{y} + n\bar{y}^2 \geq 0$$

[$\bar{y} = \frac{1}{n}\Sigma x_i$]

$$\Rightarrow \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 \geq 0$$

$$\Rightarrow \frac{\sum_{i=1}^{n} y_i^2}{n} \geq \left(\frac{1}{n}\sum_{i=1}^{n} y_i\right)^2 \quad \text{--- (1)}$$

In eq^n (1), put $y_i = |x_i - \bar{x}| \;\forall\; i$

$$\boxed{\begin{array}{l} \text{i) } |x|^2 = x^2 \\ \text{ii) } |-x| = |x| \end{array}}$$

$$\therefore \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|^2 \geq \left(\frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|\right)^2$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \geq \left(\frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|\right)^2$$

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

[taking +ve square root on both sides]

$\Rightarrow$ S.d $\geq M.D_{\bar{x}}$

Q. <u>Case of equality</u> :- when S.D = M.D ?

Ans: Equality holds when all the obsⁿs are equal or when the variable assums only two values with equal freqⁿ.

Imp. Result :- <u>M.D about median is least.</u> (6 marks)

Imp.

<u>P.f</u> :- To prove this result we shall use the following result.

If P and Q be two fixed points on a line and R be a variable point on the same line, then the sum of ~~the~~ distances of R from P and Q is least when R lies b/w P & Q.

Pf of <u>main parts</u>

Let, $x_1, x_2, \cdots, x_n$ are n obsⁿs of a variable x.

Then $M.D_c = \frac{1}{n}\sum_{i=1}^{n}|x_i - c|$

<u>case I</u>: Let, n be odd. Let, $n = 2m+1$

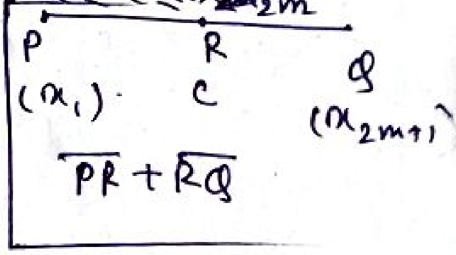
without loss of generility (WOLG), assume that,

$x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_m \leq x_{m+1} \leq \cdots \leq x_{2m} \leq x_{2m+1}$

Then using the above result we have,

$|x_1 - c| + |x_{2m+1} - c|$ is least when $x_1 \leq c \leq x_{2m+1}$

$|x_2 - c| + |x_{2m} - c|$ is " " $x_2 \leq c \leq x_{2m}$



$\vdots$

$|x_{m+1} - c|$ is least when $x_{m+1} = c$

So, we see that when $c = x_{m+1}$ it lies b/w the pair of values in each of the above sums.

& hence $M.D_c = \frac{1}{n}\sum|x_i - c|$ is least when $c = x_{m+1}$ which ~~x_{m+1}~~ is ~~our~~ the median in this case.

Case (II) :- n be even. Let, $n = 2K$

WOLG, assume that,

$x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_K \leq x_{K+1} \leq \cdots \cdots \leq x_{2K}$

$|x_1 - c| + |x_{2K} - c|$ is least $x_1 \leq c \leq x_{2K}$

$|x_2 - c| + |x_{2K-1} - c|$ " " $x_2 \leq c \leq x_{2K-1}$

$\vdots \qquad\qquad \vdots$

$|x_K - c| + |x_{K+1} - c|$ " " $x_K \leq c \leq x_{K+1}$

c lies b/w x_K & x_{K+1} if also lies b/w the pair of values ~~is~~ in each of the above sum and we know that median alaso lies b/w x_K & x_{K+1}.

So, $M.D_c = \frac{1}{n}\sum|x_i - c|$ is least then $c =$ med.

