Elements Of Statistics
Class-BCA III Semester

Dr. Vijay Kant Sharma

Assistant professor ,

Department Of Computer Application

Jagatpur P.G. College, Varanasi

Affiliated to Mahatma Gandhi Kashi vidhyapith Varanasi

Email-mzp.vijay@gmail.com

OUTLINE:-

UNIT-III

Measure of dispersion:-

Concept of measure of dispersion

Range

Semi Inter-quartile Range

Variance

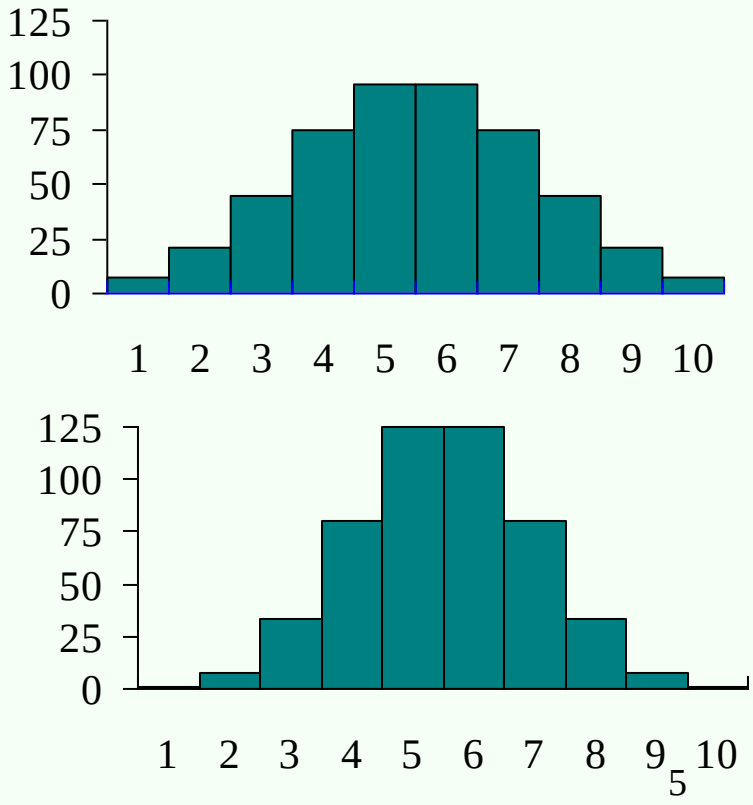
Measures of Dispersion

Definition

- *Measures of dispersion* are descriptive statistics that describe how similar a set of scores are to each other
  - The more similar the scores are to each other, the lower the measure of dispersion will be
  - The less similar the scores are to each other, the higher the measure of dispersion will be
  - In general, the more spread out a distribution is, the larger the measure of dispersion will be

Measures of Dispersion

- Which of the distributions of scores has the larger dispersion?
- The upper distribution has more dispersion because the scores are more spread out
  - That is, they are less similar to each other



Measures of Dispersion

- There are three main measures of dispersion:
  - The range
  - The semi-interquartile range (SIR)
  - Variance / standard deviation

The Range

- The *range* is defined as the difference between the largest score in the set of data and the smallest score in the set of data, $X_L - X_S$
- What is the range of the following data:
  4 8 1 6 6 2 9 3 6 9
- The largest score (X_L) is 9; the smallest score (X_S) is 1; the range is $X_L - X_S = 9 - 1 = 8$

When To Use the Range

- The range is used when
  - you have ordinal data or
  - you are presenting your results to people with little or no knowledge of statistics
- The range is rarely used in scientific work as it is fairly insensitive
  - It depends on only two scores in the set of data, X_L and X_S
  - Two very different sets of data can have the same range:
    1 1 1 1 9 vs 1 3 5 7 9

The Semi-Interquartile Range

- The *semi-interquartile range* (or *SIR*) is defined as the difference of the first and third quartiles divided by two
  - The first quartile is the 25^{th} percentile
  - The third quartile is the 75^{th} percentile
- $SIR = (Q_3 - Q_1) / 2$

SIR Example

- What is the SIR for the data to the right?
- 25 % of the scores are below 5
  - 5 is the first quartile
- 25 % of the scores are above 25
  - 25 is the third quartile
- SIR = $(Q_3 - Q_1) / 2 = (25 - 5) / 2 = 10$

2	
4	
6	← 5 = 25th %tile
8	
10	
12	
14	
20	
30	← 25 = 75th %tile
60	

When To Use the SIR

- The SIR is often used with skewed data as it is insensitive to the extreme scores

Variance

- *Variance* is defined as the average of the square deviations:

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N}$$

What Does the Variance Formula Mean?

- First, it says to subtract the mean from each of the scores
  - This difference is called a *deviate* or a *deviation score*
  - The deviate tells us how far a given score is from the typical, or average, score
  - Thus, the deviate is a measure of dispersion for a given score

What Does the Variance Formula Mean?

- Why can't we simply take the average of the deviates? That is, why isn't variance defined as:

$$\sigma^2 \neq \frac{\sum (X - \mu)}{N}$$

This is not the formula for variance!

What Does the Variance Formula Mean?

- One of the definitions of the *mean* was that it always made the sum of the scores minus the mean equal to 0
- Thus, the average of the deviates must be 0 since the sum of the deviates must equal 0
- To avoid this problem, statisticians square the deviate score prior to averaging them
  - Squaring the deviate score makes all the squared scores positive

What Does the Variance Formula Mean?

- Variance is the mean of the squared deviation scores
- The larger the variance is, the more the scores deviate, on average, away from the mean
- The smaller the variance is, the less the scores deviate, on average, from the mean

Questions:-

- What is the measure of dispersion explain?
- What is range and semi inter-quartile range?
- Define variance and elaborate it?

Reference Books:

1. S.C.Gupta - Fundamentals of Statistics - Sultan Chand & sons , Delhi.
2. D.N.Elhance - Fundamentals of statistics - Kitab Mahal, Allahabad.
3. Montgomery D.C. – Statistical Quality Control - John Wiley and Sons
4. Goon, Gupta And Dasgupta - Fundamentals of Statistics - The world press private ltd. , Kolkata.
5. Hogg R.V. and Craig R.G. – Introduction to mathematical statistics Ed 4 {1989} – Macmillan Pub. Co. New York.
6. Gupta S.P. – Statistical Methods, Sultan Chand and Sons New Delhi

Declaration

The content is exclusively meant for academic purpose and for enhancement teaching and learning.Any other use of economic/commercial purpose is strictly prohibited. The users of the content shall not distribute ,disseminate or share it with anyone else and its use is restricted to advancement of individual knowledge.The information provided in this e-content is authentic and best as per my knowledge.

Dr.Vijay kant Sharma

Assistant professor

Department of computer Application

jagatpur P.G. College Varanasi Affiliated to Mahatma Gandhi Kashi Vidyapith Varanasi

Thanks