Name: Aniruddh Kulkarni

Roll No: I081

Subject: Statistical methods

Tutorial: Tutorial 5

1) From the following table showing the wage distribution in a certain factory, determine

a) the mean wage, the median wage, the modal wage

b) the percentages of workers who earned betweeen Rs 75 and Rs 125

c) the percentage who earned more than Rs 150 per week and

d) the percentage who earned less than Rs 100 per week

Weekly wages (Rs)	No of employees	Weekly wages	No of employees
20 - 40	8	120 - 140	35
40 - 60	12	140 - 160	18
60 - 80	20	160 - 180	7
80 - 100	30	180 - 200	5
100 - 120	40		

Solution :

Weekly wages (Rs)	f_i	x_i	$x_i f_i$	$c.f_i$	$\lvert x_i - \bar{x} \rvert$	$f_i \lvert x_i - \bar{x} \rvert$	$\lvert x_i - \bar{x} \rvert^2$
20 – 40	8	30	240	8	77.02	616.16	5,932.08
40 – 60	12	50	600	20	57.02	684.24	3251.28
60 – 80	20	70	1400	40	37.02	740.04	1370.48
80 – 100	30	90	2700	70	17.02	510.6	289.68
100 – 120	40	110	4400	110	2.98	119.2	8.88
120 – 140	35	130	4550	145	22.98	804.3	528.08
140 – 160	18	150	2700	163	42.98	773.64	1847.28
160 – 180	7	170	1190	170	62.98	440.86	3966.48
180 – 200	5	190	950	175	82.98	414.9	6,885.68
	$\Sigma f_i = 175$		$\Sigma f_i x_i = 18,730$				

Mean wage $(\bar{x}) = \dfrac{\Sigma f_i x_i}{\Sigma f_i}$

$= \dfrac{18,730}{175}$

$= 107.02$

Mode = ?

Modal class = 100 – 120

$l = 100$

$h = \cancel{100}\ 20$

$f_k = 40$

$f_{k-1} = 30$

$f_{k+1} = 35$

$$\text{Mode} = l + h\left[\frac{f_k - f_{k-1}}{2f_k - f_{k-1} - f_{k+1}}\right]$$

$$= 100 + 20\left[\frac{40-30}{2(40)-30-35}\right]$$

$$= 100 + 20\left[\frac{\cancel{10}\,2}{\cancel{15}\,3}\right]$$

$$= 100 + 20\left[\frac{2}{3}\right]$$

$$= 100 + \frac{40}{3}$$

$$= \frac{300+40}{3}$$

$$= \frac{340}{3}$$

$$= 113.33$$

Mode $= 113.33$

Median = ?

$N = \Sigma f_i$

$= 175$

For odd no of observations

Median $= \left(\frac{n+1}{2}\right)^{th}$ observation

$= \left(\frac{N+1}{2}\right)^{th}$ observation

$$= \left(\frac{175+1}{2}\right)^{th} \text{ observation}$$

$$= \left(\frac{176}{2}\right)^{th} \text{ observation}$$

$$= 88^{th} \text{ observation}$$

So, the 88^{th} observation will lie in the 100 - 120 class

Hence, Modal class = 100 - 120

$l = 100$
$f = 40$
$n = 175$
$h = 20$
$c.f = 70$

$$\text{Median} = l + \left[\frac{\frac{n}{2} - cf}{f}\right] \times h$$

$$= 100 + \left[\frac{\frac{175}{2} - 70}{40}\right] \times 20$$

$$= 100 + \left[\frac{35}{2} \times \frac{1}{\cancel{40}\,2}\right] \times \cancel{20}$$

$$\text{Median} = 100 + \frac{35}{4}$$

$$= \frac{400 + 35}{4}$$

$$= 108.75$$

b) the percentage of workers who earned between Rs 75 and Rs 125

Sol:

$$\left[\left(\frac{80-75}{20}\right) \times 20\right] + 30 + 40 + \left[\frac{125-120}{20} \times 35\right]$$

