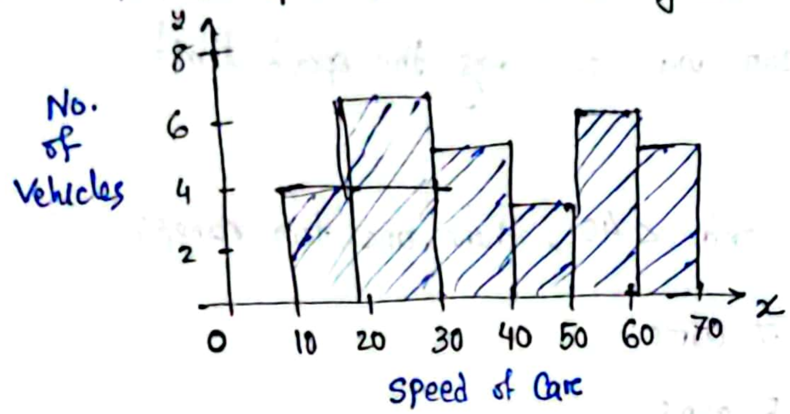
Problems on Statistics

① A Traffic police records the speeds of vehicles on a busy road with a 40 KMPH speed limit. The histogram below-



a) Estimate the average speed of the vechicles

b) Calculate the number of vehicles that were exceeding the speed limit by at least 10 Kmph

c) Calculate the number of vehicles having speed >= 20 Kmph but less than speed limit.

a) The average value would be Σ mean of every range * Num of Vehicles / Num of vehicle

$$= \frac{(4\times15)+(7\times25)+(5\times35)+(3\times45)+(6\times55)+(5\times65)}{(4+7+5+3+6+5)}$$

Formula $\left(\frac{\Sigma \text{ mean of range } * \text{ Num of vehicles}}{\text{Num of vehicles}}\right)$

$$= \frac{1200}{30}$$

$= 40$ Kmph.

b) There are two ranges that are exceeding 40 Kmph rang.

40 - 50 → have 6 cars

50 - 60 → have 5 cars

So, ~~tat~~ total 11 car were exceeding the speed limit.

c) having speed >=20 and < 40, there are two ranges

20 - 30 → 7 cars

30 - 40 → 5 cars

So, Total 7+5=12 vehicles.

② The mean of 25 observation is 36. The mean of first 13 observations is 32 and that last 13 observations is 39. What is the value of the 13th observation?

Ans:

$$\underbrace{\overbrace{t_1, t_2, t_3, \ldots t_{13}}^{32}, \underbrace{\ldots\ldots t_{24}, t_{25}}_{39}}_{36}$$

$$\sum_{i=1}^{25} ti = 36, \quad \sum_{i=1}^{13} ti = 32, \quad \sum_{i=13}^{25} ti = 39$$

$$\rightarrow \sum_{i=1}^{25} ti = 25 \times 36 = 900$$

$$\rightarrow \sum_{i=1}^{13} ti = 13 \times 32 = 416$$

$$\rightarrow \sum_{i=13}^{25} = 13 \times 39 = 507$$

Now, $\sum_{i=1}^{25} t_i + \sum_{i=13}^{25} t_i = 416+507 = 923$

$\sum_{i=1}^{25} t_i + t_{13} = 923$

$\Rightarrow t_{13} = 923 - \sum_{i=1}^{25} t_i$

$= 923 - 900$

$= 23$ Ans

(3) The following frequency table gives the values obtained in 30 rolls of a die.

Value	Freq
1	4
2	6
3	7
4	5
5	3
6	5

Find → a) sample mode b) Sample variance c) sample std.

a) Mode → 3 (Because it has max freq of 7)

b) Sample variance → $\sum f_i x_i$ ($f_i \to$ ~~value~~ freq, $x_i \to$ ~~freq~~ val)

4
12
21
20
15
30

= 102

$\bar{x} = \frac{\sum f_i x_i}{n}$

$= \frac{102}{30}$

$= 3.4$

$$\text{variance} = \frac{\sum_{i=1}^{30} (x_i - \bar{x})^2}{n-1} \quad \left[\text{For sample variance, } n-1\right]$$

$$= \frac{4(1-3.4)^2 + 6(2-3.4)^2 + 7(3-3.4)^2 + 5(4-3.4)^2 + 3(5-3.4)^2 + 5(5-3.4)^2}{30-1}$$

$= 2.731$ Ans

c) Standard deviation $= \sqrt{\text{variance}(s^2)}$

$= \sqrt{2.731}$

(4) Rayan just took his final math ~~class~~ test in his college analysis class. His prof said that he has scored 80th percentile in the class. Rayan's professor posts a list of grades, without the names on the background. There are 12 students in the class and 12 marks on the board. Grades are →

46, 85, 68, 93, 84, 70, 38, 66, 78, 75, 55, 60. What is Rayan's Grade.