Q. For two values say a & b, $a < b$ mean $= 25$, $S.D = 4$, find a & b. (C.U) (3 marks)

Ans. Given, $\bar{x} = 25$ —(i), $\sigma = 4 \Rightarrow \sigma^2 = 16$ —(ii)

From (i), $\frac{a+b}{2} = 25 \Rightarrow a + b = 50$ —— (iii)

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

From (ii), $\frac{1}{2}\left[\left(a - \frac{a+b}{2}\right)^2 + \left(b - \frac{a+b}{2}\right)^2\right] = 16$

$\Rightarrow \frac{1}{2}\left[\left(\frac{a-b}{2}\right)^2 + \left(\frac{b-a}{2}\right)^2\right] = 16$

$\Rightarrow \frac{1}{2} \times 2 \times \left(\frac{a-b}{2}\right)^2 = 16$

$\Rightarrow \frac{1}{\cancel{2}} \times \cancel{2} \times \frac{(b-a)^2}{4} = 16 \Rightarrow \frac{b-a}{2} = 4$ ~~[b>a]~~

$\Rightarrow b - a = 8$ —(iv) $[\because b > a]$

(iii) + (iv)

$\cancel{a}+b+b-\cancel{a} = 50+8 \Rightarrow 2b = 58$

$\Rightarrow b = 29$

$\therefore a = 21$.

Imp H/W

Q. Imp: Let, x takes +ve values and $x_i - \bar{x}$ is small as compared to $\bar{x}$. S.T → a) $x_g \overset{\text{nearly equal}}{\simeq} \bar{x}\left(1 - \frac{1}{2}\frac{s^2}{\bar{x}^2}\right)$

s = variance

b) $x_h \simeq \bar{x}\left(1 - \frac{s^2}{\bar{x}^2}\right)$

(5+5 = 10 marks)

Ans: a) $x_g = (x_1 \cdot x_2 \cdot \ldots \cdot x_n)^{1/n}$

$\Rightarrow \log x_g = \frac{1}{n}[\log x_1 + \log x_2 + \cdots + \log x_n]$

$= \frac{1}{n}\sum_{i=1}^{n} \log x_n$

If $\frac{a}{b}$ is small then $\left(\frac{a}{b}\right)^K$ is very small.

$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$ → log-series

Let, $\boxed{x_i - \bar{x} = x_i'}$

$\therefore \log x_g = \frac{1}{n}\sum_{i=1}^{n} \log(x_i' + \bar{x})$

$= \frac{1}{n}\sum_{i=1}^{n} \log\left[\bar{x}\left(1 + \frac{x_i'}{\bar{x}}\right)\right]$

$= \frac{1}{n}\sum_{i=1}^{n}\left[\log \bar{x} + \log\left(1 + \frac{x_i'}{\bar{x}}\right)\right]$

$= \frac{1}{n}\sum_{i=1}^{n} \log \bar{x} + \frac{1}{n}\sum_{i=1}^{n} \log\left(1 + \frac{x_i'}{\bar{x}}\right)$

$= \frac{1}{n} \times n \log \bar{x} + \frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_i'}{\bar{x}} - \frac{\left(\frac{x_i'}{\bar{x}}\right)^2}{2} + \frac{\left(\frac{x_i'}{\bar{x}}\right)^3}{3} \cdots\right]$

$= \log \bar{x} + \frac{1}{n}\left[\sum_{i=1}^{n} \frac{x_i'}{\bar{x}} - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i'}{\bar{x}}\right)^2 + \frac{1}{3}\sum_{i=1}^{n}\left(\frac{x_i'}{\bar{x}}\right)^3 \cdots\right]$

$\simeq \log \bar{x} + \frac{1}{n}\left[0 - \frac{1}{2}\frac{ns^2}{\bar{x}^2}\right]$

$\left[\because \sum_{i=1}^{n}\left(\frac{x_i'}{\bar{x}}\right) = \frac{1}{\bar{x}} \sum_{i=1}^{n}(x_i - \bar{x}) = 0\right.$

$\sum_{i=1}^{n}\left(\frac{x_i'}{\bar{x}}\right)^2 = \frac{1}{\bar{x}^2} \sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{\bar{x}^2} n . s^2 \quad \left[\because s^2 = \frac{1}{n}\sum(x_i - \bar{x})^2\right.$

