Statistics

Three types of data representation:

1. Ungrouped (T1)
   Example: 2, 14, 17

2. Grouped with single class value (T2)
   Example: 2*10, 5*15, 7*20
   Explanation: This means the element appears with its frequency.
   (e.g., 2*10 means 2 appears 10 times)

3. Grouped with range class value (T3)
   Example: (11-20)*10, (21-30)*15, (31-40)*7
   Explanation: Values between 11-20 appear 10 times, and so on.

Key Concepts:

1. Average:
   Formula: (x1 + x2 + x3 + ... + xn) / n
   Meaning: Represents the central value

2. Mode:
   Definition: The element that occurred the most number of times

3. Median:
   - Removes effect of outliers
   - For odd n: (n + 1) / 2-th value
   - For even n: Average of (n/2)-th and (n/2 + 1)-th values
   Note: Data must be sorted before calculating median

4. Dispersion / Scatterness:
   Represents how spread out the data is
   Range =maxVal-minVal
   1.Mean Deviation meaning deviation from mean
   If x is mean and x1 x2 x3 … xn are elements then
   Mean deviation is sum of all |x-xi| divided by n xi is the elements
   2.Variance and Standard Deviation
   standardDeviation =sqrt(variance)

variance=sum of (xi-x)^2/n
Standard deviation =sqrt(variance)
For T2 data and T3
variance=sum freq*(xi-x)^2/(sum of frequency)
In case of T3 xi is average of low and high value

5. Finding Average for T3:

Given:
(0-9)*2, (10-19)*10, (20-29)*12, (30-39)*8

Step:
- Convert each class range to its midpoint
  (e.g., (0-9) → (0+9)/2 = 4.5)
- Now treat it like T2:
  4.5*2, 14.5*10, 24.5*12, 34.5*8

Final Average Formula:
(x1*f1 + x2*f2 + ... + xn*fn) / (f1 + f2 + ... + fn)

6. Cumulative Frequency:
- Just the prefix sum of frequencies
- Useful for median calculation

7. Calculating Median in T3 Data:

Step-by-step:
- Find total number of values (n)
- Compute n/2 (if even) or (n+1)/2 (if odd)
- Find the class interval that contains this position
- Use the formula:

Median = L + [(n/2 − F) / f] × h

Where:
- L = lower boundary of the median class
- F = cumulative frequency before median class
- f = frequency of the median class
- h = class width

8. Calculating Mode for T3 Data:

Mode = L + [(f1 − f0) / (2f1 − f0 − f2)] × h

Where:
- L = lower limit of modal class
- f1 = frequency of modal class
- f0 = frequency of previous class
- f2 = frequency of next class
- h = class width

9.Coefficient of Dispersion

1.coeff of range=(max-min)/(max+min)
2.coeff of standard deviation=variance/mean;
3.coeff of variation= variance/mean*100;

10.Miscellaneous
Sum of all elements subtracted by mean is 0
If all elements are multiplied by y then mean becomes y*mean,variance=y^2*variance,standard deviation=y*standard deviation

If we add y to all ele the mean will increase with y like y+mean;

=>Mode=3*median-2*mean;
(xi-y)^2 will be minimum if y is mean;