Ans: 80th percentile of 12 $= \frac{80}{100} \times 12 = 9.6$

So, if we ordered the grades, he would the 10th guy.

His score is 84.

⑤ Preeti found the following ages of 8 tigers. Those tigers are randomely selected from the 20 tigers ~~from~~ at her local zoo:

5, 9, 13, 15, 17, 3, 5, 1

What is the value of std for these 8 tigers?

$$\bar{x} = \frac{5+9+13+15+17+3+5+1}{8}$$

$$= 8.5$$

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

$$= \frac{(3-8.5)^2 + (9-8.5)^2 + (13-8.5)^2 + (15-8.5)^2 + (17-8.5)^2 + (3-8.5)^2 + (5-8.5)^2 + (1-8.5)^2}{8-1}$$

$$= 35.743$$

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

$= \sqrt{35.143} = 5.928$ years.

⑥ If sample variance of a data of size 10 is 23, than what is the population variance of the data.

Sample, variance, $s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = 23$ | $s^2 = 23$, $n = 10$

$\Rightarrow s^2 \Rightarrow \sum (x_i - \bar{x})^2 = 23(10-1)$
$= 207$

Now, population variance, $\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$

$= \frac{207}{10} = 20.7$ (population Variance)

⑦ For a particular data, the value of the 10th percentile is 33.5, 25th percentile is 45, 50th percentile is 84.5 and 100th percentile is 45. What is the median of the data?

Median = 50th percentile value = 84.5.

⑧ If each value of a numerical discrete value is squared, then the mean of new data is?

ans: Cannot be determined because, we can't determine the new data mean by analysisg the old data mean.

(9) Measures of Dispersion → ~~1) Mean 2) Median 3) Variance 3) Std~~

1) Mean ✗
2) Variance ✓
3) Std Deviation ✓
4) Range ✓

(10) The five-number summary of 99 observations of a numerical variable is 25, 35, 47, 56, 78. Based on the information →

1st observation → min → 25
25th observation Q_1 → 35
50th observation → Q_2 → 47 (Median)
75 ~~76~~th observation Q_3 → 56
99th observation → Max → 78

Median would be → $\frac{99+1}{2}$ = 50th observation
= 47

P × n
P = 25% = 0.25 × 99 = 24.75 = 25 th observation
P × n
P = 50% = 0.5 × 99 = 49.5 = ~~50~~ ~~49~~ 50 th observation
P × n
P = 75 = 0.75 × 99 = 74.25 = 74 th observation

[We will take the next integer of the float value]

(11) IQR value = $(Q_3 - Q_1)$

(12) From the options, choose the outliers, if any, for the following dataset: 8.5, 8.8, 11.4, 11.5, 11.6, 11.8, 12.0, 12.1, 12.2, 12.3, 12.4, 12.8, 14.8, 15.0

Here, $n = 15$

(Q_1) ∴ 25th ~~perc~~ percentile $= 0.25 \times 15 = 3.75 \rightarrow$ 4th place $= 11.5$

(Q_3) ∴ 75th percentile $= 0.75 \times 15 = 11.25 \rightarrow$ 12th place $= 12.4$

$$\therefore IQR = (Q_3 - Q_1) = (12.4 - 11.5)$$

$$= 0.9$$

These values would be outliers → (i) $Q_1 - 1.5\,IQR$ and (ii) $Q_3 + 1.5\,IQR$

(i) $Q_1 - 1.5\,IQR$

$= 11.5 - 1.5 \times 0.9$

$= 10.15$

(ii) $Q_3 + 1.5\,IQR$

$= 12.4 + 1.5 \times 0.9$

$= 13.75$

So, values < 10.15 and $>$ ~~13~~ 13.75 are outliers

They are → 8.5, 8.8, 14.8, 15.0