$\sum_{i=1}^{n}\left(\frac{x_i'}{\bar{x}}\right)^3 = \sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{\bar{x}}\right)^3$

Since $(x_i - \bar{x})$ is small compared to $\bar{x}$. Therefore, we can assume $\left(\frac{x_i - \bar{x}}{\bar{x}}\right)^K \to 0 \;\; \forall\, K \geq 3$

$\therefore \log x_g = \log \bar{x} - \frac{s^2}{2\bar{x}^2}$

$\Rightarrow \log\left(\frac{x_g}{\bar{x}}\right) = -\frac{s^2}{2\bar{x}^2}$

$\Rightarrow \frac{x_g}{\bar{x}} = e^{-s^2/2\bar{x}^2}$

$\boxed{e^{-x} = 1 - \frac{x}{1!} + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots}$

$\Rightarrow x_g = \bar{x} e^{-s^2/2\bar{x}^2}$

$\Rightarrow x_g = \bar{x}\left[1 - \frac{s^2}{2\bar{x}^2} + \frac{\left(\frac{s^2}{2\bar{x}^2}\right)^2}{2!} - \cdots\cdots\right]$

$\Rightarrow x_g = \bar{x}\left[1 - \frac{s^2}{2\bar{x}^2} + \frac{s^4}{4\bar{x}^4} - \cdots\cdots\right]$

Now, $\frac{s^4}{\bar{x}^4} = \frac{(s^2)^2}{\bar{x}^4} = \left\{\frac{1}{n}\frac{\sum(x_i - \bar{x})^2}{\bar{x}^2}\right\}^2 \to 0$ [as our assumption]

$\therefore x_g \simeq \bar{x}\left[1 - \frac{s^2}{2\bar{x}^2}\right]$ [Proved.]

S.D of combined group :-

	Groups: I	II	combined
obs^{n}s -	n_1	n_2	$N = n_1 + n_2$
mean:	$\bar{x}_1$	$\bar{x}_2$	$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}$
s.d :-	σ_1	σ_2	σ

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

Pf :- Let, the obs^{n}s in the first group are $x_{11}, x_{12}, \cdots, x_{1n_1}$.

And " " " " second " " $x_{21}, x_{22}, \cdots, x_{2n_2}$.

Then $\bar{x}_1 = \frac{x_{11} + x_{12} + \cdots + x_{1n_1}}{n_1}$

$$\Rightarrow \bar{x}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} x_{1i}$$

& $\bar{x}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} x_{2i}$

$$\sigma_1^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(x_{1i} - \bar{x}_1)^2$$

& $\sigma_2^2 = \frac{1}{n_2}\sum_{i=2}^{n_2}(x_{2i} - \bar{x}_2)^2$

$$\sum_{i=1}^{n_1}(x_{1i} - \bar{x})^2 = \sum_{i=1}^{n_1}\{(x_{1i} - \bar{x}_1) + (\bar{x}_1 - \bar{x})\}^2$$

$$= \sum_{i=1}^{n_1}(x_{1i} - \bar{x}_1)^2 + \sum_{i=1}^{n_1}(\bar{x}_1 - \bar{x})^2 + 2\sum_{i=1}^{n_1}(x_{1i} - \bar{x})(\bar{x}_1 - \bar{x})$$

$$\Rightarrow \sum_{i=1}^{n_1}(x_{1i} - \bar{x})^2 = n_1\sigma_1^2 + n_1(\bar{x}_1 - \bar{x})^2 + 2(\bar{x}_1 - \bar{x})\sum_{i=1}^{n_1}(x_{1i} - \bar{x})$$

$$= n_1\sigma_1^2 + n_1(\bar{x}_1 - \bar{x})^2$$

Similarly,

$$\sum_{i=1}^{n_2}(x_{2i} - \bar{x})^2 = n_2\sigma_2^2 + n_2(\bar{x}_2 - \bar{x})^2$$

Therefore, combined varience $(\sigma^2) = \frac{1}{n_1+n_2}\left[\sum_{i=1}^{n_1}(x_{1i}-\bar{x})^2 + \sum_{i=1}^{n_2}(x_{2i}-\bar{x})^2\right]$

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

[Proved]

H/W

1) Let, a variable x assumes only 2 values x_1 & x_2 with equal frequency. Find mean deviation about mean $(M.D_{\bar{x}})$ & s.d.