$$= \frac{5 + 30 + 40 + 8.75}{1}$$

$$= 83.75$$

$$\% \text{ of persons} = \frac{83.75}{175} \times 100$$

$$= 47.85$$

c) > 150

$$\left[\left(\frac{160-150}{20}\right) \times 18\right] + 7 + 5$$

$9 + 7 + 5$

21

$\%$ of persons $= \frac{21}{175} \times 100$

$= 12\%$

d) < 100

$8 + 12 + 20 + 30$

70

$\%$ of persons $= \frac{70}{175} \times 100$

$= 40\%$

1) Q_3

$$Q_3 = 120 + \frac{(131.25) - 110}{35} \times 20$$

$$= 120 + 12.14$$

$$= 132.14$$

2) D_6

$$D_6 = 100 + \left(\frac{105 - 70}{40}\right) 20$$

$$= 100 + 17.5$$

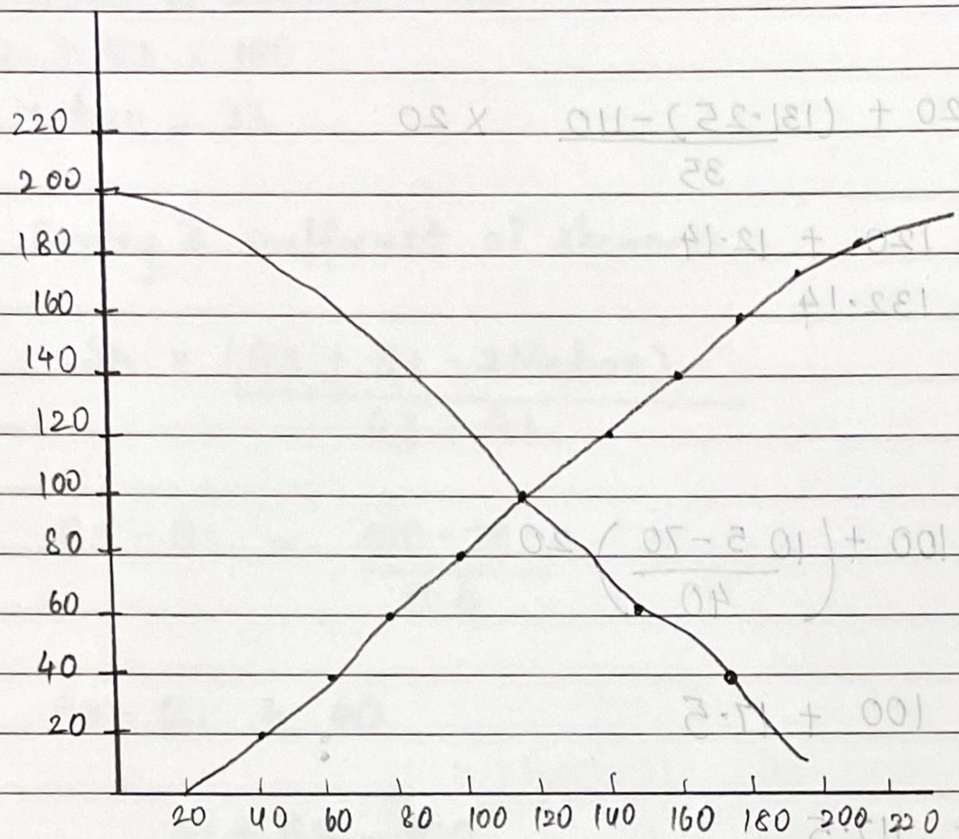
$$= 117.5$$

3) P_{66}

$$P_{66} = 120 + \frac{20(115.5 - 110)}{35}$$

$$= 120 + 3.14$$

$$= 123.14$$


220
200
180
160
140
120
100
80
60
40
20
20
40
60
80
100
120
140
160
180
200
220

2) Assume that a firm selected a random sample of 100 from its prodiction line and has obtained the data, shown in the table