2) Calculate minimum value of M.D of a variable that assumes the values 2, 3, 4, ..., 16.

1) Let, x_1 & x_2 will be obsⁿs with corresponding frequencies f_1 & f_2.

Now acc. to ques., $f_1 = f_2 = f$ (say)

$$\therefore \bar{x} = \frac{x_1 f + x_2 f}{f+f} = \frac{1}{2}(x_1+x_2)$$

$$\therefore M.D_{\bar{x}} = \frac{1}{n}\sum_{i=1}^{n}|x_i-\bar{x}|$$

$$= \frac{1}{2}\sum_{i=1}^{2}|x_i-\bar{x}|$$

$$= \frac{1}{2}\left[|x_1-\bar{x}| + |x_2-\bar{x}|\right]$$

$$= \frac{1}{2}\left[\left|x_1 - \frac{x_1+x_2}{2}\right| + \left|x_2 - \frac{x_1+x_2}{2}\right|\right]$$

$$= \frac{1}{2}\times\left[\left|\frac{x_1-x_2}{2}\right| + \left|\frac{x_2-x_1}{2}\right|\right]$$

$$= \frac{1}{2}\times\left[\left|\frac{x_1-x_2}{2}\right| + \left|\frac{x_1-x_2}{2}\right|\right]$$

$$= \frac{|x_1-x_2|}{2}$$

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

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

$$= \sqrt{\frac{1}{2} \left[(x_i - \bar{x})^2 + (x_2 - \bar{x})^2\right]}$$

$$= \sqrt{\frac{1}{2} \times \left[\left(\frac{x_1 - x_2}{2}\right)^2 + \left(\frac{x_2 - x_1}{2}\right)^2\right]}$$

$$= \sqrt{\frac{1}{\cancel{2}} \times \frac{1}{4} \times \cancel{2}(x_1 - x_2)^2}$$

$$= \frac{|(x_1 - x_2)|}{2}$$

Ans

2) The values are 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16

$n = 15$

$\tilde{x} = 9$

$$\therefore M.D_{\tilde{x}} = \frac{(9-2)+(9-3)+(9-4)+(9-5)+(9-6)+(9-7)+(9-8)+\cdots+(16-9)}{15}$$

$$= \frac{7+6+5+4+3+2+1+0+1+2+3+4+5+6+7}{15}$$

$$= \frac{2\times(1+2+\cdots+7)}{15}$$

$$= \frac{2\times 7\times 8}{15\times 2} = 3.733$$

Q. S.T $\rightarrow$ $$\sigma^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1+n_2} + \frac{n_1 n_2(\bar{x}_1 - \bar{x}_2)^2}{(n_1+n_2)^2}$$

Pf:- $$\sigma^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2 + n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2}{n_1+n_2}$$ [Reqd. to prove]

$$\Rightarrow \sigma^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1+n_2} + \frac{n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2}{n_1+n_2}$$

We have,

$$n_1(\bar{x}_1 - \bar{x})^2 = n_1\left(\bar{x}_1 - \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}\right)^2$$

$$= n_1\left(\frac{n_2\bar{x}_1 - n_2\bar{x}_2}{n_1+n_2}\right)^2$$

$$= \frac{n_1 n_2^2(\bar{x}_1 - \bar{x}_2)^2}{(n_1+n_2)^2}$$

Similarly, $$n_2(\bar{x}_2 - \bar{x})^2 = \frac{n_2 n_1^2(\bar{x}_1 - \bar{x}_2)^2}{(n_1+n_2)^2}$$

$$\therefore \sigma^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1+n_2} + \frac{\frac{n_1 n_2^2(\bar{x}_1 - \bar{x}_2)^2}{(n_1+n_2)^2} + \frac{n_2 n_1^2(\bar{x}_1 - \bar{x}_2)^2}{(n_1+n_2)^2}}{n_1+n_2}$$

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

Imp.

$$= \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1+n_2} + \frac{n_1 n_2(\bar{x}_1 - \bar{x}_2)^2}{(n_1+n_2)^2} \quad \text{[Showed]}$$

Imp. Let, R be the range, S be s.d. S.T $\to \underset{\text{lower bound}}{\frac{R^2}{2n}} \leq S^2 \leq \underset{\text{upper bound}}{\frac{R^2}{4}}$

Imp. Imp. pf:- Let, a variable x assumes the values $x_1, x_2, \ldots, x_n$. Let, a be the smallest ~~lowest~~ and b be the greatest value.

$\therefore$ Range $= b-a$

Result:- $\sum_{i=1}^{n}(x_i - c)^2$ is least when $c = \bar{x}$

We have, $\sum_{i=1}^{n}(x_i - \bar{x})^2 \leq \sum_{i=1}^{n}\left(x_i - \frac{a+b}{2}\right)^2 \quad \ldots\ldots (1)$

Now, $\sum_{i=1}^{n}\left(x_i - \frac{a+b}{2}\right)^2 = \Sigma_1\left(x_i - \frac{a+b}{2}\right)^2 + \Sigma_2\left(x_i - \frac{a+b}{2}\right)^2$,

where Σ_1 include those values of x which are less than equal ~~equal~~ to $\frac{a+b}{2}$ and Σ_2 include those values of x ~~exit~~ which are $\geq \frac{a+b}{2}$.

$$\therefore \sum_{i=1}^{n}\left(x_i - \frac{a+b}{2}\right)^2 \leq \Sigma_1\left(a - \frac{a+b}{2}\right)^2 + \Sigma_2\left(b - \frac{a+b}{2}\right)^2$$

$$= \Sigma_1\left(\frac{a-b}{2}\right)^2 + \Sigma_2\left(\frac{b-a}{2}\right)^2$$

$$= \Sigma_1\left(\frac{R}{2}\right)^2 + \Sigma_2\left(\frac{R}{2}\right)^2$$

$$\boxed{b-a = \text{range} = R}$$

$$= \left(\frac{R}{2}\right)^2 \times n$$

$$\left(\frac{R}{2}\right)^2(\Sigma_1 + \Sigma_2) = \left(\frac{R}{2}\right)^2 \times n$$

Therefore, $\sum_{i=1}^{n}\left(x_i - \frac{a+b}{2}\right)^2 \leq n\left(\frac{R}{2}\right)^2 \quad \ldots\ldots (2)$

From (1) & (2),

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \leq n \times \frac{R^2}{4}$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 \leq \frac{R^2}{4}.$$

$$\Rightarrow \quad s^2 \leq \frac{R^2}{4} \quad \text{------} \quad (A)$$

We have, $\quad ns^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2$

$$= (a-\bar{x})^2 + (b-\bar{x})^2 + \Sigma_1 (x_i-\bar{x})^2$$

here Σ_1 includes all values of x except the lowest & highest value.

$$\therefore \; ns^2 \geq (a-\bar{x})^2 + (b-\bar{x})^2 \quad \left[\because \Sigma_1 (x_i-\bar{x})^2 \geq 0\right] \quad \text{-----} \quad (iii)$$

Now, $(a-\bar{x})^2 + (b-\bar{x})^2$

$$= \frac{\left[\{(a-\bar{x})+(b-\bar{x})\}^2 + \{(a-\bar{x})-(b-\bar{x})\}^2\right]}{2}$$

$$= \frac{1}{2}\left[(a-\bar{x}+b-\bar{x})^2 + (a-\bar{x}-b+\bar{x})^2\right]$$

$$= \left[\frac{(a+b-2\bar{x})^2}{2} + \frac{(a-b)^2}{2}\right]$$

$$\therefore \; (a-\bar{x})^2 + (b-\bar{x})^2 \geq \frac{(a-b)^2}{2} \quad \left[\because (a+b-2\bar{x})^2 \geq 0\right]$$

$$\Rightarrow (a-\bar{x})^2 + (b-\bar{x})^2 \geq \frac{R^2}{2} \quad \text{----} \quad (iv)$$