C.I	f_i	x_i	$f_i x_i$	cf_i	$\lvert x_i - \bar{x} \rvert$	$f_i \lvert x_i - \bar{x} \rvert$	$\lvert x_i - \bar{x} \rvert^2$
130-134	3	132	396	3	~~0.7~~ 15.2	45.6	231.04
135-139	12	137	1644	15	10.2	122.4	104.4
140-144	21	142	~~2486~~ 2982	36	5.2	109.2	27.04
145-149	28	147	4116	64	0.2	5.6	0.04
150-154	19	152	2888	83	4.8	91.2	23.04
155-159	12	157	1884	95	9.8	117.6	96.04
160-164	5	162	810	100	14.8	74	219.04
	$\Sigma f_i = 100$		$\Sigma f_i x_i = 14720$				

Compute

a) arithmetic mean

$$\text{Mean } (\bar{x}) = \frac{\Sigma f_i x_i}{\Sigma f_i}$$

$$= \frac{14720}{100}$$

$$= 147.2$$

$f_i \lvert x_i - \bar{x} \rvert^2$

693.12
1252.8
567.84
1.12
437.76
1152.48
1095.2

4105.12

$\Sigma f_i \lvert x_i - \bar{x} \rvert^2 = \cancel{4689.264}$

b) $$\text{Variance} = \frac{\Sigma f_i \lvert x_i - \bar{x} \rvert^2}{\Sigma f_i}$$

$$= \cancel{\frac{4689.264}{100}} \quad \frac{4105.12}{100}$$

$$= \cancel{46.89} \quad 41.05$$

standard deviation = $\sqrt{Var}$

$= \cancel{\sqrt{46.89}}\ \sqrt{41.05}$

$= \cancel{6.84}\ 6.40$

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

$= 139.57 \left(\frac{5}{21}\right) (25-15)$

$= 140 + 2.38 - 0.5$

$= 141.88$

$Q_2 = 145 + \frac{5}{28} (50 - 36) - 0.5$

$= 145 + 2.5 - 0.5$

$= 147$

d) Measure of skewness $= M - Mo$

$= 147.2 - 146.6875$

$= 0.5125$

e) Coefficient of skewness $= \dfrac{M - Mo}{SD}$

$= \dfrac{0.5125}{6.40}$

$= 0.08$

Bowley's Coefficient of Skewness

$$\text{Median} = L + h \frac{\left[\frac{N}{2} - cf\right]}{f}$$

$$= 145 + 4 \frac{[60 - 36]}{28}$$

$$= 145 + 2$$

$$= 147$$

$N = 100$

$$Sk = \frac{Q3 + Q1 - 2\text{Median}}{Q3 - Q1}$$

$$= \frac{152.312 + 141.905 - 2(147)}{152.312 - 141.905}$$

$$= \frac{294.217 - 294}{10.407}$$

$$= 0.0208$$

Q3. Coefficient of skewness $= 0.6$

$Q_1 + Q_3 = 100$

Median $= 33$

By Bowley's coefficient of skewness

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

$$Q_3 - Q_1 = \frac{100 - 76}{0.6}$$

$$Q_3 - Q_1 = 40$$

$$\begin{aligned} Q_1 + Q_3 &= 100 \\ + \quad -Q_1 + Q_3 &= 40 \\ \hline 2Q_3 &= 140 \\ Q_3 &= 70 \end{aligned}$$

$$Q_1 = 100 - 70 = 30$$

Upper quartile $= \cancel{70}\ 30$

Lower quartile $= 70$

Q4. $CV = ?$

Mean = 120

Mode = 123

Karl Pearson's Coeff of skewness = −0.3

$$\sigma = \frac{M - Mo}{Sk}$$

$$= \frac{120 - 123}{-0.3}$$

$$= 10$$

$$CV = \frac{\sigma}{x} \times 100$$

$$= \cancel{\frac{10}{-0.3} \times 100} \quad \frac{10}{120} \times 100$$

$$= 8.33$$

Q5.

Morning Hours

In this scenario skewness is negative indicates longer tail towards left

Kurtosis is α_3 (P_2) this means flatter curve data is more dispersed

Evening Hours

In this scenario skewness is positive indicates longer tail towards right

Kurtosis is α_3 (P_3) means flatter curve data is less dispersed as compared to morning hours

iii) σ is more meaning data, it is more spread across the mean

iii) σ is lesser than mornmg data, is more clustered