$\therefore$ From (iii) & (iv),

$$ns^2 \geq \frac{R^2}{2} \Rightarrow s^2 \geq \frac{R^2}{2n} \quad \text{----} \quad (B)$$

$\therefore$ From (A) & (B) we get,

$$\frac{R^2}{2n} \leq s^2 \leq \frac{R^2}{4} \quad \text{[Showed.]}$$

Quartile deviation: It's a measure of dispersion based on quartiles.

We know that if the ~~vari~~ values of a variable differ much from one another the difference b/w their quartiles would be large.

Again, when the values are close to one another the difference b/w their quartiles ~~will~~ would be small.

So, difference b/w the quartiles can be taken as a measure of dispersion, given by, $Q = \frac{(Q_2 - Q_1) + (Q_3 - Q_2)}{2}$

The above measure is called Quartile deviation (Q.D)

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

Note:- The numerator $(Q_3 - Q_1)$ is basically called interquartile range.

Some imp. result:-

1) Let, $y = a + bx$ then $\boxed{var(y) = b^2 var(x)}$

Pf:- $y_i = a + bx_i \quad \forall i$

$$\therefore \bar{y} = a + b\bar{x}$$

$$\therefore y_i - \bar{y} = b(x_i - \bar{x})$$

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

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

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

$$= b^2 var(x)$$ [Proved]

H/W

1) A variable assumes only two distinct values 0 & 1 which freq. $1/3$ & $2/3$. Find mean & varience.

Ans We ~~to~~ have,

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2}{f_1 + f_2}$$

Here, $x_1 = 0$, $x_2 = 1$, $f_1 = 1/3$, $f_2 = 2/3$

$$\therefore \bar{x} = \frac{0 \times 1/3 + 1 \times 2/3}{1/3 + 2/3} = 2/3$$

Again,

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

$$= \frac{1}{2} \left[(0 - 2/3)^2 + (1 - 2/3)^2 \right]$$

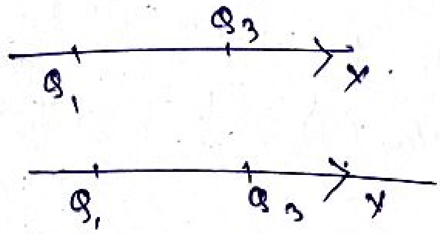
$$= \frac{1}{2} \times \left[\frac{4}{9} + \frac{1}{9} \right]$$

$$= \frac{1}{2} \times \frac{5}{9} = 5/18$$ Ans.

Imp Result:

$y = a + bx$, then $Q.D(y) = |b| \, Q.D(x)$

Pf: case I: $b > 0$.



$\therefore y = a + bx$

$\Rightarrow Q_3(y) = a + b \cdot Q_3(x)$

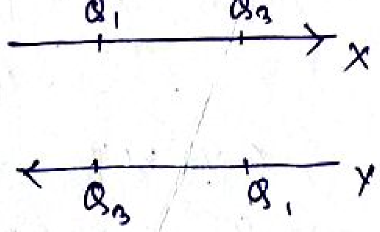
$Q_1(y) = a + b Q_1(x)$

$Q_3(y) - Q_1(y) = b\{Q_3(x) - Q_1(x)\}$

$$\Rightarrow \frac{Q_3(y) - Q_1(y)}{2} = \frac{b\{Q_3(x) - Q_1(x)\}}{2}$$

$\Rightarrow Q.D(y) = b \cdot Q.D(x)$ ----(I)

Case II :- $b<0$



$y = a+bx$

$\Rightarrow Q_3(y) = a + b\, Q_1(x)$

$Q_1(y) = a + b\, Q_3(x)$

$$\Rightarrow \frac{Q_3(y) - Q_1(y)}{2} = b\left\{\frac{Q_1(x) - Q_3(x)}{2}\right\}$$

$\Rightarrow Q.D(y) = -b\, Q.D(x)$ —— (II)

Combining eq.^n (I) & (II),

$$\boxed{Q.D(y) = |b|\, Q.D(x)}$$ [Proved.]

Imp. Imp. Imp. **Result** :- The difference b/w mean & median can't exceed S.D. $(|\bar{x} - \tilde{x}|) \leq S$

Ans- Pf:

$$|\bar{x} - \tilde{x}| = \left|\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{n}\sum_{i=1}^{n}\tilde{x}\right| = \left|\frac{1}{n}\sum_{i=1}^{n}(x_i - \tilde{x})\right|$$

$$= \frac{1}{n}\left|\sum_{i=1}^{n}(x_i - \tilde{x})\right|$$

$$\boxed{\left|\sum y_i\right| \leq \sum |y_i|}$$

$$\leq \frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}|$$

$$\leq \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}| \quad [\because M.D_{\tilde{x}} \text{ is least}]$$

$$\leq \sigma \quad [\because M.D_{\bar{x}} \leq \sigma]$$

$$\therefore |\bar{x} - \tilde{x}| \leq \sigma$$

Co-efficient of variation (C.V) : It is a measure of relative dispersion.

C.V is given by, $C.V = \dfrac{\sigma}{\bar{x}} \times 100\%$.

Q. A.M = 10, C.V = 50%. Find V(5-2x).

Sol^n:

$$Var(5-2x) = (-2)^2 var(x) \quad [\because Var(a+bx) = b^2 var(x)]$$
$$= 4\, var(x) \quad ----(1)$$

C.V = 50%

$$\Rightarrow \frac{\sigma}{\bar{x}} \times 100 = 50$$

$$\Rightarrow \frac{\sigma}{10} \times 100 = 50 \Rightarrow \sigma^2 = 25$$

$\therefore$ From (1), $Var(5-2x) = 4 \times 25 = 100$ Ans

Q. A variable ~~assu~~ assume the values a & b & (n-2) other values all equal to $\frac{a+b}{2}$. Find S.D, M.D$\bar{x}$. (2+2) (C.U)

Ans ~~Calcu~~ Calculation of S.D

x_i	f_i	$x_i f_i$
a	1	a
b	1	b
$\frac{a+b}{2}$	$(n-2)$	$\left(\frac{a+b}{2}\right)(n-2)$
Total	n	$\frac{n}{2}(a+b)$

$$(a+b) + \frac{a+b}{2}(n-2)$$
$$(a+b)\left(1+\frac{n-2}{2}\right)$$
$$\frac{2+n-2}{2}$$

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{\frac{n}{2}(a+b)}{n} = \frac{a+b}{2}$$

$$Var(x) = \frac{1}{N}\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot f_i \quad , N = \sum f_i$$

$$= \frac{1}{n}\left[\left(a-\frac{a+b}{2}\right)^2 \cdot 1 + \left(b-\frac{a+b}{2}\right)^2 \cdot 1 + \left(\frac{a+b}{2}-\frac{a+b}{2}\right)^2 \cdot (n-2)\right]$$

$$= \frac{1}{n}\left[\left(\frac{a-b}{2}\right)^2 + \left(\frac{b-a}{2}\right)^2\right]$$

$$\Rightarrow var(x) = \frac{1}{n} \times \frac{(a-b)^2}{2}$$

$$\therefore S.d(x) = \frac{|(a-b)|}{\sqrt{2n}}$$

~~Q. Calculate~~

$$M.D_{\bar{x}} = \frac{1}{N}\sum_{i=1}^{n} |x_i - \bar{x}| f_i \quad [N = \Sigma f_i]$$

$$= \frac{1}{n}\left[\left|a - \frac{a+b}{2}\right|.1 + \left|b - \frac{a+b}{2}\right| . \left|\frac{a+b}{2} - \frac{a+b}{2}\right| (n-2)\right]$$

$$= \frac{1}{n}\left[\left|\frac{a-b}{2}\right| + \left|\frac{b-a}{2}\right|\right]$$

$$= \frac{1}{n} 2(a-b)$$ Ans.

Q. Find the min. value of M.D of a variable that assums the values 2, 3, 4, ... , 16.

Sol- We know that,

$M.D_{\tilde{x}}$ is least.

So to get min. value we shall calculate $M.D_{\tilde{x}}$.

Here, $n = 15$, which is odd.

$$\therefore \tilde{x} = \left(\frac{n+1}{2}\right)^{th} \text{ obs}^{n}s$$

$$= \left(\frac{16}{2}\right)^{th} \quad "$$

$$= 8^{th} \quad "$$

$$= 9$$

Calculation of $M.D_{\tilde{x}}$

x_i	$x_i - \tilde{x}$	$\lvert x_i - \tilde{x}\rvert$
2	-7	7
3	-6	6
4	-5	5
5	-4	4
6	-3	3
7	-2	2
8	-1	1
9	0	0
10	1	1
11	2	2
12	3	3
13	4	4
14	5	5
15	6	6
16	7	7
Total	0	56

$\therefore M.D_{\tilde{x}} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \tilde{x}|$

$= \frac{56}{15} = 3.733$

1) A variable assumes the values 1, 2, ..., n with corresponding freq^n 1, 2, ..., n. Find S.d.

2) Let, $y = 2 - 3x$ be the relation b/w x & y. If Q.D(x) is 1, Find Q.D(y).

2) $y = 2 - 3x \Rightarrow y = 2 + (-3)x$

$\therefore Q.D(y) = |-3| \, Q.D(x)$

$= 3 \times 1 = 3$ Ans

1) $\sigma^2 = \frac{1}{N} \Sigma (x_i - \bar{x})^2 \cdot f_i$

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

$= \frac{\Sigma x_i^2 f_i}{N} - 2\bar{x} \frac{\Sigma x_i f_i}{N} + \bar{x}^2 \frac{\Sigma f_i}{N}$

$= \frac{1}{N} \Sigma x_i^2 f_i - 2\bar{x}^2 + \bar{x}^2 \quad [\because \Sigma f_i = N]$

$= \frac{1}{N} \Sigma x_i^2 f_i - \bar{x}^2$

$\bar{x} = \frac{1 \times 1 + 2 \times 2 + \cdots + n \times n}{1 + 2 + \cdots + n} = \frac{\cancel{n}\cancel{(n+1)}(2n+1)}{\cancel{6}\,3} \times \frac{\cancel{2}}{\cancel{n}\cancel{(n+1)}}$

$= \frac{2n+1}{3}$

$\Sigma x_i^2 f_i = 1^2 \times 1 + 2^2 \times 2 + \cdots + n^2 \times n$

$= \left\{ \frac{n(n+1)}{2} \right\}^2 = \frac{n^2 (n+1)^2}{4}$

$\boxed{\begin{array}{l} 4n^2+4n+1 \\ n^2+2n-n-2 \\ (n+2)(n-1) \end{array}}$

$N = \Sigma f_i = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$

$\therefore \sigma^2 = \frac{\cancel{2}}{\cancel{n}\cancel{(n+1)}} \times \frac{n^{\cancel{2}}(n+1)^{\cancel{2}}}{\cancel{4}\,2} - \frac{(2n+1)^2}{9}$

$= \frac{9n(n+1) - 2(2n+1)^2}{18} = \frac{9n^2 + 9n - 8n^2 - 8n - 2}{18}$

$= \frac{n^2 + n - 2}{18} = \frac{(n+2)(n-1)}{18}$

$\therefore$ S.d, $\sigma = \frac{\sqrt{(n+2)(n-1)}}{3\sqrt{2}}$